# Beyond Automation:
# Machine Learning-Based Systems and Human Behavior in the Personalization Economy

Niklas Eder[*]

ABSTRACT

*Personalization has long been a feature of online services, shaping targeted advertisement and online manipulation. Corporations now seek to exploit the enormous economic potential of personalization beyond the confinements of the online space. In recent years, the proliferation of machine learning-based decision-making has led to personalization in all spheres of life. Corporations rely on machine learning-based systems to decide if and under what conditions they contract with individuals. They determine who is invited for job interviews and who is eligible for loans. They shape how we are perceived and tailor the way in which we are treated. The implications of these systems are already immense, and they foreshadow a larger transformation. Over the course of the 21st century, ubiquitous, machine learning-based personalization will likely permeate the economy and become a fundamental condition of human existence. The shape of this transformation is still uncertain; before it concretizes, we have the opportunity to guide its direction by articulating concepts that allow us to describe and critically examine it.*

*Legal scholarship needs a conceptual foundation to address urgent questions about how personalization, driven by machine learning-based decision-making, affects liberty and other liberal democratic values. This article draws from surveillance theory and develops that foundation. It constructs a novel approach to examine the normative and constitutional implications of machine learning-based decision systems and ubiquitous personalization. The article builds on the concepts of panopticism and the surveillance assemblage to analyze how corporate machine learning-based decision-making affects the lives of individuals and transforms society. It is the first to develop an account of how ubiquitous personalization influences human agency and behavior. The article describes how machine learning-based decision systems amplify corporate power. It provides theoretical support for what Jack Balkin calls "normalization (or regimentation)"–the idea that algorithmic evaluations and decisions will govern human behavior. The article shows that existing legal responses focusing on rights, explainability, and transparency fail to prevent the already fragile balance of power between individuals and corporations from tipping in corporations' interest. It argues that legal scholarship, to adequately respond to machine learning-based decision-making, must overcome its individualistic focus and engage in a debate on the legitimacy of corporate surveillance.*

*Finally, the article explains why and how we should measure legal responses to machine learning-based decision-making against standards of legitimacy. The notion of legitimacy provides a foundation to tackle one of the great challenges with which machine learning-based systems confront liberal democracies—which is to reconcile corporate power with the values and freedoms central to these democracies. A legitimacy focus suggests that for legal responses to be adequate, they must not only educate individuals on the functions of algorithmic systems and endow them with legal rights. Rather, they must also entail general principles, such as data minimization and limitation, shaping the conditions based on which the personalization economy operates.*

TABLE OF CONTENTS

## I. INTRODUCTION

The idea that algorithms transform society, creating what has poignantly been referred to as the "scored,"[1] "algorithmic,"[2] or "black-box"[3] society, is widespread.[4] Automated decision-making is one of the most important algorithmic applications fueling this transformation.[5] The

---

[1] Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1 (2014).

[2] Jack M. Balkin, *The Three Laws of Robotics in the Age of Big Data,* 78 OHIO ST. L.J. 1217, 1219-21 (2017) [hereinafter Balkin, *Three Laws of Robotics*]; Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U.C. DAVIS L. REV. 1149, 1154-56 (2018) [hereinafter Balkin, *Free Speech in the Algorithmic Society*].

[3] FRANK PASQUALE, THE BLACK BOX SOCIETY (2015).

[4] For a brief but insightful account on the complex interplay of innovations, humans, and the social order, see Jack M. Balkin, *The Path of Robotics Law*, 6 CALIF. L. REV. 45 (2015).

