2022

Data Privacy, Human Rights, and Algorithmic Opacity

Sylvia Lu

Data Privacy, Human Rights, and Algorithmic Opacity

Sylvia Lu*

Decades ago, it was difficult to imagine a reality in which artificial intelligence (AI) could penetrate every corner of our lives to monitor our innermost selves for commercial interests. Within just a few decades, the private sector has seen a wild proliferation of AI systems, many of which are more powerful and penetrating than anticipated. In many cases, AI systems have become "the power behind the throne," tracking user activities and making fateful decisions through predictive analysis of personal information. Despite the growing power of AI, proprietary algorithmic systems can be technically complex, legally claimed as trade secrets, and managerially invisible to outsiders, creating an opacity that hinders oversight of AI systems. Accordingly, many AI-based services and products have been found to be invasive, manipulative, and biased, eroding data privacy rules, human rights, and democratic norms in modern society.

The emergence of AI systems has thus generated a deep tension between algorithmic secrecy and data privacy. Yet, in today's policy debate, algorithmic transparency in a privacy context is an equally important issue that is nonetheless managerially disregarded, commercially evasive, and legally unactualized. This Note is the first to illustrate how regulators should rethink strategies regarding data privacy through the interplay of human rights, algorithmic disclosures, and whistleblowing systems. As the world increasingly

DOI: https://doi.org/10.15779/Z38804XM07

* Doctoral Candidate and Lloyd M. Robbins Fellow at the University of California, Berkeley School of Law. This piece won the Aldo J. Test Prize for Best Berkeley Submission in the 2021 *Berkeley Technology Law Journal* Writing Competition. The author is deeply grateful to Sonia K. Katyal, Laurent Mayali, and Paul M. Schwartz for their inspiration and insightful feedback on prior drafts. For stimulating discussions and comments, thanks to Kenneth Bamberger, Elizabeth Edenberg, Chris Hoofnagle, Daniel Jellins, Emily McReynolds, Chinmayi Sharma, Robert Sloan, Jeffrey Vagle, and the participants at the 2022 Privacy Law Scholars Conference (PLSC). The author wishes to thank Andrew Selbst and Gianclaudio Malgieri, whose commentaries at the 2022 PLSC were incredibly helpful. Finally, the author owes thanks to the thoughtful editors of the *California Law Review* for their support and excellent suggestions.

looks to the European Union's (EU) data protection law—the General Data Protection Regulation (GDPR)—as a regulatory frame of reference, this piece assesses the effectiveness of the GDPR's response to data protection issues raised by opaque AI systems. Based on a case study of Google's AI applications and privacy disclosures, this piece demonstrates that even the EU fails to enforce data protection rules to address issues caused by algorithmic opacity.

This Note argues that as algorithmic opacity has become a primary barrier to oversight and enforcement, regulators in the EU, the United States, and elsewhere should not overprotect the secrecy of every aspect of AI applications that implicate public concerns. Rather, policymakers should consider imposing a duty of algorithmic disclosures through sustainability reporting and whistleblower protection on firms deploying AI to maximize effective enforcement of data privacy laws, human rights, and other democratic values.

INTRODUCTION

Decades ago, it was difficult to imagine a reality in which artificial intelligence (AI) could penetrate every corner of our private lives to track, commodify, and trade our inner selves for commercial interests.[1] Within just a few decades, the private sector has seen a wild proliferation of AI systems, many of which are more powerful and penetrating than anticipated.[2] As a radically

1. *See* William Magnuson, *Artificial Financial Intelligence*, 10 HARV. BUS. L. REV. 337 (2020) (discussing the impact of artificial intelligence on the field of finance and beyond).

2. *See* Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES (Feb. 16, 2012), https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html [https://perma.cc/QEN7-FU4U]; James Maguire, *Top Performing Artificial Intelligence Companies of 2022*, DATAMATION

disruptive innovation, AI enables firms to complete complex and time-consuming tasks at record speeds. These new capabilities, however, come with AI's overwhelming potential for omnipresent but covert surveillance.[3] With this potential, AI systems can collect, process, and transfer private information that we would not otherwise disclose.[4] For instance, facial recognition technologies have been used to capture large quantities of biometric data from numerous user photos and videos.[5] Tracking apps have been secretly tracking user behavior to collect a large volume of location information.[6] Social media algorithms have been developed to correlate one's language usage with sensitive personal data, including political views, religious beliefs, trade union membership, and more.[7]

In many cases, AI systems have become the power behind the throne—they lurk in the background, yet make crucial decisions through predictive analysis of personal data.[8] Firms have used AI to decide what should be seen in online search results or even who should be given employment opportunities, and they do so at the cost of data privacy.[9] For instance, Google has deployed algorithms to track user behavior and change search results.[10] Hospital systems and insurance companies have used medical algorithms to determine individual

(Apr. 9, 2021), https://www.datamation.com/artificial-intelligence/top-artificial-intelligence-companies.html [https://perma.cc/8GEU-6X8M].

3. *See* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 21–30 (2015) (discussing the impacts of advanced algorithms on society).

4. Karl Manheim & Lyric Kaplan, *Artificial Intelligence: Risks to Privacy and Democracy*, 21 YALE J.L. & TECH. 106, 106 (2019).

5. *Id.* at 122–23. The Internet of Things (IoT) is a form of digital devices based on sensors, software, and other technologies that transfer data to other electronic devices over a network.

6. Iman Ghosh, *AIoT: When Artificial Intelligence Meets the Internet of Things*, VISUAL CAPITALIST (Aug. 12, 2020), https://www.visualcapitalist.com/aiot-when-ai-meets-iot-technology/ [https://perma.cc/TN57-2FMW].

7. *See, e.g.*, Eliza Mackintosh, *Facebook Knew It Was Being Used to Incite Violence in Ethiopia. It Did Little to Stop the Spread, Documents Show*, CNN (Oct. 25, 2021), https://www.cnn.com/2021/10/25/business/ethiopia-violence-facebook-papers-cmd-intl/index.html [https://perma.cc/R495-PQYE]; Josh Constine, *Facebook's New DeepText AI Categorizes Everything You Write*, TECHCRUNCH (June 1, 2016), https://techcrunch.com/2016/06/01/facebook-deep-text/ [https://perma.cc/7838-V8EN].

8. *See* Magnuson, *supra* note 1, at 337.

9. *See How Results Are Automatically Generated*, GOOGLE (last visited Aug. 31, 2022), https://www.google.com/intl/en_uk/search/howsearchworks/how-search-works/ranking-results/ [https://perma.cc/LGK9-Z4AY] ("Information such as your location, past Search history and Search settings all help us to ensure that your results are what is most useful and relevant for you in that moment. . . . Search also includes some features that personalise results based on the activity in your Google Account."); Nathan R. Kuncel, Deniz S. Ones & David M. Klieger, *In Hiring, Algorithms Beat Instinct*, HARV. BUS. REV. (May 2014), https://hbr.org/2014/05/in-hiring-algorithms-beat-instinct [https://perma.cc/MQ85-S7JV].

10. Kirsten Grind, Sam Schechner, Robert McMillan & John West, *How Google Interferes with Its Search Algorithms and Changes Your Results*, WALL ST. J. (Nov. 15, 2019), https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753 [https://perma.cc/QCD7-3CGL]; Connor Finnegan, *How Facebook and Google Track Your Online Behavior*, MEDIUM (Feb. 13, 2019), https://medium.com/@ConnorFinnegan/how-facebook-and-google-track-your-online-behavior-26f161d370ab [https://perma.cc/74MR-EXAX].

health risks and eligibility to receive medical treatment.[11] Politicians and interest groups have leveraged social media algorithms to spread manipulated extremist speech to users whose posts on social media platforms implied similar political views.[12] Without adequate oversight, such algorithmic decision-making processes enable firms to influence consumer behavior, exacerbate gross social disparities, and even worsen unjust treatment.[13]

The use of AI in turn not only raises data privacy issues, but also has a significant impact on the public domain, in addition to a lack of accountability and legitimacy in corporate AI deployment.[14] As firms have used AI to penetrate private lives, individual control over AI has been reduced under opaque conditions, leading to a number of broader social implications. On the one hand, the implications of AI can be as imperceptible as covertly collecting personal data without consent, secretly manipulating individuals into buying a product, or spreading extreme information to targeted users. On the other hand, opaque AI systems can also be life-changing when they are used to decide who should receive medical treatment, who should get a job, and who should remain incarcerated.[15] Accordingly, opaque AI business applications are invasive,

---

11. Starre Vartan, *Racial Bias Found in a Major Health Care Risk Algorithm*, SCI. AM. (Oct. 24, 2019), https://www.scientificamerican.com/article/racial-bias-found-in-a-major-health-care-risk-algorithm/ [https://perma.cc/YU8B-UDFA] (describing how AI rates health risks according to predicted costs based on historically biased health records); Chris Poulin, Brian Shiner, Paul Thompson, Linas Vepstas, Yinong Young-Xu, Benjamin Goertzel, Bradley Watts, Laura Flashman & Thomas McAllister, *Predicting the Risk of Suicide by Analyzing the Text of Clinical Notes*, 9 PLOS ONE 1 (Jan. 28, 2014).

12. *See, e.g.*, Gilad Edelman, *How Facebook's Political Ad System Is Designed to Polarize*, WIRED (Dec. 13, 2019), https://www.wired.com/story/facebook-political-ad-system-designed-polarize/ [https://perma.cc/KH3B-SNHQ]; Nick Corasaniti, *Political Campaigns Can Still Target You on Facebook*, N.Y. TIMES (Nov. 11, 2021), https://www.nytimes.com/2021/11/11/us/politics/facebook-political-ads.html [https://perma.cc/FGM9-MRDL].

13. *See generally* CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2016) (discussing how companies have used unregulated advanced algorithms to perpetuate inequality through algorithmic decision-making about citizens' access to loans, employment, health, and more).

14. *See* 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, 633 (2017); Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement*, 69 FLA. L. REV. 181, 181 (2018).

15. *See* Charlotte Jee, *A Biased Medical Algorithm Favored White People for Health-Care Programs*, MIT TECH. REV. (Oct. 25, 2019), https://www.technologyreview.com/2019/10/25/132184/a-biased-medical-algorithm-favored-white-people-for-healthcare-programs/ [https://perma.cc/K6FK-VGBR]; *AI at Work: Staff 'Hired and Fired by Algorithm,'* BBC (Mar. 25, 2021), https://www.bbc.com/news/technology-56515827 [https://perma.cc/LDT6-8WNP]; Han-Wei Liu, Ching-Fu Lin & Yu-Jie-Chen, *Beyond* State v. Loomis*: Artificial Intelligence, Government Algorithmization and Accountability*, 27 INT'L J.L. & INFO. TECH. 122, 125–26 (2019) (discussing how automated systems have been applied to decide "policing, pretrial bail sentencing, and post-trial sentencing" and their far-reaching implications for the criminal justice system).

manipulative, instigative, biased, and unfair,[16] posing threats to data privacy and other democratic values like autonomy, fairness, and transparency.[17]

In addition to legal and regulatory challenges, algorithmic systems have produced unprecedented problems for corporate governance. Since the advent of big data, firms have tackled a multitude of issues concerning protection of data privacy and governance of AI innovation. Firms are directly confronting a changing landscape of data privacy legislation that requires them to adopt amended business models and privacy policies that respect individuals' rights. At the same time, firms may struggle to comply with complex data privacy rules that can disrupt business applications of AI and increase risk of violations.

In the absence of a scheme to monitor how a firm tackles data privacy issues, a given enterprise is likely to face increased legal proceedings, corresponding reputational harm, and reduced shareholder investment, all of which endanger the long-term operation of a business. Take the Facebook-Cambridge Analytica data scandal as an example. Due to insufficient data security, Cambridge Analytica's algorithms processed—without authorization—the data of up to 87 million people.[18] In the wake of this scandal, Cambridge Analytica—a British political consulting company—faced class-action stakeholders lawsuits, received administrative complaints with the U.S. Federal Trade Commission (FTC),[19] filed for bankruptcy liquidation,[20] and ultimately ceased their business operations.[21] Another tech giant, Google, has also faced ceaseless data privacy lawsuits over its AI-based covert user tracking,

16. *See* Aylin Caliskan, Joanna J. Bryson & Arvind Narayanan, *Semantics Derived Automatically from Language Corpora Contain Human-Like Biases*, 356 SCIENCE 183, 183 (2017); Moritz Hardt, *How Big Data Is Unfair*, MEDIUM (Sept. 26, 2014), https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de [https://perma.cc/U4XF-55V4].

17. *See* Anupam Chander, *The Racist Algorithm?*, 115 MICH. L. REV. 1023, 1044 (2017).

18. Paolo Zialcita, *Facebook Pays $643,000 Fine for Role in Cambridge Analytica Scandal*, NPR (Oct. 30, 2019), https://www.npr.org/2019/10/30/774749376/facebook-pays-643-000-fine-for-role-in-cambridge-analytica-scandal [https://perma.cc/VVS8-HCGB]; Carole Cadwalladr & Emma Graham-Harrison, *How Cambridge Analytica Turned Facebook 'Likes' into a Lucrative Political Tool*, GUARDIAN (Mar. 17, 2018), https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm [https://perma.cc/45S6-QFB7].

19. *FTC Sues Cambridge Analytica, Settles with Former CEO and App Developer*, FED. TRADE COMM'N (July 24, 2019), https://www.ftc.gov/news-events/press-releases/2019/07/ftc-sues-cambridge-analytica-settles-former-ceo-app-developer [https://perma.cc/PL42-RJWW] (explaining that the FTC filed a complaint against Cambridge Analytica for a breach of data privacy); *See generally* CHRIS JAY HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY 145–192 (2016) (introducing the FTC's role in regulating online data privacy in the US).

20. Nathan Bomey, *Cambridge Analytica Files for Chapter 7 Bankruptcy Liquidation After Debilitating Scandal*, USA TODAY (May 18, 2018), https://www.usatoday.com/story/money/2018/05/18/cambridge-analytica-bankruptcy/622303002/ [https://perma.cc/AF3D-N4UA].

21. Mike Snider, *Cambridge Analytica Shutting Down in Wake of Facebook Data Crisis*, USA TODAY (May 2, 2018), https://www.usatoday.com/story/tech/news/2018/05/02/cambridge-analytica-shutting-down-wake-facebook-data-crisis/573963002/ [https://perma.cc/3CTE-JJQV].

negative search results, and lack of transparency in targeted advertising.[22] As of May 15, 2022, several enforcement agencies have fined Google at least €205 million for its data protection violations.[23] These cases illustrate how both firms and citizens suffer from firms' failure to establish adequate privacy policies and data practices, given the emerging data privacy risks posed by the use of big data and AI.

Collectively and individually, the uncomfortable paradox of AI creates a tension between corporate data capitalism and civic data protection. As a deluge of data boosts an unprecedented AI economy, data privacy law and corporate governance both face enormous challenges due to the extraordinarily rapid development of AI. To prevent the erosion of data privacy through unlimited use of AI applications in the private sector, the EU has established a fundamental rights-based data protection standard under the General Data Protection Regulation (GDPR) that echoes the EU's longstanding ambition to develop a human-centered, ethical, and secure AI ecosystem.[24] Unlike the United States' sectoral approach to privacy protection that barely restrains firms from processing big data, the EU sees data protection as a human right in response to emerging AI threats to data privacy.[25] However, while the EU and its Member States strive to protect data privacy in general, invasive AI systems continue to operate in the shadow of algorithmic opacity—the opacity of AI that prevents stakeholder view due to their technical complexity, trade secrets protection, and managerial invisibility. Today, firms operate their AI in opaque conditions without sufficient stakeholder review, and consequently escape data privacy enforcement and other democratic norms. Since firms are not required to disclose their use of AI systematically, external stakeholders, such as individuals, advocacy groups, communities, shareholders, and researchers all lack comprehensive internal information on corporate use of AI. External stakeholders have no idea how firms use AI systems to process individuals' data, how firms protect individual rights through data management measures, and how

22. *GDPR Enforcement Tracker*, CMX L. TAX FUTURE (2021), https://www.enforcementtracker.com/ [https://perma.cc/C3KW-3FJV].

23. The French data protection authority fined Google Ireland €90 million and Google LLC €60 million on December 31, 2021; the Hungary data protection authority fined Google Ireland €28 million on July 16, 2020; the Belgian data protection authority fined Google Belgium €600,000 on July 14, 2020; the Swedish data protection authority fined Google LLC €5 million on March 11, 2020; France's data regulator fined Google LLC €50 million on January 21, 2019. *Id.*

24. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), 2016 O.J. (L. 119) 1–88 [hereinafter General Data Protection Regulation (GDPR)].

25. *See generally*, Paul M. Schwartz & Karl-Nikolaus Peifer, *Transatlantic Data Privacy*, 116 GEO. L.J. 115 (2017) (highlighting the EU's privacy culture and focus on protecting data subject rights in contrast to the United States' privacy framework); James Whitman, *The Two Western Cultures of Privacy: Dignity Versus Liberty* 113 YALE L.J. 1151, 1157 (2004) (describing the European "fundamental right to privacy" that is absent in the United States).

firms address major risks derived from AI applications. Although improper data management measures can lead to breaches, fines, and reputational harm, shareholders have no way to understand what anticipated legal and managerial challenges arise when using AI applications. Without such information, private AI systems often operate without external oversight until a scandal occurs, harming individuals, communities, shareholders, and firms without adequate warning. As of this writing, even the EU, which is at the forefront of both data privacy and AI regulations,[26] lacks a specific regulatory framework for algorithmic transparency that enables comprehensive stakeholder surveillance and effective enforcement of data privacy. Existing EU data protection rules applicable to AI applications, particularly in a corporate context, are also inadequate at preventing firms from sacrificing data privacy for commercial profit in opaque conditions.

This piece argues that as algorithmic opacity causes growing public concerns with respect to data privacy and other democratic values, mandated algorithmic disclosures should be pursued to remove the primary barrier to stakeholder oversight, safeguarding enforcement of data privacy and other fundamental values. In a previous Article, I discussed the legal issues caused by corporate use of AI in the U.S. regulatory context.[27] As the world increasingly looks to the EU as a regulatory reference in terms of data privacy, this piece further investigates the effectiveness of EU law to yield valuable insights for suggested regulatory actions. To provide an in-depth assessment of available regulatory tools, this piece evaluates the EU's fundamental rights-based approach to algorithmic opacity, as the EU adopts by far one of the most legally robust, widely recognized, and globally influential data privacy laws. The EU's approach rightfully serves as a reference point for regulators worldwide to examine the implications of algorithmic opacity for data privacy, human rights, and democratic norms.[28]

This piece proceeds in four parts. First, after an introduction to AI, Part I defines and contextualizes the concept of algorithmic opacity, which prevents stakeholder view due to AI's technical complexity, trade secrecy, and managerial invisibility. Part II discusses how AI systems erode data privacy, impede

26. Paul M. Schwartz, *Global Data Privacy: The EU Way*, 94 N.Y.U. L. REV. 771, 771 (2019) (describing the EU as "the world's privacy cop"). As of August 2022, there is a proposed Regulation on AI (the AI Act) undergoing review, *see* Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, 39 EUR. COMM'N (Apr. 21, 2021) [hereinafter Proposed EU AI Act].

27. *See generally* Sylvia Lu, *Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence*, 23 VAND. J. ENTER. TECH. L. 99, 108–110 (2021).

28. Schwartz, *supra* note 26, at 771 ("[T]he General Data Protection Regulation (GDPR), is now widely regarded as a privacy law not just for the EU, but for the world."). *But see generally* Anupam Chander, Margot E. Kaminski & William McGeveran, *Catalyzing Privacy Law*, 105 MINN. L. REV. 1733 (2021) (arguing that the California Consumer Privacy Act is catalyzing privacy norms across the United States).

stakeholder surveillance, and hinder enforcement of democratic values owing to their algorithmic opacity. Part III assesses the effectiveness of the EU's response to data privacy issues raised by opaque AI systems based on a case study of Google's privacy disclosures, demonstrating that the GDPR alone might be insufficient to reduce algorithmic opacity. Part IV proposes new transparency strategies through sustainability reporting and whistleblowing systems to strengthen the enforcement of data privacy, human rights, and democratic norms. Finally, it explores a broad array of policy implications and suggests critical policy moves.