[5] Automated decision-making provokes great scholarly interest. *See, e.g.*, Kiel Brennan-Marquez et al., *Strange Loops: Apparent versus Actual Human Involvement in Automated Decision-Making*, 34 BERKELEY TECH. L.J. 745 (2019) (engaging with the question whether it matters if humans appear to be in the loop of automated decision-making); Deven R. Desai & Joshua A. Kroll, *Trust but Verify: A Guide to Algorithms and the Law*, 31 HARV. J.L. & TECH. 1 (2017) (providing an insightful introduction and arguing that law must regulate software and algorithms); Tal Zarsky, *The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision-Making*, 41 SCI., TECH. & HUM. VALUES 118 (2015) (aspiring to develop a roadmap for regulators); Citron & Pasquale, *supra* note 1 (carving out the great impact of scores on human lives and suggesting ways to guarantee due process). For an earlier account, see Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249 (2008); PASQUALE, *supra* note 3 (serving as a standard reference engaging with issues of algorithmic opacity and power); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93 (2014) (astutely describing the impact of profiles on individual's lives and livelihood and laying out a mode of regulation); Frank Pasquale, *Restoring Transparency to Automated Authority Internet Governance*, 9 J. ON TELECOMMS. & HIGH TECH. L. 235 (2011) (describing algorithmic systems as "automated authority" and emphasizing the importance of transparency in response to that authority); Bart W. Schermer, *The Limits of Privacy in Automated Profiling and Data Mining*, 27 COMP. L. & SEC. REV. 45 (2011) (an insightful account of how automated systems circumvent conventional privacy

common label of "automated decision-making," however, is misleading. To understand the implications of algorithmic decision-systems, we have to look beyond automation. What shapes these systems, and ultimately the effects they have on the individual and society, are machine learning algorithms.[6] They are therefore best referred to as machine learning-based (MLB) decision-systems.[7]

Corporations rely on MLB decision-systems to minimize risks by personalizing the conditions based on which they interact with individuals. They allow corporations to individualize the offline space and constitute the foundation of what is best referred to as the "personalization economy." Relying on MLB systems to increase profits is not a new phenomenon. Corporations have relied on MLB systems for years to decide what advertisements to show to users and how to personalize their online

---

protections). For perspectives with a focus on the European Data Protection Regulation's response to automated-decision-making, see Margot E. Kaminski, *The Right to Explanation, Explained*, 34 BERKELEY TECH. L.J. 189 (2019); Sandra Wachter et al., *Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR*, 31 HARV. J.L. & TECH. 841 (2018); Bryan Casey et al., *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); Sandra Wachter et al., *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 INT'L. DATA PRIV. L. 76 (2017) [hereinafter Wachter et al., *Right to Explanation*]; Isaak Mendoza & Lee A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling, in* EU INTERNET LAW 77 (Tatiana-Eleni Synodinou et al. eds., 2017).

[6] For a concise and accessible account of how machine learning algorithms create knowledge, see Mireille Hildebrandt, *Profiling: From Data to Knowledge*, 30 DATENSCHUTZ UND DATENSICHERHEIT 548 (2006); danah boyd & Kate Crawford, *Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon*, 15 INFO., COMMC'N & SOC'Y 662, 665-66 (2012) (focusing on the particularities of knowledge created by machine learning algorithms); for more detailed analysis of the functionality of machine learning and the notion of Big Data see VIKTOR MAYER-SCHÖNBERGER, BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK AND THINK (1st ed. 2013); Emily Berman, *A Government of Laws and Not of Machines*, 98 B.U. L. REV. 1277 (2018).

[7] The notion of MLB decision-making invites us to draw on the rich literature that engages with the way machine learning algorithms generate knowledge and the implications of that knowledge. *See* boyd & Crawford, *supra* note 6 (providing an especially insightful account on how algorithms change the "definition of knowledge"); Mireille Hildebrandt & Bert-Jaap Koops, *The Challenges of Ambient Law and Legal Protection in the Profiling Era*, 73 MOD. L. REV. 428 (2010) (serving as an early account with illuminating introductions to the most important notions, including "Artificial Intelligence," and introducing the idea of "ambient intelligence"); Lanah Kammourieh et. al., *Group Privacy in the Age of Big Data, in* GROUP PRIVACY IN THE AGE OF BIG DATA (Linnet Taylor et al. eds., 2017) (relying on the notion of grouping to both understand the functionality of as well as respond to new ways of generating knowledge); Ira S. Rubinstein, *Big Data: The End of Privacy or a New Beginning?*, 3 INT'L DATA PRIV. L. 74 (2013) (introducing briefly but concisely some of the most important challenges); Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact,* 104 CALIF. L. REV. 671 (2016) (demonstrating powerfully how algorithms reproduce inequalities); Solon Barocas & Helen Nissenbaum, *Big Data's End Run Around Anonymity and Consent*, *in* PRIVACY, BIG DATA, AND THE PUBLIC GOOD: FRAMEWORKS FOR ENGAGEMENT (Helen Nissenbaum et al. eds., 2014) (demonstrating how Big Data challenges conventional approaches to protect privacy); Balkin, *Three Laws of Robotics*, *supra* note 2; Neil M. Richards & Jonathan H. King, *Big Data Ethics*, 49 WAKE FOREST L. REV. 393 (2014).

experience—practices well studied under the notion of "surveillance capitalism."[8] As corporations start exploiting the economic potential of MLB systems in all spheres of life, the environment in which we live begins to transform. We are witnessing the beginning of a profound transformation: the ubiquitous personalization of life through MLB decision-making.

Our understanding of what this transformation entails has remained limited. The limitations of our understanding, in turn, constrain our ability to normatively assess this transformation and to suggest adequate legal responses. While these limitations are partially inherent to the subject matter with which we engage—a transformation that is just beginning to take shape and that cannot easily be empirically assessed yet—they are also caused by a lack of interdisciplinary perspectives, concepts, and vocabulary to describe them.

This article builds on anthropological and sociological work on surveillance to develop analytical foundations that allow theorizing still uncertain effects of MLB decision-making and thus provides a basis on which legal scholarship can build. Building on these foundations, it assesses whether ubiquitous personalization is likely to influence our intuition, and with it one of the key determinants of human behavior.[9] It describes how MLB decision-making amplifies corporate power and argues that existing legal responses focusing on rights, explainability, and transparency do not account for all harms of MLB decision-making. To adequately respond to the effects of MLB decision-making, the article argues, legal scholarship must overcome its individualistic focus and engage in a debate on the legitimacy of corporate surveillance. The article explains why and how we should measure legal responses to MLB decision-making against standards of legitimacy. It argues that the notion of legitimacy provides a foundation through which liberal democracies can tackle one of the great challenges presented by steadily improving data processing capabilities: reconciling corporate power with the values and freedoms central to these democracies.

---

[8] *See* SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER (2019); *see also* S. C. Matz et al., *Psychological Targeting as an Effective Approach to Digital Mass Persuasion*, 114 PNAS 12714 (2017); Karen Yeung, *'Hypernudge': Big Data as a Mode of Regulation by Design*, 20 INFO., COMMC'N & SOC'Y 118 (2017); Cass R. Sunstein, *Fifty Shades of Manipulation*, 1 J. MKT. BEHAV. 213 (2016); Tal Z. Zarsky, *Privacy and Manipulation in the Digital Age*, 20 THEORET. INQUIRIES L. 157 (2019); Micah L. Berman, *Manipulative Marketing and the First Amendment*, 103 GEO. L.J. 497 (2015); Chris Jay Hoofnagle et al., *Behavioral Advertising: The Offer You Cannot Refuse*, 6 HARV. L. & POL'Y REV. 273 (2012); Kirstie Ball, *All Consuming Surveillance*: *Surveillance as Marketplace Icon*, 20 CONSUMPT. MKTS. & CULTURE 95 (2017).

[9] See Kahneman's influential work on intuition and human behavior: DANIEL KAHNEMAN, THINKING, FAST AND SLOW (1st ed. 2011).