Finally, with respect to the terminology used in this piece, the terms "privacy" and "data privacy" are used to refer generally to the area of "privacy and data protection." In Europe, the rights to privacy and data protection are both considered fundamental rights.[29] Data protection is a standard term originating that the notion of privacy that concerns dignity and the right to private life, including a right to autonomy and control over data.[30] When this piece uses "data protection," it refers to the protection of personal data in the GDPR regulatory context. When this paper discusses the concept of AI, "AI systems," "algorithmic systems," and "AI" are used interchangeably to refer to technology based on a set of algorithms that possess features of human intelligence, including cognition, rationality, prediction, and decision-making power.

I.
THE RISE OF ARTIFICIAL INTELLIGENCE AND ALGORITHMIC OPACITY

In today's modern society, the public has increasingly counted on private systems for everything that can be fulfilled by AI.[31] Among global leaders in AI, the EU leads the world in setting data protection and AI regulations.[32] Before exploring how AI systems pose threats to data protection in the EU, this section provides context on AI and algorithmic opacity. It explains the nature and application of AI, discusses the emergence of algorithmic opacity, and reveals implications for data privacy and democracy.

---

29. European Convention of Human Rights, art. 8; European Charter of Fundamental Rights, arts. 7 & 8.

30. *Data Protection*, EUR. DATA PROT. SUPERVISOR, https://edps.europa.eu/data-protection/data-protection_en [https://perma.cc/EPM3-S2ZP]. When this piece refers to privacy issues in the U.S. regulatory context, the term privacy or data privacy refers to "a right of personhood, intimacy, secrecy, limited access to the self, [or] control over information." DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 43 (2018).

31. *See* Danielle Citron, *Open Code Governance*, 2008 U. CHIC. LEG. F. 355 (2015) (a foundational work on how agencies increasingly rely on information systems).

32. Schwartz, *supra* note 26, at 771; Will Knight, *AI Is All the Rage. So Why Aren't More Businesses Using It?*, WIRED (July 30, 2020), https://www.wired.com/story/ai-why-not-more-businesses-use/ [https://perma.cc/P86S-4JR3]; Proposed EU AI Act, *supra* note 26.

A. The Nature and Application of Artificial Intelligence

The definition of AI has shifted over time as technologies continue to evolve and accomplish tasks previously thought to be impossible.[33] Today, there is no agreement as to what constitutes the concept of AI among experts in the field.[34] The EU's draft AI Act defines "AI system[s]" as "software that is developed with one or more of [certain] . . . techniques and approaches" and that can "for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."[35] Another European Commission (EC) policy initiative defined AI as human-designed systems that can achieve a given goal by using techniques such as machine learning, machine reasoning, and robotics to decide the best actions in a cyber or physical dimension.[36] To the EC, one of the crucial definitional concepts of AI systems is rationality.[37] AI systems with a substantial degree of rationality can perform intellectual tasks through reasoning to optimize their logical decisions. Many rational AI systems are based on machine learning algorithms that, when provided with data, identify patterns, create models, learn from experiences, and achieve solutions without explicit rules or human intervention.[38] Deep learning algorithms, a branch of machine learning algorithms,[39] feed massive quantities of personal data and utilize multiple layers of human-like neural networks to classify unstructured data, grasp concepts, decide criteria, identify correlations, and make decisions without human supervision.[40] The complexity and dynamism of this algorithmic system have prevented developing a clear explanation of algorithmic reasoning.[41]

---

33. Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 360 (2016).

34. *Id.*

35. Proposed EU AI Act, *supra* note 26, at 18, 39.

36. *A Definition of Artificial Intelligence: Main Capabilities and Scientific Disciplines*, *in* EUROPEAN COMMISSION'S HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE (2018) [hereinafter *A Definition of Artificial Intelligence*], https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf [https://perma.cc/V5GJ-QTTV].

37. *Id.* at 2.

38. For recent advances in machine learning algorithms, see Andreas Holzinger, Markus Plass, Michael Kickmeier-Rust, Katharina Holzinger, Golria Cerasela Crisan, Camelia-M. Pintea & Vasile Palade, *Interactive Machine Learning: Experimental Evidence for the Human in the Algorithmic Loop*, 49 APPLIED INTELLIGENCE 2401 (2019); Louis Columbus, *State of AI and Machine Learning in 2019*, FORBES (Sept. 8, 2019), https://www.forbes.com/sites/louiscolumbus/2019/09/08/state-of-ai-and-machine-learning-in-2019/?sh=73f59dd81a8d [https://perma.cc/2UW9-QNRM].

39. Bernard Widrow & Michael A. Lehr, *30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation*, 78 PROCS. IEEE 9 (1990).

40. Yann LeCun, Yoshua Bengio & Geoffrey Hinton, *Deep Learning*, 521 (7553) NATURE 436 (2015) ("Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction.").

41. *See* Charlotte A. Tschiderd, *Beyond the "Black Box"*, 98 DENV. L. REV. 683, 705–06 (2021).

The development of AI depends on the use of data and computing technologies. As stated by the EC's data strategy, data-driven technologies can greatly influence the capabilities of AI systems.[42] As an illustration, machine learning algorithms require tremendous amounts of data to analyze human reactions for industrial applications.[43] Additionally, effective AI systems need advanced computing technologies that can carry out operations to address complex problems at lightning speed.[44] With progressive data collecting and computing technologies, firms across EU Member States have derived substantial benefits from AI.[45]

In Europe, more and more aspects of real life have come to be reflected and simulated in AI business applications. AI systems have functioned as voice assistants in phones, Internet search engines, or hiring algorithms in the workplace.[46] These commercial applications of AI systems can be used not only to represent the real world, but also to generate new insights into it. The French startup Qucit is using AI to quantify city activities to improve urban transportation and reduce environmental pollution.[47] Likewise, Braingineers, an Amsterdam-based firm, applies AI to emotion analytics to help firms understand how user emotions and actions can be influenced by content.[48] Google, an American multinational technology giant that has dominated the European search engine market, also uses AI to offer customized search results and targeted advertisements.[49] As AI systems transform the world, even more

42. *A European Approach to Artificial Intelligence*, EUROPEAN COMMISSION, https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence [https://perma.cc/H3EV-EWW7] (last visited Sept. 13, 2022).

43. Joe McKendrick, *The Data Paradox: Artificial Intelligence Needs Data; Data Needs AI*, FORBES (June 27, 2021), https://www.forbes.com/sites/joemckendrick/2021/06/27/the-data-paradox-artificial-intelligence-needs-data-data-needs-ai/ [https://perma.cc/GY3S-46K8].

44. MAX CRAGLIA, EUROPÄISCHE GEMEINSCHAFTEN & GEMEINSAME FORSCHUNGSSTELLE, ARTIFICIAL INTELLIGENCE A EUROPEAN PERSPECTIVE (2018).

45. In October 2020, the European company Atos launched "LEONARDO," a supercomputer that can make more than 248 million billion calculations per second, preparing for the next generation of AI systems. *EU High-Performance Computing: One of the Fastest AI Supercomputers in the World Launched in Italy*, EUROPEAN COMMISSION DIGITAL STRATEGY (2020), https://digital-strategy.ec.europa.eu/en/news/eu-high-performance-computing-one-fastest-ai-supercomputers-world-launched-italy [https://perma.cc/PKG8-DRST].

46. *See How Mobile Apps Are Leveraging the Internet of Things (IoT)*, BUS. APPS (May 17, 2018), https://www.businessofapps.com/news/how-mobile-apps-are-leveraging-the-internet-of-things-iot/ [https://perma.cc/QHJ5-JDZX].

47. *Our Mission*, QUCIT, https://qucit.com/en/mission [https://perma.cc/5LSG-2796].

48. Ben Dickson, *5 European Companies that Are Advancing AI*, NEXTWEB (Mar. 29. 2019), https://thenextweb.com/news/5-european-companies-advancing-ai [https://perma.cc/N43P-HD6V].

49. Joseph Johnson, *Google: Search Engine Market Share in Selected Countries 2021*, STATISTA (Mar. 1, 2021), https://www.statista.com/statistics/220534/googles-share-of-search-market-in-selected-countries/ [https://perma.cc/KE7B-ZJLS]; Nick Statt, *Google Personalizes Search Results Even When You're Logged Out, New Study Claims*, VERGE (Dec. 4, 2018), https://www.theverge.com/2018/12/4/18124718/google-search-results-personalized-unique-duckduckgo-filter-bubble [https://perma.cc/QLU8-PRU9].

intriguing than the benefits of AI are the threats posed by the "black box" issue of algorithmic opacity.[50]

B. The Emergence of Algorithmic Opacity

Algorithmic opacity refers to the technical complexity, trade secrecy, and managerial invisibility of AI systems that hamper stakeholder inspection of the inner workings and operating results of AI.[51] Following recent research on the opacity of algorithms, this paper categorizes algorithmic opacity into three major categories.[52] The first concerns technical opacity, which arises when the computing operations of algorithmic systems like machine learning become too complex or intricate to comprehend.[53] Machine learning algorithms continue to transform their inner workings and data structure as they operate and make decisions. Thus, it can be difficult to grasp a machine learning system's dynamic operating processes and computing results.[54] For instance, when a deep learning algorithm aims to reach a goal, multiple layers of the neural network transform simultaneously, making its structure increasingly indecipherable to human cognition.[55] Moreover, because the neural network can learn from experiences and adapt its reasoning accordingly, the operating consequences of a deep learning algorithm become even more difficult to foresee.[56] Additionally, algorithmic processing can be technically complex for anyone without technical expertise in AI—consumers, regulators, shareholders, and the broader public—to comprehend. As the EU has pointed out, although AI systems are experts at making decisions for humans, technical opacity makes it difficult to assess the patterns and estimate decisions made by algorithms,[57] creating a barrier to regulators in monitoring the industrial operations of AI systems that raise a broad range of data privacy concerns.[58]

---

50. *See* PASQUALE, *supra* note 3, at 3.

51. Lu, *supra* note 27, at 114–15. *See* PASQUALE, *supra* note 3, at 15.

52. *See, e.g.*, Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*, 3 BIG DATA & SOC'Y 1, 6 (2016).

53. *See* IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, DEEP LEARNING 1–4 (2016).

54. *Id.*

55. Deep learning refers to a subset of machine learning AI techniques that utilize artificial neural networks (ANNs) with multiple layers to learn and make algorithmic decisions. ANNs consist of connected units that mimic the neurons in a human brain. *See* Davide Castelvecchi, *Can We Open the Black Box of AI?*, NATURE (Oct. 5, 2016), https://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731 [https://perma.cc/F6KJ-P3PA].

56. Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 889, 897 (2018).

57. Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 532, 539 (2015) (discussing the unpredictable behaviors of robots that "can lead to solutions no human would have come to on her own"); Siddhartha Mukherjee, *A.I. Versus M.D.*, NEW YORKER (Apr. 3, 2017), https://www.newyorker.com/magazine/2017/04/03/ai-versus-md [https://perma.cc/DW7C-KTVF].

58. *A Definition of Artificial Intelligence*, *supra* note 36.

The second aspect of algorithmic opacity, legal opacity, can affect any subset of algorithms and AI.[59] Legal opacity originated in trade secrets law that protects commercial seclusion in proprietary algorithms and associated algorithmic practices.[60] In the EU regulatory context, since algorithms can hardly be protected under copyright law or patent law, trade secrets law remains the primary path to protecting AI systems.[61] According to the EU Trade Secrets Directive, there are three elements of a trade secret: (1) secrecy of the information that is not generally known or readily accessible, (2) commercial value derived from secrecy, and (3) a trade secret holder's reasonable efforts to maintain secrecy.[62] Under this definition, trade secrets can include a broad spectrum of information, including know-how or information that has actual or potential commercial value and has been treated as confidential.[63] Google's search engine—an algorithmic application that decides the number of links, the relationships among webpages, and the optimization of search results—can serve as a well-known example of a trade secret. In the context of industrial application of AI, even if some of the information does not qualify as a trade secret, however, firms can nonetheless claim any information related to their AI applications—from the inner workings of algorithmic practices to business applications of algorithms—as trade secrets.[64] This situation has resulted from a flaw in intellectual property law that protects commercial opacity over public transparency in private AI systems.[65]

The third form of opacity, managerial invisibility or organizational secrecy, arises due to a lack of access to information on a firm's data management measures and organizational structure developed for compliance with regulations.[66] The lack of access to managerial and legal aspects of algorithmic practices is a result of current laws' failure to require algorithmic disclosures that

---

59. For a theorized perspective on various forms of legal opacity leading to AI's black box issue in the U.S. regulatory context, see Charlotte A. Tschider, *Legal Opacity: Artificial Intelligence's Sticky Wicket*, 106 IOWA L. REV. ONLINE 126 (2021).

60. *See* Sonia K. Katyal, *The Paradox of Source Code Secrecy*, 104 CORNELL L. REV. 1183, 1238 (2019) (discussing in-depth how trade secrets law becomes a default venue to protect source code).

61. *See* Katarina Foss-Solbrekk, *Three Routes to Protecting AI Systems and Their Algorithms Under IP Law: The Good, the Bad and the Ugly*, 16 J. INTELL. PROP. L. & PRACTICE 247 (2021). For a discussion of copyright and patent law protections for software, see Pamela Samuelson, *Staking the Boundaries of Software Copyrights in the Shadow of Patents*, 71 FLA. L. REV. 243 (2019).

62. Directive 2016/943 of the European Parliament and of the Council of 8 June 2016 on the Protection of Undisclosed Know-How and Business Information (Trade Secrets) Against Their Unlawful Acquisition, Use and Disclosure, 2016 O.J. (L. 157/1) 18 [hereinafter Trade Secrets Directive].

63. *Id.*

64. *See generally* Charles Tait Graves & Sonia K. Katyal, *From Trade Secrecy to Seclusion*, 109 GEO. L.J. 1337 (2021).

65. Katyal, *supra* note 60, at 1188 (arguing that "the very substance of what is secluded often stems from the most public of origins, and often produces the most public of implications. It is the shortcomings of intellectual property law that have made this possible").

66. I am grateful to Andrew Selbst for the idea of adding organizational opacity as the new category of opacity.

enable comprehensive risk assessment and stakeholder oversight. For instance, if Google is not required to disclose its AI-based services, the negative impacts of its algorithmic systems, breaches of data privacy laws, or its development plans and resource allocation for managing risks, it is unlikely to disclose such information voluntarily, because doing so is time-consuming, costly, and may cause consumer concerns. Without access to information on associated risk management measures, consumers, shareholders, regulators, and other stakeholders cannot know whether the firm operates its algorithmic systems in a lawful manner that meets their expectations. The secrecy of important managerial, organizational, and legal aspects of algorithmic practices can thus shield improper governance systems from stakeholder view, hindering effective detection of AI's erosion of data privacy.[67]

II.
THE THREATS OF ALGORITHMIC OPACITY TO DATA PRIVACY

AI systems can both benefit and harm society, not only helping firms complete complex tasks and offer better services with greater efficiency, but also critically exacerbating the erosion of data privacy and other democratic values. As many scholars have observed, opacity that conceals the operations of AI systems is a hindrance to private accountability,[68] as the core elements of AI applications that ensure comprehension of AI are seldom available for stakeholder review.[69] The failure to monitor corporate use of AI has thus imperiled enforcement of data privacy and other fundamental rights. The following section discusses the current concept of data privacy in Europe, illustrating how opaque AI systems have compromised data privacy, human rights, and democratic norms.

A. Concepts of Privacy and Data Protection in Europe

The concept of data privacy has been hard to define, and the meaning of privacy and data protection often varies depending on jurisdictions and their social contexts.[70] In the EU, the rights to privacy and data protection are both

67. *See e.g.*, Robert H. Sloan & Richard Warner, *Beyond Bias: Artificial Intelligence and Social Justice*, 24 VA. J.L. & TECH. 1, 23-26 (arguing "lack of information is presumptively unfair" for consumer privacy).

68. *See, e.g.*, Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WA. L. REV., 1, 1–33 (2014); Natalie Ram, *Innovating Criminal Justice*, 112(4) NW. U. L. REV. 659, 659–724, (2018); Rebecca Wexler, *Life, Liberty, and Trade Secrets*, 70 STAN. L. REV. 1343, 1343–1429 (2018).

69. *See* Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*, 3 BIG DATA SOC. 1, 1–12 (2016) (discussing three forms of opacity that arise from machine learning systems).

70. Robert Post, *Three Concepts of Privacy*, 89 GEO. L.J. 2087, 2087 (2001); Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 480-81 (2006); Min Sunitha Abhay Jain, *Artificial Intelligence: A Threat to Privacy*, 8 NIRMA U. L. J. 21, 25 (2019).

considered fundamental rights.[71] The notion of privacy concerns dignity and the right to private life, including a right to autonomy and control over data.[72] In today's Europe, the term privacy has been fused with the concept of "data protection" to imply protection of "information relating to an identified or identifiable natural person (data subject)."[73] Currently, the EU's definition of personal data encompasses all types of data relating to a data subject, including intimate relationships, health information, and personal preferences.[74] As stated by the GDPR, all personal data processing shall be based on six key principles that require data controllers to protect personal data to the greatest extent possible.[75] These principles ask that all types of personal data are handled in a lawful, fair, and transparent way;[76] gathered and processed solely for defined purposes;[77] collected and processed only when necessary to accomplish defined purposes;[78] kept accurate and up-to-date;[79] stored for no longer than necessary;[80] and handled in a manner that protects the security of personal data.[81] The GDPR holds firms responsible for demonstrating that their data practices follow all the principles mentioned above.[82] The Regulation's right to data protection thus ensures protection against unauthorized access to the self, as well as greater individual autonomy from the invisible hands of enterprises.[83] Furthermore, the expanded GDPR principles and duties require that various democratic values be taken into account—including fairness, transparency, and accountability.

B. Data Privacy and Democratic Values Compromised by Algorithmic Opacity

Despite EU regulators' conscious efforts to protect data privacy, many opaque AI-associated data practices have been found to neglect privacy and data protection principles, causing intractable problems with fairness, transparency, and accountability, among others. This section examines the opacity issue through four lenses: (1) data collection, processing, and analysis; (2) algorithmic manipulation; (3) algorithmic bias, misinformation, and incitement; and (4) data management measures.

---

71. *See Data Protection*, *supra* note 30 ("The notion of data protection originates from the right to privacy and both are instrumental in preserving and promoting fundamental values and rights.").
72. *Id.*
73. *Id.*; *General Data Protection Regulation (GDPR)*, *supra* note 24, at art. 4(1).
74. General Data Protection Regulation (GDPR), *supra* note 24, at art. 4(1).
75. *Id.* at art. 5.
76. *Id.* at art. 5.1.(a).
77. *Id.* at art. 5.1.(b).
78. *Id.* at art. 5.1.(c).
79. *Id.* at art. 5.1.(d).
80. *Id.* at art. 5.1.(e).
81. *Id.* at art. 5.1.(f).
82. *Id.* at art. 5.2.
83. Manheim and Kaplan, *supra* note 4, at 118; Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085, 1108 (2018) ("Considered as a whole, they begin to sound like the general idea of due process in all its expansiveness.").

1. Opaque Data Collection, Processing, and Analysis

Data is the engine behind AI systems.[84] Developing useful AI systems requires rich and meaningful data to feed and train algorithms. Through data, AI systems can reproduce historical events and establish predictive models to help firms estimate customer needs.[85] Thus, firms strive to find the best personal data source to create profitable AI-based services or products.[86] Due to the immense economic value of data and the difficulty of complete compliance with regulations, firms may be strongly motivated to evade legal limitations and trade data to business partners for commercial profit.[87] The business applications of AI systems have therefore given rise to a wide range of data privacy issues.[88]

In Europe, for example, many firms use tracking technologies that integrate AI techniques to connect digital devices and process data for innovative AI applications at the cost of data privacy.[89] For now, the Artificial Intelligence of Things (AIoT)—a technology that combines AI and IoT techniques to process data and make automated decisions without human interference—has been applied across various devices to collect data, create personalized services, and solve problems independently.[90] Such AI-powered technologies can present serious threats to data privacy.[91] Using tracking algorithms, firms have been covertly collecting data that are regarded as personal, sensitive, or secret from mobile phones, wearable devices, and other smart devices in homes, workplaces, and public places.[92] With data fusion techniques, firms can merge, organize, and analyze individual data points to profile individuals for obscure business applications. Furthermore, corporate AI techniques have enlarged the scale of

84. *Id.* at 111–12.