This article engages with the development that all spheres of life increasingly resemble the online space, a space in which behavior is permanently tracked, analyzed, and ultimately converted into ubiquitous personalized treatments.[10] MLB decision-making consists in practices in which algorithms are used to evaluate individuals, to create predictions about their future, and ultimately to treat them in particular ways.[11] Whether we are regarded as psychologically fragile or mentally stable, as financially trustworthy or unworthy of credit, as a productive worker or a liability to the firm, increasingly depends on the evaluations of machine learning algorithms.[12] MLB decision-systems "sort"[13] individuals according to their digital reputations[14] and set the conditions on the basis of which corporations interact with them. We are at the beginning of an era in which the opportunities we have in life and the environments we live in are shaped by MLB decision-making. One of the hardest and most important

---

[10] Already today, algorithmically generated knowledge is used to personalize the conditions for receiving loans. *See* Mikella Hurley & Julius Adebayo, *Credit Scoring in the Age of Big Data*, 18 YALE J.L. & TECH. 148 (2016). For deciding whom to consider for a vacancy, see David Angrave et al., *HR and Analytics: Why HR Is Set to Fail the Big Data Challenge*, 26 HUM. RES. MGMT. J. 1 (2016). For deciding whom to consider as tenants, see Katy McLaughlin, *Robots Are Taking Over (the Rental Screening Process)*, WALL ST. J. (Nov. 21, 2019), https://perma.cc/WK4A-W5RP. For deciding whom to support in job seeking processes, see Doris Allhutter et al., *Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective*, FRONTIERS BIG DATA (Feb. 21 2020), https://perma.cc/9XCG-EEJX.

[11] Algorithms are, for example, used to infer the socioeconomic statuses of individuals on the basis of their Facebook data: Astra Taylor & Jathan Sadowski, *How Companies Turn Your Facebook Activity Into a Credit Score*, THE NATION (May 27, 2015), https://perma.cc/RRK3-K686; to deduce health information on the basis of speech patterns recorded by digital assistants: James Cook, *Amazon Patents New Alexa Feature Knows When You're Ill and Offers You Medicine*, TELEGRAPH (Oct. 9, 2018), https://perma.cc/AHF3-3R4T; to detect pregnancies on the basis of shopping behavior: Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES (Feb. 16, 2012), https://perma.cc/N9E8-P8T9; to predict depressions and suicides on the basis of pictures uploaded to Instagram: Andrew G. Reece & Christopher M. Danforth, *Instagram Photos Reveal Predictive Markers of Depression*, 6 EPJ DATA SCI. 21 (2017); and to reveal habits on the basis of credit card histories: Josh Lauer, *Plastic Surveillance: Payment Cards and the History of Transactional Data, 1888 to Present*, 7 BIG DATA & SOC'Y 1 (2020).

[12] For scoring practices in the financial services sector, see Taylor & Sadowski, *supra* note 11; Chris Cuomo et al., *'GMA' Gets Answers: Some Credit Card Companies Financially Profiling Customers*, ABC NEWS (Jan. 22, 2009), https://perma.cc/MU5S-4XYM. Regarding the health sector, see Ryen W. White, et al., *Detecting Neurodegenerative Disorders from Web Search Signals*, DIGITAL MEDICINE (Apr. 23, 2018), https://perma.cc/C5BS-Z3DA; Leslie Scism, *New York Insurers Can Evaluate Your Social Media Use—If They Can Prove Why It's Needed,* WALL. ST. J. (Jan. 30, 2019), https://perma.cc/68C2-T8VD. Regarding the housing sector, see McLaughlin, *supra* note 10. Regarding employment, see Angrave et al., *supra* note 10.

[13] David Lyon, *Surveillance as Social Sorting: Computer Codes and Mobile Bodies*, *in* SURVEILLANCE AS SOCIAL SORTING 13 (David Lyon ed., 2003).

[14] For an early and comprehensive account of automated scoring systems, see Citron & Pasquale, *supra* note 1. On the notion of digital reputations, see PASQUALE, *supra* note 3, at 19-58.

questions relating to this development is if and how it will affect human consciousness and behavior, our freedom and self-determination.

Scholars have begun speculating on this question, expecting profound effects, but providing little theoretical background to substantiate their expectations. Viktor Mayer-Schönberger has argued that "subsequent generations may have a 'big data consciousness'—the presumption that there is a quantitative component to all that we do."[15] Jack Balkin, relying on the notion of "normalization (or regimentation)," reflects on how individuals might self-police their behavior in response to digital and algorithmic surveillance.[16] Balkin describes a dynamic in which an "algorithm causes you to internalize its classifications and assessments of risk, causing you to alter your behavior in order to avoid surveillance or avoid being categorized as risky."[17] He argues that algorithmic systems may cause people to alter their "identity, behavior, or other aspects of personal self-presentation in order to appear less risky to the algorithm, or to fall into a different category."[18] Frank Pasquale describes the penalizing nature of a system in which a credit card company relies on the algorithmic analysis of tracked behavior to individualize interest rates.[19] He argues that once the workings of such systems are disclosed, people might refrain from behavior that algorithms view as a "signal" for financial distress, even if that behavior might in fact be crucial to their well-being.[20]

The idea that MLB decision-systems will shape the way we think and behave is both alarming and fascinating. It is high time that legal scholarship engages in depth with this idea and the question of if and how MLB decision-making will affect the behavior of individuals. Just as critical, legal scholars must engage in a debate over the normative implications of this development, on both an individual and societal level. This article is the first to provide foundations for this urgently needed debate.

---

[15] MAYER-SCHÖNBERGER, *supra* note 6, at 97.

[16] Balkin, *Three Laws of Robotics*, *supra* note 2, at 1238.

[17] *Id.*

[18] *Id.* at 1238-39. Scholars have also focused on discovering behavioral changes online. Engaging with the effects of online state surveillance, scholars have recently empirically examined the emergence of chilling effects on the internet. Jonathon W. Penney, *Chilling Effects: Online Surveillance and Wikipedia Use*, 31 BERKELEY TECH. L.J. 117 (2016); Elizabeth Stoycheff et al., *Privacy and the Panopticon: Online Mass Surveillance's Deterrence and Chilling Effects*, 21 NEW MEDIA SOC'Y 602 (2019). Analyzing how individuals self-present in online environments, scholars have identified how the way one is perceived in these environments affects people's identity and selfhood and how individuals strive to shape the perception of themselves. Alison Hearn, *Verified: Self-Presentation, Identity Management, and Selfhood in the Age of Big Data*, 15 POPULAR COMMC'N 62 (2017).

[19] PASQUALE, *supra* note 3, at 23.

[20] *Id.* at 31.

To provide these foundations, the article draws on surveillance studies. Building on the sociological and anthropological research field of surveillance studies allows us to develop an in-depth understanding of the effects MLB decision-making has on individuals and on society. [21] Surveillance scholars have developed concepts that allow us to theorize on how technology gives rise to new forms of surveillance, and how these practices affect the behavior of individuals and impact society. Surveillance theory, as Julie Cohen has astutely explained, provides a toolkit that legal scholars can rely on when describing the harms which their legal responses are supposed to address.[22] This article introduces legal scholars to that toolkit and shows ways to apply it to MLB decision-making. It demonstrates how legal scholarship, engaging with the social and societal implications of algorithmic systems, can rely on surveillance theory to further refine its analyses.

Building on surveillance studies to enhance our understanding of the normative and legal issues arising from MLB decision-making is a promising endeavor. However, it also entails significant challenges. There is no single theory that legal scholarship can borrow and readily apply to MLB decision-making. The field of surveillance studies is characterized by controversies on what the best analytical models are to observe the effects of surveillance. These controversies cannot be set aside by legal scholars who wish to build on surveillance theory. Different concepts of surveillance lead to different understandings of the effects of surveillance in general, and of MLB decision-making in particular. What's more, different ways to conceptualize the effects of MLB decision-making ultimately suggest different legal responses. Legal scholarship, which builds on surveillance studies, must therefore engage with these controversies, too. By engaging with the two most impactful and conflicting models used to analyze surveillance, this article discusses important controversies within surveillance studies and explains their consequences on legal and constitutional assessments.