85. Paul M. Schwartz, *Privacy and Security Law: What Korean Companies Need to Know*, PAUL HASTINGS, https://paulschwartz.net/wp-content/uploads/2018/12/PH_Perspective_Korea-Privacy-and-Security-Law-Schwartz.pdf [https://perma.cc/VRK2-H672] ("Personal data is the gold of the information economy—a new profit source for companies seeking to know more about their customers and better meet their needs.") (last visited Sept. 2, 2022).

86. Manheim and Kaplan, *supra* note 4, at 111–12.

87. *Id.*

88. Joseph Zulick, *AIoT: The Power of Combining AI with the IoT*, RELIABLEPLANT, https://www.reliableplant.com/Read/31799/combining-ai-iot [https://perma.cc/3STU-J8X2].

89. Janakiram MSV, *Why AIoT Is Emerging as the Future of Industry 4.0*, FORBES (Aug. 12, 2019), https://www.forbes.com/sites/janakirammsv/2019/08/12/why-aiot-is-emerging-as-the-future-of-industry-40/?sh=ad7e50e619be [https://perma.cc/FR52-CWHD] (discussing how Artificial Intelligence of Things (AIoT) can serve as the brain of connected devices).

90. Bernard Marr, *What Is the Artificial Intelligence of Things? When AI Meets IoT*, FORBES (Dec. 20, 2019), https://www.forbes.com/sites/bernardmarr/2019/12/20/what-is-the-artificial-intelligence-of-things-when-ai-meets-iot/ [https://perma.cc/F5RP-5LHP].

91. Manheim and Kaplan, *supra* note 4 at 111–12.

92. *See* Jennifer M. Urban, Chris Jay Hoofnagle & Su Li, MOBILE PHONES AND PRIVACY, BERKELEY CONSUMER PRIVACY SURVEY & BCLT RESEARCH PAPER 2, 4–5 (July 11, 2012), https://www.ftc.gov/system/files/documents/public_comments/2013/12/00007-89101.pdf [https://perma.cc/7CKB-NC28].

data collection, processing, and transferring for automatic decision-making and other purposes unknown to users.[93]

In this context, algorithmic opacity has become a barrier to the identification of algorithmic practices that collect, process, or transfer personal data without legal basis. Considering trade secrecy and other commercial interests, firms may refuse to disclose the role of algorithms in their business models and how their algorithmic practices pose risks to data privacy. Without sufficient description of data privacy risks derived from algorithmic practices and associated risk management measures, organizational secrecy prevents stakeholders from discovering any potential unauthorized processing and improper data management through current privacy disclosures. The technical opacity of algorithmic systems also impedes comprehension of algorithmic decision-making processes and results.[94] Even if a firm discloses that its smart home devices deploy privacy-preserving techniques to handle personal data, stakeholders have no way to measure how effectively such techniques prevent users' personal data from unauthorized processing. Without adequate regulation that requires meaningful disclosures on managerial and technical aspects of AI applications, clandestine data collection, processing, and transfer from millions of websites and digital devices has constantly intruded on individuals' private lives.[95] As a result, the use of AI has constituted a form of commercial seclusion that results in power imbalances, surrenders the protection of personal data, and risks eroding trust between private enterprises and individuals. [96]

2. Opaque Algorithmic Manipulation

Algorithmic manipulation is the process of using AI systems to analyze personal data, detect individual vulnerabilities, and offer tailored content to directly or indirectly manipulate one's decision-making processes.[97] One of the typical forms of algorithmic manipulation concerns using data for behavioral advertising and responsive search ads.[98] With collected data, firms can use predictive analytics to estimate user actions and trade extracted insights for

93. Mark Lippett, *Fixing the Biggest IoT Issue—Data Security*, INFOSECURITY MAG., (Aug. 7, 2020), https://www.infosecurity-magazine.com/opinions/fixing-biggest-iot-data/ [https://perma.cc/XBW7-D5B8] (discussing how the cloud has led to unprecedented data access that contributes to data protection concerns).

94. However, AIoT can bring many benefits to consumers and firms, for example see Dinesh Soundararajan, *AIoT: The Powerful Convergence of AI and the IoT*, IOT NOW (Apr. 10, 2020), https://www.iot-now.com/2020/04/10/102236-aiot-the-powerful-convergence-of-ai-and-the-iot/ [https://perma.cc/WVG9-JAKP].

95. Zulick, *supra* note 88.

96. Lu, *supra* note 27, at 118–21.

97. Jon Whittle, *AI Can Now Learn to Manipulate Human Behaviour*, THE CONVERSATION (Feb. 10, 2021), https://theconversation.com/ai-can-now-learn-to-manipulate-human-behaviour-155031 [https://perma.cc/65H7-4U62].

98. *See* Matthew Crain & Anthony Nadler, *Political Manipulation and Internet Advertising Infrastructure*, 9 J. INFO. POL'Y 370, 370 (2019).

commercial manipulation.[99] For instance, before the 2016 U.S. presidential election, Cambridge Analytica misused the data of millions of Facebook users to send targeted online ads in campaigns that manipulated swing voters into supporting Donald Trump.[100] Such algorithmic manipulation is merely the tip of the iceberg.[101]

Today, online behavioral advertising is regarded as the largest market for data collectors, where firms such as Spotify and Google have begun to use AI to create personalized content profiles.[102] To influence consumers in their shopping choices, many firms collect a broad spectrum of real-time information about users through algorithms to examine user actions and predict their needs.[103] Furthermore, firms use AI to integrate such data with other information previously obtained to create personalized ads or content profiles. By doing so, firms can display specific advertisements or content to the best targets according to users' personality profiles.[104] This enhances market efficiency by connecting consumers and their potentially preferred merchants while stimulating the digital economy by bringing financial gain to platforms and advertisers.[105]

However, the use of personalized content can quickly turn into a form of algorithmic coercion for influence.[106] Allowing the use of personalized content profiles means that firms can categorize individuals into personality clusters and then decide what users can and cannot see.[107] Although some companies claim that they have attempted to avoid third-party tracking for targeted advertising, many of them still operate algorithms in a way that exacerbates algorithmic manipulation. For instance, Google's use of federated learning of cohorts (FLoC)—a so-called privacy-preserving online tracking technique that gathers users' webpage visit history and groups them into "cohorts" to prevent the creation of "personal profiles"—is nevertheless found to perpetuate erosion of

99. *How Artificial Intelligence (AI) Is Used in Targeted Marketing*, AZATI (2020), https://azati.ai/artificial-intelligence-targeted-marketing/ [https://perma.cc/9JVX-4DKM].

100. Dipayan Ghosh & Ben Scott, *Facebook's New Controversy Shows How Easily Online Political Ads Can Manipulate You*, TIME (Mar. 19, 2018), https://time.com/5197255/facebook-cambridge-analytica-donald-trump-ads-data/ [https://perma.cc/R36L-4LBQ].

101. Joe Westby, *'The Great Hack': Cambridge Analytica Is Just the Tip of The Iceberg*, AMNESTY INT'L (July 24, 2019), https://www.amnesty.org/en/latest/news/2019/07/the-great-hack-facebook-cambridge-analytica/ [https://perma.cc/R6PE-YGCS].

102. *Id.*

103. *See, e.g.*, IBM Watson Advertising, *Five Benefits of Machine Learning in Advertising*, IBM (Feb. 24, 2021), https://www.ibm.com/watson-advertising/thought-leadership/benefits-of-machine-learning-in-advertising [https://perma.cc/3DWE-C7JF].

104. For a further discussion of the application of online behavior advertising, see Sophie C. Boerman, Sanne Kruikemeier & Frederik K. Zuiderveen Borgesius, *Online Behavioral Advertising: A Literature Review and Research Agenda*, 46 J. ADVERT. 363 (2017).

105. *Id.* at 363.

106. *See How Artificial Intelligence (AI) Is Used in Targeted Marketing*, *supra* note 99.

107. *See* Javier Parra-Arnau, David Rebollo-Monedero & Jordi Forné, *Measuring the Privacy of User Profiles in Personalized Information Systems*, 33 FUTURE GENERATION COMPUT. SYST. 53, 61 (2014).

data privacy.[108] As the FLoC assigns consumers to over thirty thousand interest groups, firms can not only extract sensitive information about their personal interests, but also obtain data on users' browser types to create consumer profiles with detailed information for behavioral targeting.[109] To address such privacy problems, Google proposed to replace the FLoC with Topics, a new technique that categorizes people into around 350 interest groups and avoids grouping individuals into "sensitive categories" based on race or gender.[110] Nevertheless, Google says that advertisers can still combine information offered by Topics with additional data to infer sensitive information about consumers for targeted marketing.[111] In this context, consumers tend to receive manipulated information offered by algorithms instead of open and unbiased access to a marketplace.

Today, firms can use AI for targeted marketing purposes with opacity. Currently, firms are not obligated to disclose whether their AI applications constitute manipulation or how they manage to protect users from manipulation against their best interests. Although many firms admit that they use algorithms to offer personalized services, organizational opacity allows them to avoid disclosures of the controversial existence of manipulation and any negative impact of such manipulative practices. As a result, individuals have no practical way to challenge the personality profiles created for them or to resist such manipulation of information. Algorithmic opacity thus permits AI applications to decide the content individuals see and, by extension, erode individual autonomy to manipulate personal decisions without stakeholder oversight.[112]

3. Opaque Algorithmic Bias, Misinformation, and Incitement

AI systems not only process data to influence individuals' online choices, but also improperly utilize data to influence their thinking, actions, and opportunities in real life.[113] Data constructs a virtual world that represents the human world, a physical realm that is far from neutral. Likewise, algorithms fed by data reflect the training data, which is representative of the biased human

---

108. Chetna Bindra, *Building a Privacy-First Future for Web Advertising*, GOOGLE ADS COM. BLOG (Jan. 25, 2021), https://blog.google/products/ads-commerce/2021-01-privacy-sandbox/ [https://perma.cc/2GMH-Z3SG]; Amber Neely, *Google Starts Funneling Chrome Users into Groups to Ease Targeted Advertising*, APPLE INSIDER (Apr. 7, 2021), https://appleinsider.com/articles/21/04/07/google-starts-funneling-chrome-users-into-groups-to-ease-targeted-advertising [https://perma.cc/5S87-FMP9].

109. Neely, *supra* note 108.

110. *Topics*, GOOGLE: PRIVACY SANDBOX (Jan. 25, 2022), https://privacysandbox.com/intl/en_us/proposals/topics [https://perma.cc/372Y-7FLY].

111. Matt Burgess, *Google Has a New Plan to Kill Cookies. People Are Still Mad*, WIRED (Jan. 27, 2022), https://www.wired.com/story/google-floc-cookies-chrome-topics/#:~:text=Google's%20plan%20to%20remove%20third,is%20back%20with%20another%20plan [https://perma.cc/6K6U-YGFL].

112. Lu, *supra* note 27, at 119.

113. Knowledge at Wharton, *Who Made That Decision: You or an Algorithm?*, WHARTON SCH. OF THE UNIV. OF PA. (Mar. 25, 2019), https://knowledge.wharton.upenn.edu/article/algorithms-decision-making/ [https://perma.cc/8C5D-8Q56].

world.[114] In the process of creating algorithms, firms may collect data based on biased selection criteria; data trainers and designers may embed biased values into these algorithms.[115] Thus, AI systems are more likely to produce bias if not monitored to ensure fairness. The resulting AI-based services have neglected data for certain groups, replicated unfair decision-making processes, and perpetuated social segregation.

In Europe, for instance, cardiovascular diseases have long been considered to be exclusive to men, so the majority of cardiovascular disease data has been collected from male medical records.[116] AI applications that rely on medical records will thus decide that a man's heart attack is caused by cardiovascular diseases, yet a woman with the same symptoms is likely to be diagnosed with a mental illness by the AI program, although scientific research has found that women are also likely to suffer from heart attacks.[117] Despite AI systems being fed with seemingly comprehensive data, as a director of eHealth at Charité of the Berlin Institute of Health pointed out, "There are huge data gaps regarding the lives and bodies of women."[118] As a significant amount of healthcare data is collected from military sources, where women only represent 6 percent, such an incredibly small sample is likely to replicate inaccurate predictions sourced from gender bias embedded in algorithms.[119] In addition to gender, structural race, ethnic, and class discrimination also emerge in AI services or products.[120] Firms using AI to select prospective job candidates have found that their systems discriminate against women or people of color because of the purported undesirability of such applicants based on training data.[121]

Also, biased algorithms have been revealed to facilitate hate speech, incite violence, and create outbreaks of social disorder. In the astounding January 6, 2021 insurrection at the U.S. Capitol, tech giants such as Facebook, Twitter, and YouTube were blamed for using algorithms that favored the spread of biased and extreme information.[122] Facebook has been known to deploy its AI systems "in ways that favor extreme speech and behavior," leaving users vulnerable to

114. *See* Anupam Chander, *The Racist Algorithm?*, 115 MICH. L. REV. 1023, 1034–36 (2017).

115. *Id.* at 1036.

116. Carmen Niethammer, *AI Bias Could Put Women's Lives at Risk—A Challenge for Regulators*, FORBES (Mar. 2, 2020), https://www.forbes.com/sites/carmenniethammer/2020/03/02/ai-bias-could-put-womens-lives-at-riska-challenge-for-regulators/?sh=12d96889534f [https://perma.cc/URN3-TPDB].

117. *Id.*

118. *Id.*

119. *Id.*

120. *See* Latanya Sweeney, *Discrimination in Online Ad Delivery*, ARXIV:1301.6822 [CS.IR] 2 (2013).

121. FREDERIK ZUIDERVEEN BORGESIUS, DISCRIMINATION, ARTIFICIAL INTELLIGENCE, AND ALGORITHMIC DECISION-MAKING 20, 25 (2018), https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73 [https://perma.cc/V73Y-5BB4].

122. Roger McNamee, *Platforms Must Pay for Their Role in the Insurrection*, WIRED (Jan. 7, 2021), https://www.wired.com/story/opinion-platforms-must-pay-for-their-role-in-the-insurrection/. [https://perma.cc/B8MT-EJDJ].

manipulation.[123] Indeed, according to Facebook's own research, "64 percent of the time a person joins an extremist Facebook Group, they do so because the platform recommended it."[124] Research also pointed out that social media algorithms have exacerbated the spread of deepfakes and misinformation to online users.[125] Without regulatory tools reviewing the algorithms that produce biased and extreme misinformation, algorithms can play a harmful role in provoking divisions, violence, and terrorism.

In all these cases, firms are not obligated to disclose how they build algorithms and AI to ensure fairness and accuracy. Algorithmic opacity permits them to develop and deploy AI-based products or services without stakeholder surveillance. Absent disclosure rules, organizational secrecy impedes the stakeholder detection of biased, extreme, inaccurate, or inciting AI applications. Since firms are not compelled to explain the risks involved and what they have done to ensure the quality of algorithms, outsiders cannot readily identify the existence of bias and error in AI applications. Even if a user intends to challenge certain algorithmic decisions, the computing processes and operating results of these algorithms can be claimed as trade secrets and are otherwise too technically complicated for the user to detect their flaws. Without critical information on algorithmic practices, stakeholders may find it difficult to investigate the cause of discriminatory, invasive, or inciting algorithmic operations. Accordingly, algorithmic opacity can amplify the marginalization of minorities, breaking rules of fairness, lawfulness, and transparency in a data protection context.

4. Opaque Data Management Measures

Data management measures refer to the ways firms consider their business models, risk control systems, and managerial approaches in light of data privacy protection.[126] The enforcement of data privacy rules relies on data management measures that ensure respect for user privacy.[127] Evaluating a firm's data management measures is crucial to the sustainability of a business, as inappropriate data management measures can lead to data privacy incidents,

123. *Id.*

124. *Id.*

125. Deepfakes refer to synthetic media that use machine learning algorithms to overlay faces on videos of other individuals, with fake—but highly convincing—results. Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, CALIF. L. REV. 1753, 1758 (2019) ("[Deepfake technology] leverages machine learning algorithms to insert faces and voices into video and audio recordings of actual people and enables the creation of realistic impersonations out of digital whole cloth.").

126. For a thorough discussion of privacy management tools and approaches, see Maryline Laurent & Claire Levallois-Barth, *Privacy Management and Protection of Personal Data*, *in* DIGIT. IDENTITY MGMT. 137–205 (Maryline Laurent & Samia Bousefrane eds., 2015).

127. *Id.* at 143–164, 171–184.

serious violations, massive fines, and impaired public trust in a firm.[128] Shareholders are also increasingly demanding information on corporate approaches to the protection of human rights and democratic values, including data privacy.[129] Inadequate data management measures can adversely affect a firm financially and have social impacts on the public.[130] Since the GDPR took effect, data management measures have played a key role in influencing a firm's financial conditions, since firms can be fined up to €20 million or 4% of their annual global revenue for data privacy violations.[131] As of May 15, 2022, there were at least 1,079 fines in the total amount of €1.62 billion imposed on firms for breaking GDPR rules,[132] and it is estimated that there will be a large number of firms liable for future hefty GDPR fines.[133] One survey indicated that only 20% of firms in the United States and the EU are compliant with GDPR rules, indicating that corporate governance and complete compliance with data privacy law are far from uniform.[134] These data privacy incidents have negative impacts not only on citizens' data privacy, but also on firms' financial conditions, social reputation, and capital funds, which are all foundational aspects of corporate sustainability.

However, without information on a firm's data management measures, algorithmic opacity hides essential data practices from stakeholder scrutiny and tacitly makes avoidance of corporate accountability a default setting.[135] Although firms rely heavily on privacy policies and notices to demonstrate that they value data privacy, few if any of them account for their data governance in response to data privacy risks posed by AI.[136] Due to algorithmic opacity disguised as trade secrecy and organizational invisibility in AI systems, the outcomes of firms' policies for AI applications remain opaque to the public. From a financial perspective, such opacity prevents stakeholders from estimating the effectiveness of data management measures and hinders investors from making informed investment decisions, forming barriers in the capital market as

128. Sonia Cheng & Steve McNew, *Information Governance: Achieving Data Ethics, Privacy and Trust*, ETHICAL BOARDROOM (2019), https://ethicalboardroom.com/information-governance-achieving-data-ethics-privacy-and-trust/ [https://perma.cc/JW39-NWFR].

129. *See* David Hess, *The Transparency Trap: Non-Financial Disclosure and the Responsibility of Business to Respect Human Rights*, 56 AM. BUS. L.J. 5, 49 (2019).

130. Cheng & McNew, *supra* note 128.

131. General Data Protection Regulation (GDPR), *supra* note 24, at art. 83(5).

132. *GDPR Enforcement Tracker*, CMS.LAW, https://www.enforcementtracker.com [https://perma.cc/C3KW-3FJV].

133. *5 Staggering GDPR Fines in 2020: Where Are We Headed in 2021?*, AXIOMQ (Jan. 12, 2021), https://axiomq.com/blog/5-staggering-gdpr-fines-in-2020-where-are-we-headed-in-2021/ [https://perma.cc/E98B-TC4W].

134. *20% of Companies Report Being GDPR Compliant Post May 25 Deadline*, TRUSTARC (July 12, 2018), https://trustarc.com/20-of-companies-report-being-gdpr-compliant-post-may-25-deadline/ [https://perma.cc/5QU2-ZA45].

135. Rebecca Wexler, *Code of Silence*, WASH. MONTHLY (June 11, 2017), https://washingtonmonthly.com/2017/06/11/code-of-silence/ [https://perma.cc/7AS9-WD5E].

136. See Ari Ezra Waldman, *Privacy Law's False Promise*, 97 WASH. U. L. REV. 773 (2020).

a result.[137] Moreover, inadequate data practices can bring about legal proceedings that cause tarnished reputation, litigation costs, and reduced external investment in firms, ultimately frustrating firms' ability to sustainably innovate AI-based products and services. From a social perspective, without accessible information on data management measures, firms may pay less attention to improving data governance. Consequently, algorithmic opacity derived from invisible data management measures has led to a lack of private accountability for corporate use of AI, hurting not only citizens but also firms themselves.

III.
THE EU RESPONSE TO AI'S RISKS TO DATA PRIVACY IN THE SHADOW OF ALGORITHMIC OPACITY

As algorithmic opacity hinders the monitoring of corporate AI applications, individuals continue to suffer from unauthorized data processing, manipulative information, unfair decision-making, and more without sufficient information to redress their harm; shareholders are increasingly afflicted with violations, fines, reputational damage, and decreases in share price due to lack of information regarding firms' inadequate management measures; regulators are burdened with lack of information on corporate algorithmic practices for enforcement of data privacy rules. Owing to algorithmic opacity that impedes stakeholder view of private AI, corporate misconduct that has been investigated reveals merely a small part of the problem.

In response to data privacy risks posed by AI in the business application context, the EU has established stricter data protection law that underlines the high status of data protection as a fundamental right.[138] On May 25, 2018, the GDPR took effect, replacing the Data Protection Directive and becoming the primary EU privacy law binding on all Member States.[139] As noted, the GDPR has become a leading data privacy regulation worldwide, as it seeks to meet the needs of citizens in the new digital era and regulate a large economic market.[140] Additionally, as a regulation, the GDPR harmonizes EU data protection

137. Directive (EU) 2017/828 of the European Parliament and of the Council of 17 May 2017 amending Directive 2007/36/EC as regards the encouragement of long-term shareholder engagement, 2017 O.J. (L.132/1). For a more detailed discussion of the financial impact of algorithmic opacity, see Lu, *supra* note 27, at 104–06.

138. *See generally* STOA, THE IMPACT OF THE GENERAL DATA PROTECTION REGULATION (GDPR) ON ARTIFICIAL INTELLIGENCE, EUROPEAN PARLIAMENTARY RESEARCH SERVICE (2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf ("The GDPR generally provides meaningful indications for data protection in the context of AI applications.").

139. *See The History of the General Data Protection Regulation*, EURO. DATA PROTECTION SUPERVISOR, https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en [https://perma.cc/3WF9-DZ77] (last visited Sept. 2, 2022).

140. As an illustration, due to the "Brussels effects," even the United Kingdom chose to enact the GDPR into its domestic law after leaving the EU. *See* Paul M. Schwartz, *The Data Privacy Law of Brexit: Theories of Preference Change*, 22.2 THEORETICAL INQUIRIES L. 111, 113, 127 (2021).

legislation and sets out stronger enforceable rules through a coherent regulatory framework.[141] It overrides domestic regulations with direct legal effect and adheres to previous European principles that consider data protection as a human right, while also enhancing individual rights to data protection in the digital age.[142] Nonetheless, as analyzed in this section, the GDPR still falls short of addressing the various opacity issues raised by AI. The following section assesses the GDPR's regulatory solutions to challenges posed by AI—increased individual control, rights related to automatic decision-making, transparency duties, and data protection by design—and demonstrates their inadequacy in mitigating the risks of algorithmic opacity.

A. Increased Control and Its Inadequacy

As a first step in strengthening individual rights to data protection, the GDPR gives individuals increased control over their information.[143] Through lawfulness and transparency requirements, the GDPR imposes stricter restrictions on the processing of individual data to enhance user control.[144] As required by Article 6 of the GDPR, when firms intend to process personal data, they are required to have a legal ground for such processing, such as consent, contract performance, or legitimate interest.[145] In many AI applications, firms obtain consent from users in an understandable and transparent manner through privacy notices. According to Article 4(11) of the GDPR, such consent must be a "freely given, specific, informed and unambiguous indication of the data subject's . . . agreement to the processing of personal data . . . ."[146] This marks an opt-in approach that demonstrates the EU's pursuit of a higher standard of data protection, especially compared to the opt-out approach commonly adopted in U.S. federal privacy regulations. The European Data Protection Board (EDPB) advises, "affirmative opt-in methods might include signing a consent statement, oral confirmation, a binary choice presented with equal prominence, or switching technical settings away from the default."[147] In practice, to give users data control, firms can no longer bundle consent hidden in click-wrap boxes or provide inferior services due to declined consent. Individuals must clearly and

141. Fernanda Nicola & Oreste Pollicino, *The Balkanization of Data Privacy Regulation*, 123 W. VA. L. REV. 61, 78 (2020) ("[T]he move from a Directive to a Regulation cannot be disregarded, not only from a formal standpoint, but also a substantive standpoint.").

142. For a discussion of the GDPR's regulatory implications, see Chris Jay Hoofnagle, B. van der Sloot & F. Zuiderveen Borgesius, *The European Union General Data Protection Regulation: What It Is and What It Means*, 28 INFO. & COMMC'N. TECH. L. 65 (2019).

143. General Data Protection Regulation (GDPR), *supra* note 24, at arts. 5, 32.

144. *Id.*

145. *Id.* at art. 6.

146. *Id.* at art. 4(11). *See also* art. 7 (Conditions for Consent).

147. *See What Is Valid Consent?*, INFO. COMM'R'S OFFICE, https://ico.org.uk/for-organisations/guide-to-dataprotection/guide-to-the-general-data-protection-regulation-gdpr/consent/what-is-valid-consent/ [https://perma.cc/YRJ7-GYDF].

affirmatively confirm that such consent is freely granted and that they are informed of how their data will be used.[148]

Despite the increased control offered by the Regulation, however, algorithmic opacity impairs its adequacy in protecting personal data in the face of AI. The first issue relates to the incompleteness of privacy notices, which are the primary source of transparency in giving users increased control. Under the GDPR, firms rely on privacy policies and notices to disclose their use of data and present privacy choices. However, disclosures through privacy notices and policies alone have proven ineffective in providing useful information that meaningfully improves accountability. Studies have revealed that most online users do not read privacy notices in detail.[149] Reading privacy policies and notices requires a significant amount of time and knowledge of privacy terms, and forming and expressing a preference for each service exhausts individuals. Even for users who are willing to read privacy notices carefully, the information provided by such disclosures is insufficient to make careful choices. Moreover, privacy notices often use terms that are ambiguous, open-ended, over-simplified or technical, making it difficult for users to understand actual data practices and AI applications fully.[150] According to a survey conducted by the EC, around 60% of Europeans read privacy policies; yet, due to the great length and complexity of privacy statements, only 13% of them read them in full.[151] Additionally, firms are not required to describe how they use AI to process personal data or how they manage the risks involved in algorithmic practices to comply with data protection rules. Hence, users who have read multiple privacy disclosures may not easily understand how their data will be ultimately used by AI operating in opaque conditions—whether their data will be used to reveal other sensitive information, whether their data will be used for new purposes once a firm sells user data, or whether their data can be used by the requesting firm in new contexts. As a result, when users manage their privacy settings through privacy notices, their choices are seldom based on an adequate understanding of privacy disclosures and actual algorithmic practices.[152]

The second issue concerns the uncertain truthfulness of privacy disclosures in the shadow of algorithmic opacity. Given that corporate use of AI is protected as organizational secrecy or trade secrets, outsiders can seldom determine whether what a firm does with their data meets their expectations. It remains

148. *Id.*

149. *See* Daniel J. Solove, *Privacy Self-Management and the Consent Dilemma*, 126 HARV. L REV. 1880, 1884 (2013).

150. *Id.* at 1885.

151. *Data Protection Regulation One Year On: 73% of Europeans Have Heard of at Least One of Their Rights*, EUR. COMM'N (June 13, 2019), https://ec.europa.eu/commission/presscorner/detail/en/IP_19_2956 [https://perma.cc/KT5X-YECF].

152. Rishab Bailey, Smriti Parsheera, Faiza Rahman & Renuka Sane, *Disclosures in Privacy Policies: Does "Notice and Consent" Work?*, 33 LOY. CONSUMER L. REV. 1, 1 (2021) (indicating that users usually do not truly understand the actual meaning of terms like "third party," "affiliate," and "business-partner").

unknown how individuals can learn about a firm's actual processing of data, including illegitimate repurposing of data and unlawful data practices. Today, firms usually use positive language to describe their purposes for AI-based services or products, as well as broadly define their purposes for processing data to make those data practices seemingly lawful. Since firms are not required to disclose privacy risks derived from AI applications, they may avoid describing the negative effects produced by algorithms and AI in detail. Consequently, users cannot assess privacy risks to make informed decisions about AI-based data processing when confirming their privacy choices.

For instance, Google's current privacy control system—Privacy Checkup—consists of a set of ineffective privacy disclosures.[153] Privacy Checkup requires users to review privacy settings for numerous items, including but not limited to Ad Settings, Web and App Activity settings, Location History Settings, YouTube History Settings, Networking Settings, and Third-party Access Settings, with each containing lengthy descriptions that require users to review their preferences regarding corporate use of their data.[154] The significant amount of time required for a complete setting of privacy preference under this model makes it difficult for users to control their information effectively or offer meaningful consent. Additionally, Google defines privacy terms and the purpose of data processing in a very broad manner that avoids accountability. In Google's privacy policy, privacy terms like "personal information" and "third party" are so vaguely defined that the consumer cannot anticipate how their information is collected and handled by Google and unknown third parties.[155] Also, according to Google's privacy policy, the purpose of its use of data is "to provide better services to all . . . users."[156] With such a broadly defined purpose, Google can process its data for almost anything that can be said to help "provide better services to all . . . users."[157] Moreover, Google avoids describing any risks derived from its AI-based services in its privacy notices. Without mentioning a single word about the role of AI in tracking user behaviors and associated privacy risks, Google's disclosure involves a general and open-ended statement of its purpose for data use. As Google only provides a set of positive, lengthy, vague

153. Username: rozepz, *What is Google Privacy Checkup? Everything You Need to Know*, TOM'S GUIDE (Apr. 26, 2018), https://forums.tomsguide.com/faq/what-is-google-privacy-checkup-everything-you-need-to-know.188110/ [https://perma.cc/2AWS-QQVA]. *See, e.g.*, *Google Account: Privacy Checkup*, GOOGLE, https://myaccount.google.com/privacycheckup?pli=1 [https://perma.cc/DZG8-G8LC] (last visited Sept. 2, 2022).

154. Username: rozepz, *supra* note 153.

155. Bjarki Valtysson, Rikke Frank Jørgensen & Johan Lau Munkholm, *Co-Constitutive Complexity: Unpacking Google's Privacy Policy and Terms of Service Post-GDPR*, 42 NORDICOM R. 124, 137 (2021) ("[E]xplanations of key terms such as 'personal information' and 'third party' are over-simplified, making it more difficult for users to understand exactly how and when their personal data is processed, and by whom.").

156. *See Google Privacy & Terms*, GOOGLE https://policies.google.com/privacy/archive/20191015?hl=en [https://perma.cc/ENU2-NRHG] (last visited Sept. 2, 2022).

157. *Id.*

terms describing its AI-based services, users cannot make informed decisions about their privacy settings based on these privacy policies and notices. These factors might weaken the effectiveness of the control mechanism offered by the GDPR.[158] Hence, despite the GDPR's attempt to give individuals increased control over their data, existing privacy disclosures and consent models are not largely helpful in reducing organizational secrecy. As managerial invisibility and trade secrecy hide actual algorithmic practices, it remains difficult for individuals to prevent omnipresent corporate surveillance from threatening data protection rules.

B. Increased Automatic Decision-Making Rights and Their Inadequacy

Besides increased control over personal data, the Regulation also grants individuals the right not to be subject to automated decision-making.[159] Under the GDPR Article 22, automated decision-making uses a program that makes decisions through algorithms without any human intervention, when those decisions result in direct legal effects on the individual.[160] Considering the far-reaching impacts of algorithmic decisions, the GDPR restricts the use of algorithmic decisions and grants individuals a right to require human intervention or object to those decisions. According to the EDPB, automated decision-making is prohibited regardless of any action taken by a data subject.[161] The Court of Justice of the European Union (CJEU) will interpret this provision to determine whether the GDPR prohibits automated decisions by default or grants individuals a right that should be exercised by individuals.[162]

Although Article 22 of the GDPR protects individuals against unjust automated decisions, it does not apply to corporate algorithmic decisions in a variety of contexts for the following reasons. First, this provision only regulates algorithmic decisions made by machines, meaning that algorithmic systems that do not make decisions for an individual do not qualify as automated decision-making.[163]

---

158. *See* Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 J.L. POL'Y INF. SOC. 543, 564–65 (2008).

159. General Data Protection Regulation (GDPR), *supra* note 24, at art. 22.

160. *Id.*

161. *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679*, EUR. COMM'N (Oct. 3, 2017) [hereinafter *Guidelines on Automated Individual Decision-Making*], https://ec.europa.eu/newsroom/article29/redirection/document/49826 [https://perma.cc/SMC2-AJG2] ("Article 22(1) establishes a general prohibition for decision-making based solely on automated processing. This prohibition applies whether or not the data subject takes an action regarding the processing of their personal data.").

162. Sebastião Barros Vale & Gabriela Zanfir-Fortuna, *Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities*, FUTURE PRIV. F. 7 (2022), https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf [https://perma.cc/UX3S-DFQC].

163. *Id.* at 8.

Second, Article 22 of the GDPR applies to algorithmic decisions that produce legal effects or similarly "significant effects" on the data subject.[164] The GDPR uses the "significant effects" framework to grant greater protections to individuals in fields that directly implicate their legal rights. Algorithmic decisions regarding a person's legal rights, such as access to employment or medical treatment, warrant heightened protection because they produce "significant effects" on the individual, as compared to the less serious but still cognizable privacy interests present in online advertising and personalized content. According to the Article 29 Data Protection Working Party, decisions regarding a person's legal rights such as access to employment or medical treatment are those having similarly significant and important effects on the individual, whereas most online advertising and personalized content may *not* fall into this category.[165] Without a clear definition of "significant effects," this provision may exclude the effects that individuals may suffer from after being continuously exposed to deliberately constructed content through algorithms that favor biased or extreme material.[166]

Third, the law imposes a ban only on decisions "based solely" on automated processes, implying that algorithmic decisions may be allowed without limitations if they are not the *only* means of reaching the results.[167] The Article 29 Working Party further clarified that "based solely" means that algorithmic decisions are made without meaningful human involvement.[168] Thus, a free pass is possible when firms claim that their staff is involved in the decision-making process.[169] Although courts and enforcers have found some instances to meet the criteria set by Article 22 of the GDPR, the threshold for AI applications to fall into the category of automated decision-making remains high.[170]

Besides, algorithmic opacity can hinder effective enforcement of this provision. Firms are not obligated to explain to stakeholders whether and how

164. *Id.*; Sandra Wachter, Brent Mittelstadt & Luciano Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 INT. DATA PRIV. L. 76, 76 (2017).

165. *Guidelines on Automated Individual Decision-Making*, *supra* note 161, at 22, ("In many typical cases the decision to present targeted advertising based on profiling will not have a similarly significant effect on individuals.").

166. *But see* Gianclaudio Malgieri & Giovanni Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, 7 INT'L DATA PRIV. L. 243, 243 (2017) (arguing "the envisaged 'significant effects' on individuals (Article 22(1)) can encompass as well marketing manipulation, price discrimination, etc").

167. Wachter, Mittelstadt & Floridi, *supra* note 164, at 88; Céline Castets-Renard, *Accountability of Algorithms in the GDPR and Beyond: A European Legal Framework on Automated Decision-Making*, 30 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 91, 122 (2019).

168. *Guidelines on Automated Individual Decision-Making*, *supra* note 161, at 20–21.

169. *But see* Malgieri & Comandé, *supra* note 166, at 243 (arguing that "the threshold of minimum human intervention required so that the decision-making is 'solely' automated (Article 22(1)) can also include nominal human intervention").

170. Sebastião Barros Vale & Gabriela Zanfir-Fortuna, *supra* note 162, at 8–9. I am grateful to Gianclaudio Malgieri for this source and his insightful comments on Articles 22 and 25 of the GDPR.

their AI applications produce significant effects on an individual. Meanwhile, they are not required to explain to stakeholders how the staff conducts such algorithmic processing and to what extent the staff have exerted appropriate influence on algorithmic decisions.[171] Without disclosing such information, firms have discretion in deciding what forms of algorithmic processing qualify as automated decision-making, and they may exclude many unfair, manipulated, or inaccurate automated processing without being noticed.[172]

Additionally, Article 22 of the GDPR does not apply to three scenarios mentioned in the exception rules: situations where automated decisions are (1) "necessary for entering into, or performance of, a contract between the data subject and a data controller"; (2) authorized by "[the] law to which the controller is subject" and that adopts appropriate measures to protect the data subject; or (3) "based on the data subject's explicit consent."[173] The exception rules give firms as data controllers some leeway in automatic decision-making. For instance, firms may gain explicit consent from an individual either desperate for certain opportunities, unaware of the substantial risks caused by automatic decisions, or simply too tired to consider privacy risks due to consent fatigue.[174]

In Google's privacy policy, the firm mentions that it uses automated systems to assess user information for a more personalized user experience.[175] Although Google deploys algorithms to track user activities for customized search results, such algorithmic processing may not fall under the purview of the GDPR's automated decision-making requirements, as the firm can claim that its algorithmic decisions involve some degree of human intervention and do not produce direct legal effects on users.[176] Due to managerial invisibility and trade secrecy in Google's actual algorithmic practices, although the firm holds the algorithmic power to determine the accessibility of online resources for billions of users, users may not have an opportunity to obtain information on its manipulations or unfair algorithmic processing.[177] Even if the rest of the GDPR data protection principles are applied to algorithmic processing regardless of whether such processing qualifies as automated decision-making,[178] organizational secrecy makes it difficult for stakeholders to monitor the lawfulness of AI applications and the adequacy of associated algorithmic governance. As biased and inciting information remains pervasive in the automatic processing embedded in algorithms, individuals are still likely to

171. Bryce Goodman & Seth Flaxman, *European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"*, 38 AI MAG. 50, 56 (2017).

172. Tal Z. Zarsky, *Incompatible: The GDPR in the Age of Big Data*, 47 SETON HALL L. REV. 995, 1016 (2016).

173. General Data Protection Regulation (GDPR), *supra* note 24, at art. 22.

174. *See* McDonald & Cranor, *supra* note 158.

175. *See Google Privacy & Terms*, *supra* note 156.

176. *Id.*

177. *Id.* ("We use automated systems that analyze your content to provide you with things like customized search results, personalized ads, or other features tailored to how you use our services.").

178. Sebastião Barros Vale & Gabriela Zanfir-Fortuna, *supra* note 162, at 13–28.

experience adverse results or receive inferior services without proper stakeholder surveillance.[179] Algorithmic opacity may thus grant firms a free pass for their problematic AI systems, allowing them to replicate the discrimination, inequality, incitement, and hatred that have been deeply rooted in society for decades.

C. Increased Algorithmic Transparency and Its Inadequacy

In response to concerns regarding the invisible corporate use of big data, Articles 5(1)(a) and 12 to 15 of the GDPR establish stricter requirements for transparency, requiring firms to manage data in a transparent and fair manner, offer notices to data subjects, and give individuals access to their personal data.[180] To mitigate risks posed by opaque algorithmic decisions, Article 22 of the GDPR also requires firms to provide meaningful information on the existence of algorithmic decision-making, how data is processed, the logic involved, and its possible outcomes.[181]

There are some concerns and limitations regarding the rules on algorithmic transparency under the GDPR. First, firms are not required to disclose the existence and explanations of many AI applications that do not constitute decisions or produce legal effects, or that are not made solely by algorithms.[182] Moreover, the duty to provide information is only applicable to algorithms that feed personal data, excluding algorithms that feed anonymized data. According to the GDPR Articles 13 and 14, the transparency duty specific to automated decision-making applies "at least" to algorithmic decisions covered by the GDPR Article 22. In other words, firms operating AI in a manner that does not meet the Article 22 criteria are encouraged but not obligated to follow the GDPR transparency duties specific to automated decision-making.[183]

Second, for automated decisions that fall under the purview of Article 22, the information provision duty required by the GDPR seems to be an "*ex ante* notification" that asks for disclosure of meaningful information about the algorithmic decision-making process. According to the EDPB, meaningful information includes "the categories of data that have been or will be used,"

179. Cynthia Dwork & Deirdre K. Mulligan, *It's Not Privacy, and It's Not Fair*, STANFORD L. REV. ONLINE 35, 36 (Sept. 3, 2013).

180. General Data Protection Regulation (GDPR), *supra* note 24, at arts. 5(1)(a) & 12–15.

181. *Id.* at arts. 13, 14, 15, & 22.

182. For scholarly debate on the right to explanation under Article 22 of the GDPR, see Bryan Casey, Ashkon Farhang & Roland Vogl, *Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise*, 34 BERKELEY TECH. L.J. 143 (2019); Margot E. Kaminski, *The Right to Explanation, Explained*, 34 BERKELEY TECH. L.J. 189 (2019); Ronan Hamon, Henrik Junklewitz, Gianclaudio Malgieri & Paul De Hert, *Impossible Explanations?: Beyond Explainable AI in the GDPR from a COVID-19 Use Case Scenario*, FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency 549 (2021).

183. Maja Brkan & Grégory Bonnet, *Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions: Of Black Boxes, White Boxes and Fata Morganas*, 11 EUR. J. RISK REG. 18, 21–22 (2020).

"why these categories are considered pertinent," "how any profile used in the automated decision-making process is built," and "how it is used for a decision concerning the data subject."[184] Such information does not include periodic reporting of whether and how a firm's AI process personal data without the risk of unauthorized processing, bias, manipulation, and more. The category and detail level of the information firms should disclose to individuals is also insufficient to enable the detection of unauthorized data processing, unfairness, manipulation, incitement, or improper data management measures.[185] In practice, enforcers generally strictly interpret the provision and resist expanding transparency duties under the GDPR.[186] As individuals need additional information about algorithmic practices to assess risks involved in AI,[187] transparency obligations under current law cannot reduce the organizational secrecy of algorithmic systems effectively.[188]

Third, due to the technical complexity or inscrutability of algorithmic systems, it can be technically difficult for firms to explain the causes and correlations of many algorithmic decisions.[189] Using technical opacity in AI, firms can give technical descriptions of decision-making processes and avoid disclosure of controversial practices they want to keep secret. Firms may even avoid offering comprehensible explanations about technical aspects of AI applications, such as mathematical formulas and computing systems, as they are not required to disclose such information under existing transparency obligations.[190] For these reasons, under the GDPR, individuals cannot access comprehensive information on AI applications to reduce algorithmic opacity and further detect, contest, or challenge bias and other problems produced by AI.[191]

As external stakeholders, individuals usually struggle with insufficient transparency in AI applications, regardless of whether the algorithmic system

184. Cary Coglianese & David Lehr, *Transparency and Algorithmic Governance*, 71 ADMIN. L. REV. 1, 49 (2019).

185. *See* General Data Protection Regulation (GDPR), *supra* note 24, at arts. 5(1)(a) & 12–15. Some may argue that the general transparency requirements under the GDPR apply to any algorithmic processing, whether it qualifies as automated decision-making. However, under the general transparency duties, individuals can only obtain incomplete information through notices and access their personal data processed by AI to verify accuracy and lawfulness, as firms are merely required to describe the existence of algorithmic processing and the types of personal data handled for that processing. Sebastião Barros Vale & Gabriela Zanfir-Fortuna, *supra* note 162, at 20.

186. Sebastião Barros Vale & Gabriela Zanfir-Fortuna, *supra* note 162, at 20.

187. For a more detailed discussion of this requirement, see Andrew Burt, *Is There a "Right to Explanation" for Machine Learning in the GDPR?*, IAPP (June 1, 2017), https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/ [https://perma.cc/Z2P9-LS7F].

188. Sebastião Barros Vale & Gabriela Zanfir-Fortuna, *supra* note 162, at 19 ("When [DPA] enforce the right of data subjects to obtain transparency about the "logic involved" in qualifying ADM in specific cases, . . . [m]athematical explanations, access to algorithms, or detailed information about computing systems are never considered in these cases.").

189. Jim Shook, Robyn Smith, Alex Antonio, *Transparency and Fairness in Machine Learning Applications*, 4 TEX. A&M J. PROP. L. 443, 451 (2018).

190. Sebastião Barros Vale & Gabriela Zanfir-Fortuna, *supra* note 162, at 19.

191. Wachter, Mittelstadt, & Floridi, *supra* note 164.

meets the Article 22 criteria. For instance, according to Google's disclosed information, Google users can obtain only general information on Google's algorithmic practices. Although Google mentioned that it deploys automated systems in its privacy policy, the firm only offered a few short and vague descriptions of its use of AI.[192] Google claimed that (1) it protects user privacy by using federated learning to keep data processed through users' devices rather than uploaded to cloud spaces;[193] (2) it adopts automated systems to offer "customized search results, personalized ads, or other features tailored to how [users] use [their] services";[194] (3) it uses algorithms to "recognize patterns in data," as it uses Google Translate algorithms to identify language patterns that help users communicate across languages.[195] In these descriptions, Google establishes a broad purpose—enhancing user experience—for its use of AI without disclosing any information on the specific risks posed by its algorithmic systems, although its algorithmic practices have made it the subject of numerous investigations and litigations.[196] In the United States, Google was sued for privacy violations when its algorithms surreptitiously tracked the browsing activities of millions of users who browsed in "private" incognito mode.[197] In the EU, regulators are investigating how Google has used algorithms to track and transfer user location data to other service providers without a legitimate legal basis.[198] Despite Google's privacy disclosures, organizational opacity makes users unaware of how Google processes algorithms in a manner that potentially compromises data protection and other democratic values. Without specific transparency requirements that reveal the risks involved in AI applications and their associated management measures, algorithmic opacity prevents individuals from opening the black box of AI to avoid unlawful data processing.[199]

192. *Google Privacy & Terms*, *supra* note 156.

193. *Id.*

194. *Id.*

195. *Id.*

196. Jonathan Stempel, *Google Faces $5 Billion Lawsuit in U.S. for Tracking 'Private' Internet Use*, REUTERS (June 2, 2020), https://www.reuters.com/article/us-alphabet-google-privacy-lawsuit/google-faces-5-billion-lawsuit-in-u-s-for-tracking-private-internet-use-idUSKBN23933H [https://perma.cc/AV6X-MQB4].

197. Paresh Dave, *Google CEO Sought to Keep Incognito Mode Issues Out of Spotlight, Lawsuit Alleges*, REUTERS (Sept. 24, 2021), https://www.reuters.com/technology/google-ceo-sought-keep-incognito-mode-issues-out-spotlight-lawsuit-alleges-2021-09-24/ [https://perma.cc/8NL3-WFKZ].

198. Natasha Lomas, *Google's Location Tracking Finally Under Formal Probe in Europe*, TECHCRUNCH (Feb. 4, 2020), https://techcrunch.com/2020/02/04/googles-location-tracking-finally-under-formal-probe-in-europe/ [https://perma.cc/5RQ8-YVDE].

199. Even the EU's proposed AI Act has not required comprehensive disclosures on algorithmic practices for data protection. *See* Proposed EU AI Act, *supra* note 26, at 14–15, 17. ("Transparency obligations will apply for systems that (i) interact with humans, (ii) are used to detect emotions or determine association with (social) categories based on biometric data, or (iii) generate or manipulate content ('deep fakes').")

D. Data Protection by Design and by Default and Its Inadequacy

Through Article 25 of the GDPR, the EU has demonstrated its aim to protect personal data in the face of the AI revolution by requiring firms to consider data protection from the beginning of the AI design process.[200] Over the years, privacy and data protection by design has received much systematic scholarly attention,[201] recognizing the role of technological configurations and system design choice in the protection of data privacy.[202] The GDPR requires firms to embed data protection rules into the design of technologies and minimize the amount of data they collect and process,[203] aiming to prioritize the protection of data privacy and avoid unnecessary data processing throughout the lifecycle of business applications.[204]

However, the broad and vague standards laid out by the GDPR have made it difficult to know how this rule would be implemented on the ground.[205] Despite the vagueness of the notion of data protection by design, the EU's guidance suggests only that firms that "place privacy and data protection at the forefront of product development will be well placed to ensure that their goods and services respect the principles of privacy by design."[206] Given the lack of detailed guidance on implementation, data protection by design can involve either translating vague privacy concepts into code using technologies that reflect privacy values, adopting techniques like data pseudonymization, or slightly improving data management measures.[207] Firms are not required to disclose what kind of approach they have taken to achieve data protection by design. As one of data protection by design measures include the linking of legal competences with engineering skills, stakeholders may be confused as to how firms engineer data protection rules into AI systems, especially under changing regulatory environments.[208] Also, given the vague standard of implementing

200. General Data Protection Regulation (GDPR), *supra* note 24, at art. 25.

201. *See* WOODROW HARTZOG, PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES (2018) (discussing how the law could "require software and hardware makers to respect privacy in the design of their products") (I am grateful to Paul M. Schwartz for this source).

202. *See* ANN CAVOUKIAN, PRIVACY BY DESIGN: THE 7 FOUNDATIONAL PRINCIPLES (2011), https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf [https://perma.cc/9X7D-KL6W] (establishing a framework for privacy by design).

203. General Data Protection Regulation (GDPR), *supra* note 24, at art. 25.

204. *Id.*

205. Alex Mihaildis & Liane Colonna, *A Methodological Approach to Privacy by Design Within the Context of Lifelogging Technologies*, 46 RUTGERS COMPUT. TECH. L.J. 1, 9 (2020).

206. Article 29 Data Protection Working Party, *Opinion 8/2014 on the Recent Developments on the Internet of Things*, THE WORKING PARTY ON THE PROTECTION OF INDIVIDUALS WITH REGARD TO THE PROCESSING OF PERSONAL DATA 3 (Sept. 16, 2014), https://www.pdpjournals.com/docs/88440.pdf [https://perma.cc/Y4B5-RUST].

207. Firms will have to guess the meanings of vague concepts by themselves. *See* Connally v. Gen. Const. Co., 269 U.S. 385, 391 (1926).

208. Ari Ezra Waldman, *Data Protection by Design? A Critique of Article 25 of the GDPR*, 53 CORNELL INT'L L. J. 147, 166 (2020) ("As it is today, Article 25(1) is so vague that it cannot inform corporate behavior.").

data protection by design, firms may be allowed to adopt any privacy-enhancing technologies that are far from effective while claiming that they have met this duty. Without algorithmic disclosures that require firms to explain their data protection by design measures, organizational opacity makes it difficult for stakeholders to make firms accountable for the outcome of their data management measures.

In Google's privacy policies, the firm has offered a few positive statements regarding its use of technology in protecting data privacy.[209] Google emphasized that it adopts federated learning as a data minimization technology to protect user privacy, as federated learning can de-identify personal data and allow personal information to be processed through user devices.[210] Besides federated learning, Google has not mentioned other specific data protection by design measures for its use of AI.[211] Although Google's algorithms have been accused of intruding on users' data privacy, providing biased content, and making manipulated algorithmic decisions, the firm has not explained how its data protection by design measures address these issues.[212] As Google has promised to improve web tracking through its new privacy approaches like FLoC or Topics, the fact that advertisers are permitted to infer and store users' sensitive data repeatedly remains unmentioned in its current privacy policies.[213] At the Google I/O conference on May 19, 2021, Google's CEO, Sundar Pichai, proclaimed, "We strictly uphold responsible data practices so every product we build is private by design." He disclosed only positive things they have done without discussing major data privacy issues and associated landmark litigations against Google, hiding essentially negative issues in a black box.[214] Despite the requirement of data protection by design, stakeholders cannot assess the effectiveness of a firm's approaches to data protection. Due to organizational secrecy in data management measures and trade secrecy in AI applications, how the principle of data protection by design can reduce AI's erosion of data privacy remains in doubt. If firms are not obligated to disclose how they achieve data protection by design, algorithmic opacity allows firms to take advantage of these vague and soft rules

209. *Helping Make the Online World Safer*, GOOGLE SAFETY CTR., https://safety.google/intl/en/security/security-leadership/ [https://perma.cc/2NC5-DHRF] (last visited Sept. 2, 2022).

210. *Id.*

211. *Google Privacy & Terms*, *supra* note 156.

212. Natasha Lomas, *France Fines Google $120M and Amazon $42M for Dropping Tracking Cookies Without Consent*, TECHCRUNCH (Dec. 10, 2020), https://techcrunch.com/2020/12/10/france-fines-google-120m-and-amazon-42m-for-dropping-tracking-cookies-without-consent/ [https://perma.cc/E893-47JT].

213. Rob Pegoraro, *Google Touts 'Privacy by Design' at I/O Conference, but Privacy from Whom*?, USA TODAY (May 20, 2021), https://www.usatoday.com/restricted/?return=https%3A%2F%2Fwww.usatoday.com%2Fstory%2Ftech%2Fcolumnist%2F2021%2F05%2F20%2Fgoogle-i-o-2021-privacy-upgrades-coming-android-chrome-browser%2F5187880001%2F [https://perma.cc/B8MT- EJDJ]; Burgess, *supra* note 111.

214. Sundar Pichai, *Google I/O 2021: Being Helpful in Moments that Matter*, GOOGLE (May 18, 2021), https://blog.google/technology/developers/io21-helpful-google/ [https://perma.cc/6EPK-R5P5].

under the GDPR. Consequently, such a well-intended provision may fail to balance the commercial interests of corporations with the human rights of individuals, frustrating its original goal of setting up a human-centered and ethical AI ecosystem.[215]

IV.
TOWARD ALGORITHMIC TRANSPARENCY FOR DATA PRIVACY, HUMAN RIGHTS, AND DEMOCRACY

The EU fundamental-based approach under the GDPR regime has enforced one of the most rigorous standards for data protection. However, even with such an ambitious approach, algorithmic opacity has perpetuated information asymmetry that impedes stakeholder oversight and enforcement of the GDPR rules. Despite the GDPR's attempts to increase user control over personal data, incomplete privacy disclosures allow algorithmic opacity to hinder users' perception of risks involved in AI applications, creating a barrier to actual user control.[216] Although the GDPR requires disclosures of some aspects of automated decision-making processes, its *ex ante* notification approach that does not mandate provision of comprehensive and detailed information has perpetuated organizational invisibility, technical opacity, and legal opacity, allowing many unlawful AI applications and improper data management measures to escape stakeholder review.[217] Currently, problematic AI applications can still be shielded by opacity that does not permit outsiders to access the risks of algorithmic processes and associated data practices.[218] Accordingly, most AI systems have been developed without stakeholder surveillance in mind and remain free from law enforcement.[219] This section proposes regulatory tools for algorithmic transparency to maximize effective stakeholder oversight and law enforcement. Then, it provides assessments of their implications for policy considerations.

A. Regulatory Tools for Algorithmic Transparency

In light of the asymmetrical power relationships between firms (as data controllers) and individuals (as data subjects), regulators should consider the adoption of a complementary transparency framework, which comprises (1) corporate disclosure obligations and (2) employee reporting protections, to

215. Waldman, *supra* note 208.
216. *See supra* Part III.A.
217. Wachter, Mittelstadt, & Floridi, *supra* note 164.
218. Frank Pasquale, *Secret Algorithms Threaten the Rule of Law*, 2017 MIT TECH. REV. 1 (2017).
219. Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel F. Reidenberg, David G. Robinson & Harlan Yu, *Accountable Algorithms*, 165 U. PA. L. REV. 633, 638 (2017).

prevent algorithmic opacity from hurting data privacy and other democratic values in a business context.[220]

In recent years, scholars in the field of algorithmic transparency have proposed a number of regulatory solutions to the issue of algorithmic opacity, such as algorithmic impact assessments and audits.[221] As many scholars have argued, impact assessments can enhance accountability by increasing meaningful documentation for regulatory investigation and facilitating risk management before development of AI systems.[222] Similarly, internal and external audits can achieve accountability through audits that evaluate the quality and impact of algorithmic systems on the broader public.[223] All these proposals are useful solutions to accountability in AI, but both impact assessments and audits are rarely disclosed to important external stakeholders like shareholders, communities, and human rights advocates to assess their truthfulness and adequacy.[224] Even if some aspects of algorithmic impact assessments may be disclosed to individuals subjected to automated decision-making, stakeholders need additional information on managerial, legal, and technical aspects of algorithmic systems to assess systemic risks.

To overcome the constraints of impact assessments and auditing, this section proposes a complementary transparency regime—corporate algorithmic disclosures through sustainability reporting and whistleblowing mechanisms. Corporate sustainability disclosures reduce technical opacity, legal opacity, and

220. Jack M. Balkin, *Sidley Austin Distinguished Lecture on Big Data Law and Policy: The Three Laws of Robotics in the Age of Big Data*, 78 OHIO ST. L.J. 1217, 1228 (2017).

221. *See* ADA LOVELACE INSTITUTE, EXAMINING THE BLACK BOX: TOOLS FOR ASSESSING ALGORITHMIC SYSTEMS 7 (2020) ("There are two methodologies that have seen wide reference in popular, academic, policy and industry discourse around the use of data and algorithms in decision making: algorithm audit and algorithmic impact assessment."); Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 115 (2019) (a proposed human impact statement); Margot E. Kaminski, *Understanding Transparency in Algorithmic Accountability*, *in* CAMBRIDGE HANDBOOK OF THE LAW OF ALGORITHMS 121, 121–22 (2020).

222. Andrew D. Selbst, *An Institutional View of Algorithmic Impact Assessments*, 35 HARV. J.L. & TECH. 117, 122 (2021) (discussing the goals of algorithmic impact assessments).

223. Casey, Farhang & Vogl, *supra* note 182; Inioluwa Deborah Raji, Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron & Parker Barnes, *Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing*, ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY 2.4 (2020).

224. Article 29 Data Protection Working Party, *Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is 'Likely to Result in a High Risk' for the Purposes of Regulation 2016/679*, EUR. COMM'N (Apr. 4, 2017, as last revised and adopted on Oct. 4, 2018), https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47711 [https://perma.cc/U6D2-V9ZR] ("Publishing a DPIA is not a legal requirement of the GDPR."); Katharine Miller, *Radical Proposal: Third-Party Auditor Access for AI Accountability*, STANFORD HAI (Oct. 20, 2021), https://hai.stanford.edu/news/radical-proposal-third-party-auditor-access-ai-accountability [https://perma.cc/42RM-LRUV] ("[I]nternal audits are rarely publicized[.]"); Selbst, *supra* note 222, at 117 ("In practice, an impact assessment framework relies on the expertise and information to which only the creators of the project have access."); Mona Sloane, *The Algorithmic Auditing Trap*, MEDIUM (Mar. 17, 2021), https://onezero.medium.com/the-algorithmic-auditing-trap-9a6f2d4d461d [https://perma.cc/SVP9-76Y2].

managerial invisibility by requiring firms to disclose managerial, legal, and technical aspects of algorithmic practices under a series of topics. Whistleblowing systems offer internal oversight by allowing employees to assess the truthfulness of disclosures and reveal the unlawfulness of corporate behaviors that have not been disclosed by firms. As explained below, these recommended transparency rules can be useful in reducing algorithmic opacity to facilitate stakeholder oversight, risk assessments, and enforcement of law, ultimately safeguarding the protection of data privacy, human rights, and democratic norms.

1. Corporate Disclosure Requirements

Recent evidence finds that around half of large-sized companies in Europe have been using or plan to use AI technologies.[225] Yet, the opacity of their use of algorithmic systems and associated data practices have indicated many data protection issues, causing growing concerns among citizens, investors, and consumers.[226] To truly enhance data privacy in the shadow of algorithmic opacity, regulators should consider integrating algorithmic disclosures into sustainability disclosure duties because commercial applications of AI has posed unprecedented threats to fundamental values, and society increasingly demands disclosure regulations for private accountability.[227] Absent algorithmic disclosures, individuals cannot access comprehensive information on algorithmic practices that enables them to detect systematic risks before exercising their individual rights; shareholders are less likely to detect financial risks before making investment decisions; regulators can have a harder time detecting rule-breaking behaviors and launching an investigation—all of these can hinder enforcement of data protection rules and other democratic norms.

As a beacon of valuable legislation that leads countries worldwide in safeguarding data privacy and human rights, the EU should be one of the first legal systems to consider sustainability reporting as a way to require algorithmic disclosures. At the EU level, the institution's Non-Financial Reporting Directive (NFRD) requires mandatory disclosures in corporate annual reports on social

225. This survey was conducted by Ipsos for the European Commission. *See European Enterprise Survey on the Use of Technologies Based on Artificial Intelligence*, EURO. COMM'N 4, 6 (2020), https://www.ipsos.com/sites/default/files/ct/publication/documents/2020-09/european-enterprise-survey-and-ai-executive-summary.pdf [https://perma.cc/5ZA7-GV5S] (In 2020, "42% of businesses hav[e] adopted at least one of these ten AI technologies"; "18% of enterprises that do not currently use AI plan to adopt at least one technology in the next two years").

226. *Artificial Intelligence and Privacy*, DATATILSYNET (2018), https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf [https://perma.cc/26W6-US4Z].

227. *The Non-Financial Reporting Directive: What You Need to Know*, DATAMARAN (2022), https://www.datamaran.com/non-financial-reporting-directive/ ("The number of regulatory initiatives requiring non-financial disclosure is growing rapidly. From 2013 to 2018, there has been a 72% increase in the number of recorded regulations concerning non-financial issues. And this trend looks set to continue.").

matters regarding "(1) environmental, social and employee, (2) respect for human rights, and (3) anti-corruption and bribery" to promote private accountability and sustainable businesses.[228] Issues concerning privacy and data protection—fundamental rights recognized in the EU context—can fall within the Directive's purview, as they can be considered to fit into the category of "respect for human rights."[229] Although the NFRD will be amended by the proposed Corporate Sustainability Reporting Directive (CSRD), the disclosure topics required by the CSRD are largely based on the NFRD. Most importantly, the NFRD's disclosure rules provide by far one of the most adopted sets of guidelines and useful insights for regulators who might envision suggested algorithmic sustainability disclosures.[230] Meanwhile, because the EU is faced with rising AI economic development and unprecedented data privacy issues caused by algorithmic systems, policymakers might want to monitor firms' data practices through full-fledged requirements, including disclosure obligations on commercial applications of AI. The following section proposes a set of disclosure duties designed to mitigate the negative effects of algorithmic opacity inherent in AI systems for policymakers in the EU, the United States, and elsewhere.

a. Principles of Disclosures on AI Systems

According to the Guidelines on the NFRD, large public interest entities with more than 500 employees must periodically disclose non-financial information based on a materiality standard to stakeholders in a balanced, comprehensive, forward-looking, and consistent manner.[231] When certain kinds

228. Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups, 2014 O.J. (L 330/1) preamble paras. 2, 3, 4, 6, 18 and arts. 1(1) & 1(3) (hereinafter Non-Financial Reporting Directive); Virginia Ho & Stephen Park, *ESG Disclosure in Comparative Perspective: Optimizing Private Ordering in Public Reporting*, 41 U. PA. J. INT. L. 249 (2019).

229. Non-Financial Reporting Directive, *supra* note 228, at art. 19a.1; The rights to privacy and data protection are both considered fundamental rights in the EU. *See Data Protection*, *supra* note 30.

230. On April 21, 2021, the EC proposed the CSRD, which will amend the NFRD with more detailed disclosure duties. *See Proposal for a Directive of the European Parliament and of the Council, amending Directive 2013/34/EU, Directive 2004/109/EC, Directive 2006/43/EC and Regulation (EU) No 537/2014, as Regards Corporate Sustainability Reporting*, COM (2021) 189 final (Apr. 21, 2021), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0189&from=EN [https://perma.cc/6CV7-HPSP].

231. EUR. COMM'N, GUIDELINES ON NON-FINANCIAL REPORTING (METHODOLOGY FOR REPORTING NON-FINANCIAL INFORMATION 6–9 (3.2–3.6) (2017) (hereinafter GUIDELINES ON NON-FINANCIAL REPORTING). The NFRD sets a minimum standard for the applicable scope of mandatory disclosure, allowing countries to establish stricter standards for corporate disclosures. As an illustration, Sweden requires firms of all types with more than 250 employees to carry out NFRD disclosure duties. SWEDISH AGENCY FOR GROWTH POLICY ANALYSIS, FROM VOLUNTARY TO MANDATORY SUSTAINABILITY REPORTING 13 (2018) ("The Swedish application is therefore broader than the minimum levels required in the NFR Directive, partly in that it covers all companies with over 250

of business applications can bring about massive negative social and financial impacts on communities, the EC establishes additional disclosure guidelines for such applications. For instance, the Commission published climate change disclosure guidelines in 2019 in response to the grave and damaging effects of climate change produced by commercial activities.[232] As business applications of AI systems can also incur significant social risk to human rights and pose financial risk to corporate sustainability, especially when fines could run into the millions if not billions of euros, policymakers should soon establish guidelines for algorithmic disclosures that apply to large firms to proactively address the long-term AI risks. Below are suggested algorithmic disclosure guidelines designed to address problems posed by algorithmic opacity.

i. Materiality Standard for Algorithmic Disclosures

Pursuant to the NFRD, materiality refers to information that is necessary for stakeholders to understand a firm's "development, performance, position, and the impact of its activities."[233] In contrast to the U.S. investor-centric approach, the EU definition of materiality takes both investors and stakeholders into account.[234] According to the NFRD, a firm is required to disclose information that is considered material.[235] Under this materiality standard, materiality implies a financial dimension and a social dimension. Financial materiality refers to a firm's "development, performance [and] position" that affects its commercial value and thus being of most interest to shareholders who care about the return on their investment.[236] When AI revolutions bring about changes in markets and regulations, the use of AI will create commercial risks and opportunities that are deemed financially material. Social materiality denotes the external "impact of [the firm's] activities," which is of interest to stakeholders such as consumers, shareholders, and business partners.[237] As firms update AI business models, their business practices will produce effects on individual fundamental rights like privacy and data protection that are generally recognized as a material concern in society. When firms decide the materiality of certain business practices, they must evaluate both financial and social effects

employees – i.e. half of the total stated in the directive – and partly in that the reporting requirement applies to all companies and not just listed companies or certain financial institutions.").

232. EUR. COMM'N, GUIDELINES ON NON-FINANCIAL REPORTING: SUPPLEMENT ON REPORTING CLIMATE-RELATED INFORMATION, EUR. UNION (2019), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019XC0620(01)&from=EN [https://perma.cc/5YBK-R7LQ].

233. GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 14.

234. For the U.S. investor-centric materiality standard, see TSC Indus. v. Northway, Inc., 426 U.S. 438, 449 (1976). For the EU definition of materiality, see GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 14.

235. GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 2.2 (organizations must "disclose information . . . to the extent that such information is necessary for an understanding of the company's development, performance, position, and impact of its activities").

236. *Id.*

237. *Id.*

of AI applications on legal compliance for their stakeholders.[238] Given the far-reaching financial and social impacts of AI, corporate use of AI systems should constitute a material issue by default.

ii. The Principle of Comprehensible Algorithmic Disclosures

As discussed previously, technical opacity involved in algorithms can make dynamic, non-intuitive, and complex algorithmic processes and outcomes difficult to explain.[239] To reduce technical opacity for stakeholders, firms must provide adequate descriptions of AI techniques they use and explanations of their AI applications for stakeholders without technical expertise in AI. With comprehensible algorithmic disclosures, understanding the input-output behavior of algorithms and AI can empower stakeholders to improve them.[240] Thus, to follow the principle of comprehensible disclosures, firms must explain the design concepts and performances of AI systems for stakeholders who need to measure the risks posed by AI.[241] By doing so, the law could help stakeholders attain a broader perspective on AI applications and data governance through comprehensible explanations, reducing opacity to the extent possible.[242]

iii. The Principle of Minimum-Necessary Algorithmic Disclosures

Considering the tension between public transparency and commercial secrecy, firms must adhere to the principle of minimum necessary disclosure, whereby only information on critical aspects of how firms are using AI systems has to be disclosed.[243] Although disclosures can lead to greater transparency and accountability, complete disclosure is not a realistic approach to achieving data privacy protection.[244] Since algorithms consist of public and personal data, complete transparency will adversely affect other legal interests such as trade secrets, data privacy, and security, especially in situations where AI systems are used for services that need some extent of seclusion.[245] For instance, if a firm is required to disclose the source code of its algorithms for stakeholder review, such public disclosure is likely to compromise the trade secrecy of the algorithms. In this vein, complete transparency can be abused by opportunistic competitors for

238. *Id.*

239. *See* Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement*, 69 FLA. L. REV. 181, 188 (2018).

240. Mukund Sundararajan, Ankur Taly & Qiqi Yan, *Axiomatic Attribution for Deep Networks*, *in* PROCEEDINGS OF THE 34TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING (2017).

241. Goodman and Flaxman, *supra* note 171.

242. *Id.*

243. For a discussion of minimum necessary disclosure, see Lu, *supra* note 27, at 135.

244. Deven R Desai & Joshua A Kroll, *Trust but Verify: A Guide to Algorithms and the Law*, 31 HARV. J.L. & TECH. 1, 10 (2017) ("[S]imply disclosing or open-sourcing source code does nothing to show that the disclosed software was used in any particular decision unless that decision can be perfectly replicated from the disclosures.").

245. PASQUALE, *supra* note 3, at 142.

their own interests.[246] To avoid a complete disclosure regime, only in cases involving public concerns should firms be required to disclose some aspects of data practices and AI applications for stakeholder scrutiny.[247] Meanwhile, minimum disclosure can help prevent information overload, one of the causes of the failure of privacy policies.[248] The issue of information overload can be seen in the U.S. securities law context, where investors are sometimes offered far more information than what is required, which confuses them with an overload of irrelevant materials.[249] Similarly, in the space of data privacy, many privacy policies are loaded with lengthy and vague or obscure descriptions that hide problematic data practices in later paragraphs to deceive readers into thinking that they are being properly informed.[250] To enhance the effectiveness of disclosures, the principle of minimum necessary disclosure requires firms to describe their algorithmic practices and data management approaches succinctly, along with devoting assiduous attention to matters of data privacy concerns.

iv. The Principle of Double-Layered Algorithmic Disclosures

Since disclosures may risk hurting trade secrecy, regulators can reckon with marking AI systems based on different disclosure duties such as (1) complete disclosure to a competent authority, when full transparency is necessary to understand a risky algorithmic system's operating environment, such as disclosing its source code or databases, in order to replicate specific algorithmic decisions and clearly explain an AI system's misbehavior;[251] or (2) partial disclosure with comprehensible and minimum information released to stakeholders periodically.[252] To protect trade secret interests, complete disclosure must only apply to situations where complete disclosure of AI is necessary for investigative purposes. Unlike complete disclosure, partial disclosure requirements must apply to large-sized firms that develop AI-based

246. CHRISTIAN SANDVIG, KEVIN HAMILTON, KARRIE KARAHALIOS & CEDRIC LANGBORT, AUDITING ALGORITHMS: RESEARCH METHODS FOR DETECTING DISCRIMINATION ON INTERNET PLATFORMS 9 (2014).

247. *Id.*

248. *See* Joel R. Reidenberg, N. Cameron Russell, Alexander J. Callen, Sopfia Qasir & Thomas B. Norton, *Privacy Harms and the Effectiveness of the Notice and Choice Framework*, 112 J. L. & POL'Y INF. SOC. 486, 492 (2014).

249. Jillian Loh, *Could the Pay Ratio Disclosure Backfire? Examining the Effects of the SEC's Pay Ratio Disclosure Rule*, 4 TEX. A&M L. REV. 417, 438–39 (2017) ("Disclosures are less effective when investors become overloaded with extraneous information that is not useful--making it difficult and confusing for investors to identify the important information about a company. For example, a study of institutional investors conducted by Stanford University in 2015 revealed that the majority found proxy statements to be too long and difficult to read, and only a third of the information disclosed was relevant.").

250. Daniel J. Solove, *Introduction: Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880, 1888–89 (2013).

251. *See* Stadish v. Superior Court, 84 Cal. Rptr. 2d 350, 359 (Cal. Ct. App. 1999) (illustrating the use of protective orders in trade-secret-related litigations); Desai and Kroll, *supra* note 244, at 39.

252. Non-Financial Reporting Directive, *supra* note 228.

services or products.[253] Large-sized firms are not only equipped with more resources to develop, profit from, and produce problems through AI, but also more likely to bear the cost of disclosure duties. Therefore, they should be held to a high standard of accountability for their algorithmic practices. This principle enables stakeholders to monitor firms through disclosed information without compromising the core secrecy of AI systems and associated data practices.

b. Topics in Disclosure on AI Systems

According to the NFRD, firms should discuss five topics for stakeholder review, which include (1) business models; (2) policies and due diligence processes; (3) outcome of policies; (4) principal risks and their management; and (5) key performance indicators.[254] For each topic, the following section proposes mandated disclosures for large-sized firms that develop AI-based services or products.[255]

i. Business Model Disclosures

Based on the NFRD, business model disclosures discuss how a firm generates lasting corporate value through its commercial activities.[256] Under the NFRD, disclosures of business models involve descriptions of a firm's products and services, competitive market, business strategies, organizational structure, and factors in future development.[257] In generic terms, business model disclosures involve how a firm operates its business in a given environment.[258] Currently, firms are not required to disclose the role of AI in their business models. In Google's 2021 NFRD report, the firm did not describe how algorithmic systems were considered in its business models, although its privacy policy mentioned that Google's services deploy algorithms and automated systems.[259]

In an algorithmic-disclosure context, business model disclosures provide basic knowledge to help stakeholders understand what kinds of AI-based services and products have been deployed by a firm. Such information enables stakeholders to be aware of AI applications that have material financial and social impacts on the firm. Therefore, firms should be required to explain the role of AI in their business models, as well as the external impacts of their business models for AI applications. The rigid rules of data protection and hefty

253. Kaminski, *supra* note 182, at 210.
254. *Id.*
255. GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 4.
256. *Id.* at 4.1.
257. Hess, *supra* note 129, at 49.
258. *Id.*; GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 4.1.
259. *See* GOOGLE, 2021 EUROPEAN UNION NFRD REPORT (2021), https://kstatic.googleusercontent.com/files/119a1ae1ee72b369468e9429fdc46153ea91af5d09cdff7347115461654bbae3ed5c733c8ea86f84c0b8cf3feddb6972ed8dab82e87f36f46874ea9d6b045be1 [https://perma.cc/R5XN-JTLU].

fines for violations under the GDPR also make it vital for firms to disclose how their business models adhere to data protection rules, and how those rules shape their business models for AI applications. Without such information, stakeholders cannot identify what kinds of AI applications they should monitor proactively, allowing algorithmic opacity to hide risky applications that further erode data privacy, human rights, and corporate interests.

In algorithmic disclosures, business model disclosures should require firms to describe AI applications that are socially and financially material to stakeholders. From the perspective of social materiality that considers the external impacts of AI applications on communities, firms should describe what AI techniques they have used, if and how their business models have complied with existing regulations, and whether there have been any damaging impacts of their business models on stakeholders, including users and shareholders.[260] Specifically, firms must be required to disclose their AI-based services and products, as well as associated AI techniques and data management measures, which include how they obtain, store, process, transfer, manage, and delete data for AI applications. When data is fed to AI and used for predictive analysis or algorithmic decisions that reveal or influence important parts of individuals' private lives, firms must also report whether their AI systems are working under applicable data privacy laws and AI regulations. In addition, firms should explain whether their AI systems are making biased, erroneous, inciting, or arbitrary decisions that may cause unfairness or disorder in society.[261] Firms making automated decisions must also explain the source of data, what categories of data are used and why, the rules of data selection and training processes, the variables considered for algorithmic decisions, the reasoning behind validated decisions, and an evaluation of anticipated harmful effects on users, not only in cases having legal effects, but also in those related to data privacy and democratic concerns.[262]

From the perspective of financial materiality, firms should be required to illustrate the financial effects of their business models, such as how their AI-based products and services have contributed to revenue and how regulatory burdens have influenced their business operations. For instance, firms must be required to explain changes in their business models made to gain access to new markets and to seize commercial opportunities related to AI. Also, firms must describe their dependencies on AI-based services and products and how they make changes in business models to comply with existing regulations and generate profits. To safeguard corporate sustainability, firms must show how

260. For proposed algorithmic disclosures on business descriptions, see Lu, *supra* note 27, at 138.

261. For the far-reaching effects of decision-making algorithms applied in other contexts, see Jessica Eaglin, *Constructing Recidivism Risk*, 67 EMORY L.J. 59 (2017).

262. *Data Protection Impact Assessment (DPIA)*, GDPR.EU (2018), https://gdpr.eu/data-protection-impact-assessment-template/ [https://perma.cc/XMR7-W4SD]; *See also* Selbst, *supra* note 222, at 139–52 (discussing the elements of AI impact assessments).

their business models for AI applications ensure compliance with all applicable laws over the long term.

ii. Policy and Due Diligence Process Disclosures

Under the NFRD, the topic of policy and due diligence process means a firm's practical approaches to main issues, goals, and plans of action.[263] Under this topic, firms usually describe their governance measures like management and board duties, resource allocations, and board oversight.[264] Disclosures of policy and due diligence processes enable the inspection of a firm's capabilities of handling issues that imply public concerns.[265] According to the NFRD, policy and due diligence information can represent both socially and financially material information for the governance and control system of a firm.[266] Such information can advance stakeholders' understanding of the robustness of a firm's plan to tackle managerial challenges and its commitment to reduce operational risks, including data privacy violations.[267] Pursuant to the NFRD, firms are not presently compelled to disclose their AI-related policies and due diligence processes. In Google's 2021 NFRD Report, the firm did not describe how its current corporate policies consider data privacy in detail.[268] Google said it amended its privacy policy *in 2018* to strengthen user control through simplified account setting and restated it would commit to compliance with the GDPR.[269] As for its managerial approaches to issues posed by AI, Google mentioned that it published Google AI Principles in June 2018 to guide its product designs and policies.[270] These descriptions are restatements of the policy it developed three years ago, rather than explanations about updated policies and due diligence processes it adopted to reduce present and imminent danger caused by AI applications. Moreover, the report did not disclose how Google's corporate policies consider the social and financial risks posed by AI and how Google could systematically implement policies to control privacy risks through its governance system. Based on the NFRD Report, stakeholders cannot understand Google's main AI challenges, responding policies, plans of actions, allocation of resources and personnel to implement policies, and its implementation of AI principles to safeguard data privacy and other fundamental values.[271]

263. GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 10–11.

264. *Id.*

265. *See id.*

266. *See* OFFICE OF THE HIGH COMMISSIONER, THE UNITED NATIONS GUIDING PRINCIPLES ON BUSINESS AND HUMAN RIGHTS 17 (2011).

267. *See* GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 10–11.

268. *See* GOOGLE, 2021 EUROPEAN UNION NFRD REPORT, *supra* note 259.

269. *Id.* at 10.

270. *Id.* at 11.

271. *Id.* at 10–11. *But see* Kent Walker, *Google AI Principles Updates, Six Months*, GOOGLE (Dec. 18, 2018), https://www.blog.google/technology/ai/google-ai-principles-updates-six-months/ [https://perma.cc/QC6M-UQ6P].

In algorithmic disclosures, policy and due diligence process disclosure can help stakeholders assess the appropriateness of a firm's AI governance and data protection by design measures. In this context, firms should be required to describe how their AI applications consider data protection in policies, due diligence plans, and operational decision-making processes. Such disclosures can reduce organizational secrecy and legal opacity that shield a firm's AI-driven data management measures, organizational structure, and governance systems from stakeholder scrutiny. To inform stakeholders on the level of the firm's awareness of and dedication to AI-related issues, firms must describe the role of the board and management in addressing legal issues derived from AI systems. Additionally, firms must explain how they allocate resources and collaborate with consumers, staff, developers, trainers, and other suppliers of AI systems to ensure data protection.[272] Firms must disclose not only their use of privacy-enhancing technologies, but also how their data protection by design approaches can contribute to compliance in their deployment of AI.[273] Otherwise, stakeholders cannot understand whether the firm adopts proper data management measures to handle anticipated business, legal, and technical risks derived from AI applications. Last but not least, firms conducting algorithmic decisions must explain how their staffs participate in surveilling the process of automated decision-making. If not, firms may avoid the application of Article 22 of the GDPR without being noticed. These disclosures will largely reduce algorithmic opacity disguised as managerial invisibility and legal opacity, deepening stakeholders' understanding of how firms periodically reform their management measures and organizational structure to prevent AI from eroding data protection, autonomy, fairness, and other democratic values.[274]

iii. Outcome Disclosures

Under the NFRD, outcome refers to the results of implementation of corporate policies and due diligence processes.[275] Outcome disclosures can provide an overview of the strengths and weaknesses of business operations.[276] Under this topic, firms should describe the social performance of corporate operations under existing management measures.[277] For now, the NFRD does not require disclosures of the effectiveness of privacy and AI policies. Firms can hide the inadequacy of their policies and due diligence processes in addressing data privacy issues raised by AI until legal proceedings occur. As an illustration,

---

272. Scherer, *supra* note 33, at 369 (a description of the problem of discreteness in AI systems).

273. *See Summary of Privacy Enhancing Technologies–A Survey of Tools & Techniques*, CRANIUM (2018), https://www.craniumusa.com/summary-of-privacy-enhancing-technologies-a-survey-of-tools-and-techniques/ [https://perma.cc/3DLT-KNFU].

274. For how the courts expand the concept of accountability to stakeholders that may be harmed by corporate products, see generally, Balkin, *supra* note 220.

275. GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 10; *id.*

276. GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 12.

277. *Id.*

Google's 2021 NFRD report mentions nothing about the results of implementing its privacy policies.[278] Although Google's privacy policy states that it values user privacy, consumers and regulators have consistently accused the firm's AI-based online services of violating data privacy regulations. However, under the NFRD, one has no way of understanding the outcomes of privacy policies and due diligence processes for AI to prevent invasive, unfair, and harmful AI applications from materializing.

As neither the GDPR nor the NFRD require mandatory disclosure on the outcomes of AI-related policies, in algorithmic disclosures, firms should be obliged to explain the outcome and impact of privacy policies on business applications of AI. Without such information, stakeholders cannot know whether a firm's algorithmic governance has enabled it to comply with existing regulations and achieve sustainable profit levels. Considering social and financial materiality, firms must describe the financial and social effects of privacy policies, including the policies' resulting contributions to AI business models to revenue and market competitiveness, as well as their consequent contributions to compliance with existing regulations. Specifically, firms must explain the effectiveness of their privacy policies and data management measures, including their adoption of data protection by design approaches against the policy goals of compliance.[279] Otherwise, stakeholders cannot know which policies and due diligence processes need further improvement to achieve desirable financial and social outcomes. Learning from the U.S. legal proceeding disclosures under Regulation S-K,[280] firms must disclose data privacy and AI-associated legal proceedings, which reflect negative financial and social outcomes of data management measures that require careful stakeholder monitoring for further improvement. As mentioned earlier, up to May 15, 2022, statistics found that European data regulators have imposed at least 1,079 fines on firms for GDPR breaches in the amount of €1.62 billion.[281] While many fines are still pending,[282] legal proceedings can be a key factor for stakeholders to decide the actual financial and social outcomes of privacy policies.[283] Firms must thus discuss how legal proceedings derived from their use of AI affect financial

278. *See* GOOGLE, 2021 EUROPEAN UNION NFRD REPORT, *supra* note 259.

279. *See generally* OFFICE OF THE PRIVACY COMMISSIONER OF CANADA, PRIVACY ENHANCING TECHNOLOGIES–A REVIEW OF TOOLS AND TECHNIQUES (2017), https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2017/pet_201711/ [https://perma.cc/T675-4CWB].

280. Energy Policy and Conservation Act of 1975 – Regulation S-K, 17 C.F.R. § 229.103 ("Describe briefly any material pending legal proceedings, other than ordinary routine litigation incidental to the business, to which the [firm] or any of its subsidiaries is a party or of which any of their property is the subject.").

281. *GDPR Enforcement Tracker*, *supra* note 22.

282. George Leopold, *Busted: Google Remains Atop List of GDPR Fines*, DATANAMI (Mar. 10, 2020), https://www.datanami.com/2020/03/10/busted-google-remains-atop-list-of-gdpr-fines/ [https://perma.cc/N5FU-9Y2X].

283. *Cf.* Lu, *supra* note 27, at 141.

performance and describe compliance measures to hold themselves more accountable for their algorithmic practices.

iv. Principal Risks and Their Management Disclosures

Pursuant to the NFRD, risk management represents a firm's approaches to continuously identifying and mitigating risks that may endanger business operations.[284] Under this topic, firms should illustrate how major risks are likely to influence their stakeholders and explain how they measure and minimize these risks.[285] Risks may relate to internal factors like corporate services and strategic decisions, or external factors like regulatory environments.[286] Currently, firms are not compelled to disclose data privacy risks caused by their use of AI under the NFRD. Google's 2021 NFRD Report, as an illustration, did not consider any data privacy risks as principal risks. [287] Therefore, stakeholders cannot adequately inspect the major risks posed by AI systems through NFRD disclosures and GDPR transparency duties, although AI applications have caused increasing legal and managerial issues that affect a firm's survival.

In algorithmic disclosures, principal risk disclosures can improve stakeholders' assessments of various risks derived from algorithmic practices. Principal risks can arise from the dynamic unpredictability of algorithms and the regulatory environment.[288] For algorithmic disclosures, firms should be required to consider both social risks and financial risks. Social risks refer to the negative social impacts of algorithmic systems on communities such as tracking algorithms that secretly track user behavior to collect, process, and transfer a large amount of location information without legal basis. Financial risks refer to adverse financial influence of AI on a business, such as applicable regulations and corporate policies that increase the cost of data management measures. Both forms of risk may arise from business applications of AI and influence other parts of the value chain. Given AI's potential to enlarge the erosion of data privacy and other democratic values, lawmakers should require firms to disclose the principal risks caused by AI applications and the affected populations. Otherwise, for fear of reputational harm and regulatory investigation, firms may tend not to disclose major risks and risk management approaches, as Google did in its 2021 NFRD Report.[289]

To enable stakeholder assessments of a firm's major risks and risk management approaches, firms must describe how they measure and manage risks derived from changes in regulations of or advances in AI systems.[290]

284. *See* GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 12–13.
285. *See id.*
286. *Id.*
287. *See* GOOGLE, 2021 EUROPEAN UNION NFRD REPORT, *supra* note 259.
288. Tschiderd, *supra* note 41, at 705–706.
289. *See* GOOGLE, 2021 EUROPEAN UNION NFRD REPORT, *supra* note 259.
290. *Cf.* Lu, *supra* note 27, at 145–46.

Absent comprehensive information about principal risks and their management, individuals subject to AI applications cannot assess data privacy risks in a systematic manner before giving consent; shareholders cannot assess financial risks to make investment decisions; regulators may be slow to identify improper or unlawful algorithmic practices in order to initiate an investigation. To maximize effective risk assessment for enforcement of data privacy law, firms must explain a number of risk factors such as (1) how they define data privacy and human rights risks associated with AI, (2) what corporate practices may foreseeably cause data privacy violations, (3) what are the principal AI-caused financial and social risks the firms have identified over the short, medium, and long term, (4) how have previous and ongoing incidents related to AI and corresponding legal proceedings financially and socially affected the firms, shareholders, consumers, and citizens, (5) how have firms reduced such risks or taken measures to prevent them from materializing and reoccurring, (6) how have firms developed their AI-based business applications to avoid data privacy invasions, bias, inaccuracy, manipulation, misinformation, and incitement, and (7) the frequency of reviews and analyses with regard to risk identification and assessment for AI systems.[291] As discussed in Part III, the GDPR's transparency rules are still inadequate to reduce algorithmic opacity that hinders risk detection and stakeholder surveillance. The proposed disclosure items under this topic provide stakeholders with additional information to conduct risk assessments and monitor corporate algorithmic practices.

v. Key Performance Indicator Disclosures

According to the NFRD, key performance indicators (KPIs) are metrics for evaluating how well a firm makes certain progress and achieves its goals.[292] Effective KPIs are precisely defined and highly relevant to a firm's primary goals. With KPIs, stakeholders can efficiently track corporate performance through quantifiable indicators to better understand the strengths and vulnerabilities of their adopted policies. Then, stakeholders can communicate with firms regarding areas that require improvement.[293] Pursuant to the NFRD, firms must publish KPIs related to their performance on social issues in their sustainability disclosures that concern outcomes of policies and risk management. Under the NFRD, however, firms are not compelled to disclose their KPIs with respect to data privacy or AI issues. According to Google's 2021 NFRD Report, the firm did not mention any statements regarding KPIs relevant to its use of AI.[294] Without useful indicators, stakeholders cannot effectively assess and compare the strengths and weaknesses of adopted policies through the NFRD disclosures to hold firms accountable for their AI applications.

291. *See id.*
292. GUIDELINES ON NON-FINANCIAL REPORTING, *supra* note 231, at 13–14.
293. *See id.*
294. *See* GOOGLE, 2021 EUROPEAN UNION NFRD REPORT, *supra* note 259.

In algorithmic disclosures, firms should be required to disclose KPIs that can be used to measure the outcomes of their privacy policies, due diligence processes, and risk management measures. When many firms adopt the same KPIs, it can be useful to facilitate comparability of corporate impacts and performance across firms. To promote stakeholder oversight for data privacy and other democratic values endangered by AI systems, effective KPIs should include at least the following items: (1) proportion of AI-based business models, (2) proportion of AI-based business revenues, (3) proportion of automated decision-making processes, (4) proportion of privacy personnel involved in the operation of AI business applications, (5) proportion of staff involved in automated decision-making processes, (6) proportion of privacy-enhancing technologies involved in AI-based services or products, (7) legal proceedings including national, European, and international litigations and investigations concerning data privacy violations, (8) number of internal reporting cases concerning potential data privacy violations, (9) response time to handle reported data privacy incidents, (10) Data Protection Impact Assessment completion rate,[295] and (11) number of third-parties involved in the developing of AI.[296] The above metrics should be subject to change to continuously represent operational progress, stakeholder needs, and commercial interests to achieve the goal of data privacy and other democratic values.

2. Corporate Whistleblowing Mechanisms

Through the lens of sustainability reporting, regulators can require firms to disclose information about the core managerial, technical, and legal aspects of their use of AI that involves data privacy and other democratic concerns. However, some may argue that firms may only disclose positive information to the public.[297] To prevent this problem, an employee reporting regime—corporate whistleblower regulations—can enhance the accuracy, comprehensiveness, and truthfulness of corporate disclosures on AI. Below, this section explains how a robust whistleblowing regime can enable employees to take part in monitoring algorithmic practices and enhancing the truthfulness of algorithmic disclosures.

a. Whistleblowing Protection in the EU

Corporate whistleblowing refers to the reporting of individuals on unlawful activities occurring in a work environment that pose threats to the public interest.[298] The reporting individuals, termed as whistleblowers, can play a vital

295. *Data Protection Impact Assessment (DPIA)*, *supra* note 262.

296. *See* Philip Brudney, *What Are Your Privacy KPIs?*, RISK3SIXTY (Nov. 4, 2019), https://risk3sixty.com/2019/11/04/what-are-your-privacy-kpis/ [https://perma.cc/8U9H-8RG2].

297. Rüdiger Hahn & Regina Lülfs, *Legitimizing Negative Aspects in GRI-Oriented Sustainability Reporting: A Qualitative Analysis of Corporate Disclosure Strategies*, 123 J. BUS. ETHICS 401, 409–13 (2014).

298. *See* Richard Calland & Guy Dehn, *Introduction*, *in* WHISTLEBLOWING AROUND THE WORLD 9 (2004).

role in promoting social welfare by disclosing breaches of law.[299] In recent years, the outbreaks of many scandals, such as the Luxembourg Leaks,[300] the Facebook-Cambridge Analytica Scandal,[301] the Paradise Papers,[302] and the latest Facebook Papers Leaks, were all revealed to the public by whistleblowers, making regulators aware of the importance of initiating whistleblower legislation at the EU level.[303] In recognition of the value of whistleblower law and uneven whistleblower legislation across Member States, the EU adopted a Whistleblower Directive that came into force in December 2019.[304] The Directive offers a broad definition of whistleblowers by which employees, self-employed individuals, and shareholders in work scenarios who have "reasonable grounds to believe that the information on breaches reported was true" can receive protections against corporate retaliation for reporting corporate wrongdoing.[305] The Whistleblower Directive requires Member States to set up a three-tiered reporting system composed of internal corporate reporting, external government reporting, and public disclosures for whistleblowers to disclose illegal activities.[306] Whistleblowers are encouraged first to report through internal reporting channels within firms,[307] though they can also report directly through external reporting channels that point to relevant competent authorities.[308] If neither internal nor external reporting channels give a timely response to the whistleblower, or if the whistleblower reasonably believes the violation involves imminent public danger, they can legitimately reveal the information to the public through the media.[309] The Whistleblower Directive protects whistleblowers from any type of harassment, threat, or retaliation by employers, including suspension, dismissal, demotion, or transfer by firms,[310]

299. *Id.*

300. *See* Christian Oliver, *EU Tax: Tough Love for Multinationals' Sweetheart Deals*, FIN. TIMES (July 13, 2015), https://www.ft.com/content/32e6a5c4-1a80-11e5-a130-2e7db721f996 [https://perma.cc/4G6Z-R2ZF].

301. *See* Nicholas Confessore, *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*, N.Y. TIMES (Apr. 4, 2018), https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html [https://perma.cc/HS9T-F6UP].

302. *See Key Revelations from the Paradise Papers*, GUARDIAN (Nov. 8, 2017), https://www.theguardian.com/news/2017/nov/08/key-revelations-from-the-paradise-papers [https://perma.cc/M8PN-43WQ].

303. Jake Smith, *More Opportunities, More Concerns: A Look at the Reactionary Nature of Whistleblower Protections in a Growing Global Economy*, 37 ARIZ. J. INT. COMP. L. 381, 400–03 (2020).

304. *See* Directive (EU) 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the Protection of Persons Who Report Breaches of Union Law, 2019 O.J. (L. 305/17) recital (1) & (4) (2019) [hereinafter Whistleblower Protection Directive].

305. *Id.* at 35, 37 (arts. 4 & 6).

306. *Id.* at 37–40, 41 (arts. 7–13 & 15).

307. *Id.* at 25 (recital 47).

308. The competent authority may include administrative agencies, judicial authorities, or ombudsmen. *Id.* at 39 (art. 10).

309. *Id.* at 41 (art. 15, paras. (a), (b)).

310. *Id.* at 43 (art. 19).

even if reporting illegal activities may surrender trade secrecy.[311] The Whistleblower Directive also penalizes firms that hinder whistleblowing, retaliate against whistleblowers or reveal their identities.[312] These all make employees safe for whistleblowing, and employees can thus be in the best position to detect corporate misconduct and uncover such facts to the public.

b. The Interplay of Whistleblowing, Data Privacy, and Algorithmic Transparency

The Whistleblower Directive can help reduce algorithmic opacity, facilitate stakeholder oversight, and ultimately promote data privacy protections for several reasons.[313] First, the Directive recognizes that whistleblowing can be used to reveal breaches of data protection rules. The Whistleblower Directive explicitly states that "[r]espect for privacy and protection of personal data . . . are other areas in which whistleblowers can help to disclose breaches, which can harm the public interest."[314] As insiders, employees can access information on AI systems that are protected as trade secrets,[315] monitor data practices, identify AI incidents that occur randomly, and report errors in algorithmic disclosures or other misconduct.[316] For instance, in the recent case of the Facebook Papers, Facebook's former product manager Frances Haugen discovered that the algorithms adopted by Facebook could replicate biased and extreme information.[317] According to Haugen, although Facebook was aware of the harmful effects of its AI-based services on users, the firm continued to prioritize commercial profit over individual privacy and autonomy.[318] When Haugen realized that Facebook did not intend to fix its controversial data practices and problematic algorithmic operations, she collected tens of thousands of internal documents as an employee for media disclosures.[319] This case demonstrates how

311. *Id.* at 44 (art. 21, paras. 2–3).

312. *Id.* at 45 (art. 23); Karin Henriksson, *One Year Until the First EU Whistleblower Protection Directive Deadline: 3 Tips for Building a Speak-Up Culture*, WHISTLEB BLOG (Dec. 22, 2020), https://whistleb.com/blog-news/one-year-until-the-first-eu-whistleblower-protection-directive-deadline/ [https://perma.cc/NA2H-D294].

313. For a pathbreaking work discussing the role of whistleblowing in the context of algorithms and how the U.S. Defend Trade Secrets Act's whistleblower immunity regime can be crucial to hold firms accountable for their algorithmic practices, see Katyal, *supra* note 221, at 126–37.

314. Whistleblower Protection Directive, *supra* note 304, at 20 (recital (14)).

315. Katyal, *supra* note 221, at 133, 136.

316. Whistleblower Protection Directive, *supra* note 304, at 24 (recital (43)).

317. For a documentation of what Haugen disclosed to media, see Jeff Horwitz, *The Facebook Files*, WALL ST. J. (Oct. 1, 2021), https://www.wsj.com/articles/the-facebook-files-11631713039 [https://perma.cc/9CTP-RCYW].

318. Karen Haoarchive, *The Facebook Whistleblower Says Its Algorithms Are Dangerous. Here's Why*, MIT TECH. REV. (Oct. 5, 2021), https://www.technologyreview.com/2021/10/05/1036519/facebook-whistleblower-frances-haugen-algorithms/ [https://perma.cc/75U8-GKLT].

319. Reed Albergotti, *Frances Haugen Took Thousands of Facebook Documents: This Is How She Did It*, WASH. POST (Oct. 26, 2021),

employees as whistleblowers can reveal violations and corporate misconduct to the larger public to safeguard data privacy.

Second, whistleblowing can motivate firms to develop and operate AI systems in a lawful and ethical manner. A whistleblower's action can prompt others in similar fields to monitor firms and come forward in the public interest.[320] Whistleblowing often brings about damaging reputational harm, investigations, and litigations. To prevent the triggering of whistleblowing, firms need to be constantly alert to possible violations of legal rules or untruthful statements in algorithmic disclosures.[321] Then, whistleblowing activities are likely to deter firms from violating data privacy rules and motivate them to establish trustworthy data management measures for their use of AI. If firms do not intend to turn away from problematic practices, their misconduct may be revealed by whistleblowers. Hence, whistleblowing protections can facilitate internal oversight that deters firms from wrongdoing, enhances compliance with law, and increases the legality and legitimacy of advanced AI that is increasingly relied upon by the public.[322]

Third, whistleblowing can help increase accountability in AI, as employees are entitled to examine the truthfulness of corporate disclosures. The reporting of whistleblowers can contribute to greater transparency and, by extension, greater accountability in AI originally operating in opaque conditions.[323] With enhanced transparency and accountability, firms would be more careful to avoid unauthorized use of personal data, prevent unjust and unethical treatment in automated decisions, and provide more accurate and balanced algorithmic disclosures for external stakeholder review.

B. Implications for Regulation and Policy Considerations

After discussing the introduction of mandated sustainability disclosures and whistleblowing protections to data privacy spaces, this concluding subsection explores the pros and cons of using the proposed transparency framework to address the issue of algorithmic opacity for regulatory considerations and suggested moves.

---

https://www.washingtonpost.com/technology/2021/10/26/frances-haugen-facebook-whistleblower-documents/ [https://perma.cc/U3B8-8NDB].

320. Katyal, *supra* note 221, at 127–28; Mengqi Sun, *More Tech Whistleblowers Are Expected, Experts Say*, Wall ST. J. (Oct. 28, 2021), https://www.wsj.com/articles/more-tech-whistleblowers-are-expected-experts-say-11635413403 [https://perma.cc/T87Q-AUJL].

321. Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 680 (2016); Katyal, *supra* note 221, at 135.

322. Katyal, *supra* note 221, at 127 ("Whistleblowing has also been shown to be particularly effective in similar situations . . . where whistleblowers have been considered to be vital to achieving greater compliance because they can help to detect areas of wrongdoing.").

323. *Id.*

1. Regulatory Considerations

This part surveys an array of policy considerations that policymakers must take into account, including the tension between trade secrecy and algorithmic transparency, costs and benefits, changes in behaviors, law enforcement, accountability, and innovation.

a. Tension Between Trade Secrets and Algorithmic Disclosures

One of the most critical issues created by algorithmic transparency is the possibility of surrendering trade secrecy in AI. However, the proposed disclosures on AI can balance corporate trade secrecy with individual data privacy through its double-layered disclosure principle, as I have suggested concerning the principle of double-layered disclosure in a previous Article.[324] This piece further argues that under this disclosure principle, complete disclosures to a competent authority are only applicable in rare situations, where the audience is required to maintain the confidentiality of information involving trade secrets. As mentioned previously, information concerning trade secrecy such as source code should be disclosed only to certain audiences under exceptional scenarios to protect the commercial interests of trade secret owners properly.[325] In general, only certain information on crucial aspects of AI and of public concern is required for mandatory disclosures.[326] Requiring firms to disclose core legal, managerial, and technical aspects of AI applications for the protection of privacy and data protection as fundamental values may be considered legitimate for two reasons. First, as explained below, the disclosed items may not involve information that qualifies as trade secrets. Second, even if firms claim that parts of disclosure items involve trade secrets, the Trade Secrets Directive allows that a trade secret is suspended if there is a need to protect "a legitimate interest recognized by Union or national law," which may include privacy and data protection.[327] Although this piece lays out the principle of comprehensible disclosure,[328] the aim of this principle is to provide information on AI applications of public concern rather than asking firms to reveal the entire design, logic, and core secrecy of AI systems.[329] Meanwhile, the principle of minimum-necessary disclosure can be used to balance trade secrecy for commercial interests and algorithmic transparency for data privacy.[330] Since most AI systems consist of personal data, trade secrecy, and

---

324. Lu, *supra* note 27, at 135.

325. *See supra* Part IV.A.1.a.iv. *See generally* Lu, *supra* note 27.

326. Lu, *supra* note 27, at 156–57.

327. Trade Secrets Directive, *supra* note 62, at 11 (art. 5(d)) ("Member States shall ensure that an application for the measures, procedures and remedies provided for in this Directive is dismissed where the alleged acquisition, use or disclosure of the trade secret was carried out in any of the following cases: (d) for the purpose of protecting a legitimate interest recognised by Union or national law.").

328. Lu, *supra* note 27, at 134.

329. Brkan & Bonnet, *supra* note 183, at 40.

330. *See generally* Lu, *supra* note 27, at 135

publicly available information, the proposed algorithmic disclosures do not encourage disclosures on unnecessary information that contains trade secrecy or personal data.[331] In the context of algorithmic opacity, AI-based services or products may hurt data privacy, human rights, and democracy in that they may be intrusive, inaccurate, and discriminative without being held accountable, putting the human rights of individuals at risk.[332] When such public concerns arise, algorithmic disclosures and whistleblowing laws should be used as regulatory tools to balance commercial interest and public interest by asking firms to disclose certain internal practices that can possibly cause harm to citizens.[333] In turn, risky corporate activities will be more likely to be exposed to stakeholder view for the sake of promoting the welfare of citizens.

b. Costs and Benefits

Disclosure duties are usually accompanied by high costs.[334] Yet, one should not neglect that the costs of opacity and non-compliance are high as well, as data privacy laws like the GDPR impose large fines on firms for non-compliance, and firms are facing reputational harm and litigation costs that can lead to loss of consumer trust and sustainable investor funds.[335] Disclosures can be beneficial for firms and stakeholders because they can rebuild mutual trust between firms and citizens through stakeholder communication, and such communication can also become a form of surveillance that can help firms improve misconduct, reduce non-compliance, avoid reputational harm, mitigate AI risks, and even gain trust from the larger public.[336] Although making periodic reports entails large costs for firms, according to recent research on GDPR compliance, more than half of firms agreed that meeting customer expectations was most significant for them, even if they have to bear huge compliance costs.[337] In the post-GDPR age in which citizens are keenly aware of the importance of data privacy, concerned about the unlimited power of omnipresent algorithms, and increasingly urge sustainability disclosure regulations, [338] a request for

331. Desai and Kroll, *supra* note 244, at 38 ("[T]ransparency is often impossible or undesirable in practice.").

332. Whistleblower Protection Directive, *supra* note 304, at 17 (recital (3)).

333. *See* Alexandros Michailidis, *Will the EU Miss Its Chance to Properly Protect Whistleblowers?*, TRANSPARENCY.ORG, https://www.transparency.org/en/news/will-the-eu-miss-its-chance-to-properly-protect-whistleblowers [https://perma.cc/WHJ5-4UPM].

334. Stephen Bainbridge, *Mandatory Disclosure: A Behavioral Analysis Corporate Law*, 68 U. CIN. L. REV. 1023, 1024 (2000).

335. Robert Bird & Stephen Park, *The Domains of Corporate Counsel in an Era of Compliance*, 53 AM. BUS. L. J. 203, 233 (2016). *See generally* Lu, *supra* note 27, at 127.

336. Adam J. Sulkowski, S.P. Parashar & Lu Wei, *Corporate Responsibility Reporting in China, India, Japan, and the West: One Mantra Does Not Fit All*, 42 N. ENGL. L. REV. 787, 798 (2008) (indicating many industrial players believed that CSR reporting can facilitate stakeholder communication and lead to trusting the reporting entity).

337. Robert Bird & Stephen Park, *Turning Corporate Compliance into Competitive Advantage*, 19 U. PA. J. BUS. L. 285, 287 (2017).

338. *The Non-Financial Reporting Directive: What You Need to Know*, *supra* note 227.

algorithmic transparency might be just a question of time. It should thus be worthwhile for firms to bear the cost of disclosures if doing so can meet the expectations of consumers and other stakeholders. In a reality in which citizens and firms see privacy and data protection as precious democratic values, bearing the cost of disclosure could enhance algorithmic accountability that benefits the larger public. Additionally, disclosures help firms build their reputations when they seem more legitimate and responsible in terms of their use of AI systems, promoting corporate sustainability as a result.

c. Organizational and Behavioral Changes

Disclosure cannot directly fix all the problems posed by AI.[339] It is only a tool to enhance transparency for oversight and regulatory purposes. However, algorithmic disclosure can be instrumental in achieving the protection of data privacy and democratic norms. Much research indicated that the existence of disclosure requirements itself can shift corporate behavior toward compliance, as firms are being monitored by others.[340] When it comes to applications of advanced AI—a highly opaque and powerful technology—organizational and managerial transparency is necessary to achieve behavioral reforms and organizational legitimacy through surveillance from stakeholders, whose expectations matter to firms when running their business sustainably.[341] Stakeholder checks and employee monitoring can stimulate acceptable behavior more effectively than legal obligations alone.[342] Because algorithmic disclosures require stakeholder review of corporate business models, policies, outcomes, risks management, and KPIs,[343] firms may be motivated to make internal procedural, substantial, and cultural changes under external pressure.[344] The information firms are required to disclose will become sources for stakeholders to redress harm and for regulators to conduct law enforcement. Meanwhile, such disclosures can open up channels of communication between firms and

339. *See generally* Lu, *supra* note 27, at 154.

340. Pauline Blondet, *How the EU Whistleblower Directive Can Help Enhance a Culture of Integrity*, CORPORATE COMPLIANCE INSIGHTS (June 27, 2019), https://www.corporatecomplianceinsights.com/whistleblower-directive-culture-integrity/ [https://perma.cc/XNQ6-AG4C].

341. David Hess, *The Three Pillars of Corporate Social Reporting as New Governance Regulation: Disclosure, Dialogue, and Development*, 18 BUS. ETHICS Q. 447, 460 (2008). *See generally* Lu, *supra* note 27, at 155–56.

342. Dhruva Krishna, *Deepfakes, Online Platforms, and A Novel Proposal for Transparency, Collaboration, and Education*, 27 RICH. L.J. & TECH. 4, 84 (2021).

343. For a discussion of proposed algorithmic disclosures in the US regulatory context under the topics of "Description of Business," "Legal Proceedings," "Risk Factor," and "Management's Discussion and Analysis of Financial Condition and Results of Operations (MD&A)," *see* Lu, *supra* note 27, at 101, 135–53.

344. *See* Hess, *The Three Pillars of Corporate Social Reporting as New Governance Regulation: Disclosure, Dialogue, and Development*, *supra* note 341.

stakeholders.[345] As a result, legal obligations on transparency can become an effective social tool that helps firms collaborate with stakeholders to begin a wholesome monitoring process for advanced, healthy, and lawful corporate practices that respect human rights and lead to sustainable business operation.[346]

In the context of whistleblowing, employees can also play a part in promoting the protection of data privacy, human rights, and democratic values. In firms where commercial seclusion is a default setting, employees can report opaque corporate misconduct to help build a more ethical working culture.[347] Additionally, with employee whistleblowing mechanisms in place, firms will be more alert to potential rule-breaking behaviors. This means that employees can also contribute to dialogue that accelerates corporate practice reforms. Also, trade unions and non-profit organizations can be instrumental in affecting the enforcement of whistleblower protections by urging follow-up on reporting.[348] Thus, disclosures and whistleblowing could in turn promote organizational changes toward public interest and corporate sustainability.[349]

d. Law Enforcement

In addition to organizational reforms, transparency requirements can serve as an excellent strategy to improve enforcement of data privacy laws. Through mandatory disclosure duties or reports presented by whistleblowers, corporate misbehavior can be identified by regulators and the broader public for enforcement purposes.[350] In the area of data privacy, algorithmic opacity has led to weaknesses in enforcement because trade secrecy, technical opacity, and managerial invisibility in AI make problematic applications of AI systems hard to detect, understand, or correct.[351] However, it is unnecessary to overprotect the secrecy of every aspect of business applications of AI if doing so hinders effective law enforcement and inflicts harm on the broader public.[352] In this

345. This is also emphasized by the UN's Special Representative of the Secretary-General, Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy." OFFICE OF THE HIGH COMMISSIONER, *supra* note 266.

346. *See* Hess, *The Transparency Trap: Non-Financial Disclosure and the Responsibility of Business to Respect Human Rights*, *supra* note 129, at 17; Stephen Kim Park, *Targeted Social Transparency as Global Corporate Strategy*, 35 NW. J. INT. L. & BUS. 52, 114 (2014); *Hess, The Three Pillars of Corporate Social Reporting as New Governance Regulation: Disclosure, Dialogue, and Developmen*t, *supra* note 341, at 460.

347. Vigjilenca Abazi, *The European Union Whistleblower Directive: A "Game Changer" for Whistleblowing Protection?*, 49 INDUS. L.J. 640, 654 (2020).

348. *See* David Lewis & Wim Vandekerckhove, *Trade Unions and the Whistleblowing Process in the UK: An Opportunity for Strategic Expansion?*, 148 J. BUS. ETHICS 835, 838-40 (2018).

349. *See* Abazi, *supra* note 347, at 653, 655.

350. Whistleblower Protection Directive, *supra* note 304, at 17 (recital 2).

351. See *id.* at 1–2 (recital 3); *see generally* Lu, *supra* note 27, at 115, 122, 130.

352. *See* Peter S. Menell, *Tailoring a Public Policy Exception to Trade Secret Protection*, 105 CALIF. L. REV. 1, 18 (2017) ("[U]ncritical protection of all secret business information can conflict with effective law enforcement."); The EU's Trade Secrets Directive also excludes trivial information from trade secret protection. *See* Trade Secrets Directive, *supra* note 62, at 4.

context, disclosure rules and whistleblowing mechanisms can not only enable regulators, shareholders, and other stakeholders to monitor AI applications, but also incentivize firms to self-monitor algorithmic business practices.[353] For instance, employees can monitor whether their firms are using AI in a manner that compromises data privacy or other democratic norms, and then turn to external reporting channels if their firms fail to address the reported issues within a reasonable time frame.[354] In doing so, disclosure duties and whistleblowing activities can promote the investigation of corporate wrongdoing and lower the cost of enforcement through a collaboration between private entities and governments.[355] Consequently, enforcement systems will be fed with a larger amount of useful information, leading to effective prosecution and understanding of the state of AI development.[356]

e. Accountability

Transparency requirements can be useful in promoting private accountability in an atmosphere of private domination of AI systems development.[357] When society increasingly depends on automated decisions made by AI, firms have gradually come to have some control over citizens' daily activities.[358] In this regard, the inner workings and algorithmic results of AI systems developed by firms need to be held accountable, whether legally, managerially, or technically.[359] Through whistleblowing protections, employees will be encouraged to review corporate use of AI systems and empowered to uncover corporate wrongdoing.[360] With algorithmic disclosure duties, firms will also be required to understand and explain the operation of AI applications that increasingly become decision-makers in various aspects of civic life.[361] In turn, transparency enables stakeholders such as consumers to be aware of problematic

353. *See* Katyal, *supra* note 221, at 129; Orly Lobel, *The Lawyer's Role in a Contemporary Democracy, Promoting the Rule of Law, Lawyering Loyalties: Speech Rights and Duties Within Twenty-First-Century New Governance*, 77 FORDHAM L. REV. 1245, 1249 (2009).

354. Scott Pelley, *Whistleblower: Facebook Is Misleading the Public on Progress Against Hate Speech, Violence, Misinformation*, CBS NEWS (Oct. 4, 2021), https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/ [https://perma.cc/4JA6-PWPF].

355. Norman D. Bishara, Elletta Sangrey Callaha & Terry Morehead Dworkin, *The Mouth of Truth*, 10 N.Y.U. J.L. & BUS., 37, 39–40 (2013); Menell, *supra* note 352, at 22.

356. *See* Whistleblower Protection Directive, *supra* note 304, at 17 (recital (2)); Katyal, *supra* note 221, at 129.

357. *See generally* David S. Levine, *The Impact of Trade Secrecy on Public Transparency, in* THE LAW AND THEORY OF TRADE SECRECY: A HANDBOOK OF CONTEMPORARY RESEARCH 406 (2011); Kaminski, *supra* note 221; Lu, *supra* note 27, at 158.

358. Katyal, *supra* note 221, at 128. *See generally* Lu, *supra* note 27, at 99, 119–20.

359. *See generally* Lu, *supra* note 27, at 99.

360. Katyal, *supra* note 221, at 127. *See* Orly Lobel, *Citizenship, Organizational Citizenship, and the Laws of Overlapping Obligations*, 97 CALIF. L. REV. 433, 473 (2009).

361. Katyal, *supra* note 221, at 128 (quoting Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1252 (2008)).

algorithmic governance and empowers them to invoke their legal rights. These all may help increase private accountability for corporate use of AI systems.[362]

f. Innovation

Greater transparency can not only require firms to share information on AI systems, but also indirectly encourage more individuals to share their information voluntarily with corporations, benefiting innovations in AI systems as a result. Evidence from a report released by Datatilsynet, the Norwegian data protection authority, revealed that European citizens were concerned about the commercial use of personal information associated with AI systems.[363] At the same time, the report indicated that AI developers were using AI systems restrictively due to a growing public distrust of corporate data practices.[364] Establishing a robust disclosure framework to foster a data sharing culture is thus of great value for the facilitation of data-driven innovation.[365] Through an algorithmic disclosure and whistleblowing regime, individuals would be more likely to believe that firms are being monitored and their personal data is being processed lawfully. Disclosures can also stimulate innovation by enabling innovators to utilize more technological and commercial information about AI systems.[366] With trust in corporate use of personal information and AI, individuals would be more willing to share information about themselves with the private sector, which would open opportunities to develop more innovative and trustworthy AI-based services and products. As emphasized in the EU's Open Source Software Strategy, greater transparency can bring about greater exposure to brilliant ideas, which will not only help produce solutions to problematic data practices but also improve the design of AI systems, encouraging innovation more broadly.[367] At the same time, the adoption of algorithmic disclosures can fix the regulatory vacuum, improve existing disclosure standards, and catalyze useful, innovative algorithmic transparency initiatives among firms and stakeholders.[368]

---

362. *See generally* Margot E. Kaminski, *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, 92 S. CAL. L. REV. 1529 (2019) (the GDPR's accountability problem).

363. DATATILSYNET, ARTIFICIAL INTELLIGENCE AND PRIVACY (2018), https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf [https://perma.cc/26W6-US4Z].

364. *Id.* at 6.

365. Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 126 (2014).

366. *See* Jeanne C. Fromer, *Dynamic Patent Disclosure*, 69 VAND. L. REV. 1715, 1715–17 (2016).

367. *See Open Source Software Strategy*, EUR. COMM'N, https://ec.europa.eu/info/departments/informatics/open-source-software-strategy_en [https://perma.cc/C3KW-3FJV].

368. *See* Ho & Park, *supra* note 228, at 302–06; Non-Financial Reporting Directive, *supra* note 228, at 14.

2. Suggested Moves

As AI systems pervasively and adversely affect social and commercial aspects of democratic societies, existing data privacy law perpetuates a status quo of algorithmic opacity that obscures unlawful AI applications and inadequate data governance from stakeholder view.[369] Since current legal responses are unlikely to offer satisfactory solutions, new legal rules are necessary to protect data privacy, human rights, and democratic norms. Although transparency can only partially address issues of AI governance,[370] it should be an excellent complement to regulations safeguarding the protection of data privacy and other democratic values. In an age in which the private sector dominates the deployment of AI systems to strengthen corporate power, algorithmic opacity has become a strategy for firms to avoid responsibility.[371] Opacity exacerbates the scale of risk posed by AI systems, including erosion of data privacy, equality, autonomy, and more.[372] However, as even the GDPR permits too much opacity in business applications of AI, policymakers should consider imposing complementary transparency duties for algorithmic systems through disclosure requirements.[373] The proposed algorithmic disclosures can require firms to provide critical information on their use of AI, put high-risk AI applications under stakeholder surveillance, strengthen law enforcement that is negatively affected by opacity, and facilitate public debate on innovative algorithmic governance, all of which are fundamental to achieving the ambition of the EU and many other countries: to develop human-centered and trustworthy AI.[374]

The proposed algorithmic disclosure framework that outlines proactive, continuous, comprehensive, and flexible transparency duties should be adopted as a legally binding regulation. First, considering algorithmic opacity's massive and pervasive impacts on corporate behaviors, policymakers should move away from flexible international disclosure standards and impose a set of corporate social disclosure duties to establish a wholesome data-centric regulatory system.[375] Otherwise, algorithmic opacity will continue posing a barrier to firms' capacity to perceive risks and receive oversight by hiding inaccurate, biased, and manipulated algorithmic practices. Second, the previously proposed mandated

369. Genevieve LeBaron & Andreas Rühmkorf, *The Domestic Politics of Corporate Accountability Legislation: Struggles Over the 2015 UK Modern Slavery Act*, 17 SOCIO-ECON. REV. 709, 713–15, 736–38 (2019).

370. Abazi, *supra* note 347, at 654.

371. Ari Ezra Waldman, *Power, Process, and Automated Decision-Making*, 88 FORDHAM L. REV. 613, 618–21 (2019).

372. *See generally* Nick Feamster, *Mitigating the Increasing Risks of an Insecure Internet of Things*, 16 COLO. TECH. L.J. 87 (2017); Scherer, *supra* note 33; Lu, *supra* note 27, at 100, 117–27.

373. *See generally* Lu, *supra* note 27.

374. Surya Mattu & Kashmir Hill, *How a Company You've Never Heard of Sends You Letters About Your Medical Condition*, GIZMODO (June 19, 2017), https://gizmodo.com/how-a-company-you-ve-never-heard-of-sends-you-letters-a-1795643539 [https://perma.cc/N688-3MPT].

375. *See* Park, *supra* note 346, at 102.

algorithmic disclosures should be audited, published annually, and made accessible online for periodic stakeholder review. Otherwise, based on existing *ex ante* privacy notices, stakeholders and consumers alike cannot adequately and regularly assess risks derived from AI-based products and services, weakening the effectiveness of data privacy rules like the GDPR; shareholders as external stakeholders will likely lack enough information to make informed investment decisions on a regular basis; firms will thus be less incentivized to improve their data management measures, accelerate substantial organizational reforms, or improve multilateral communication with the larger public.[376] Third, policymakers should consider enhancing the consistency and effectiveness of whistleblower protection as a part of data privacy and human rights protection. If not, firms are likely to disclose misleading information on their algorithmic practices, perpetuating algorithmic opacity to escape actual stakeholder surveillance. Fourth, as AI technologies are likely to evolve continuously and create new legal issues, regulators should enable multi-stakeholder engagement in algorithmic governance to update mandated disclosure standards continually. Additionally, regulators should be more proactive rather than reactive in enhancing algorithmic transparency as a critical step in regulating AI systems. Otherwise, the disclosure duties may no longer meet the social and financial needs of stakeholders. With such moves, the combination of algorithmic disclosures and whistleblowing mechanisms can provide useful information on AI to protect data privacy, human rights, and other democratic values. Based on these proposed transparency strategies, corporate AI systems that are originally invisible to outsiders would be more likely to be assessed, studied, and corrected by regulators and stakeholders at an early stage, in turn building trustworthy AI systems that benefit both firms and citizens.

CONCLUSION

In the era of AI, in which policymakers are confronted by adverse consequences of algorithmic opacity that jeopardize the public's ability to hold the private sector accountable, transparency is essential for effective stakeholder oversight and enforcement of democratic norms. Hence, it should be treated as a carrier of data privacy rules and should form an important part of policy considerations for the protection of data privacy, human rights, and democratic values. The right to data privacy cannot be assured if affected individuals cannot understand how private AI applications access, assess, and intrude upon them. The future of data privacy law rests on a new set of proactive regulatory approaches. A disclosure scheme and whistleblowing system that unveil how firms develop and deploy AI offer an excellent vehicle for such considerations, whether legally, socially, or commercially. Regulators in the EU, the United

376. Iris H.-Y. Chiu & Ernest W.K. Lim, *Managing Corporations' Risk in Adopting Artificial Intelligence: A Corporate Responsibility Paradigm*, 20 WASH. U. GLOBAL STUD. L. REV. 347, 388 (2021).

States, and elsewhere should consider imposing a duty of social transparency on firms deploying AI techniques to protect human rights and democratic values. The proposed new set of transparency rules will shed light on effective stakeholder oversight, law enforcement, and data governance, leading to trustworthy AI systems that benefit both firms and citizens. As citizens can place their trust in AI, rules of transparency can perhaps one day clear a path to a world of human-centered, privacy-compliant, and omnipresent AI systems in the post-GDPR era.