When engaging with pervasive personalization and the MLB decision-making driving it, the question arises of why to frame it as a matter of surveillance, and why to focus on corporate, rather than state, practices. First, with regard to practices related to MLB decision-making, like behavioral profiling and targeted advertisement, the notion of surveillance is already widely used. "Surveillance Capitalism" has become a common

---

[21] David Murakami Wood, *Situating Surveillance Studies*, 6 SURVEILLANCE & SOC'Y 52 (2009) (delimiting Surveillance Studies as a research field). For a brief history of surveillance theory, see Maša Galič et al., *Bentham, Deleuze and Beyond: An Overview of Surveillance Theories from the Panopticon to Participation*, 30 PHIL. & TECH. 9 (2017).

[22] *See* Julie E. Cohen, *Studying Law Studying Surveillance*, 13 SURVEILLANCE & SOC'Y 91 (2015) (calling for legal scholarship and surveillance studies to inform each other).

framework to analyze how corporations "commodify" the private sphere and manipulate individuals to generate profit.[23] Describing behavioral profiling and targeted advertisement as forms of corporate surveillance has allowed scholars to analyze how these practices are embedded in capitalism and impact society. Framing MLB decision-making as a form of corporate surveillance brings with it the potential for similar analyses. It can provide the foundation for legal scholarship that critically analyzes MLB decision-making while accounting for the complexities which come with analyzing its social and societal effects in the context of capitalism.

Second, focusing on corporate MLB decision-making, rather than state practices, is especially urgent, as corporate practices, despite profoundly impacting the lives of individuals, are not as constrained by constitutional protections.[24] The balance of power between individual and corporate freedoms was already uneven before MLB decision-making. Now, it threatens to collapse entirely to corporate benefit as machine learning algorithms put unprecedented data processing capacities in the hands of corporations. Legal scholarship must develop responses that prevent this balance from collapsing. It must suggest ways to reconcile values central to liberal democracies with the goal to exploit the undeniable economic potential of machine learning algorithms.

The article develops two perspectives on the effects of corporate MLB-decision-making and revisits responses to MLB decision-making suggested by legal scholarship on the basis of these perspectives.[25] By providing interdisciplinary engagement with the social and societal effects of MLB decision-making, it adds a crucial dimension to the legal debate on algorithmic power. The article develops concepts that allow legal scholars to critically assess the effects of MLB decision-making as they begin to take a clearer shape. It explains how MLB decision-making transforms the environment of individuals and might impact their behavior. It describes if and how MLB decision-making might create dynamics such as the ones identified by Mayer-Schönberger, Balkin, and Pasquale. The article develops a novel perspective on the canon of responses suggested by legal

---

[23] *See supra* note 8.

[24] An especially thorough analysis of MLB systems used by the state has been developed by Berman, *supra* note 6.

[25] For overviews on responses to automated decision-making, see JESSICA FJELD ET AL., PRINCIPLED ARTIFICIAL INTELLIGENCE: MAPPING CONSENSUS IN ETHICAL AND RIGHTS-BASED APPROACHES TO PRINCIPLES FOR AI (2020) (available at https://perma.cc/HQ9K-WN44); Brent Mittelstadt et al., *The Ethics of Algorithms: Mapping the Debate*, 3 BIG DATA AND SOC'Y 1 (2016). For a variety of detailed suggestions, see Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 1 (2017).

scholars and considered by regulators, including transparency, [26] explainability, [27] non-discrimination, [28] fairness, [29] reasonableness, [30] due process,[31] data protection,[32] privacy,[33] and the right not to be subjected to automated decision-making.[34] The article makes implicit assumptions on which suggested responses rest and critically assesses these assumptions. It engages with the limits of individual rights and transparency as main strategies to respond to MLB decision-making. The article shows how MLB decision-making amplifies corporate power and that it can be conceived as a system of governance—one which functions differently from conventional systems of governance but which has similar effects for those who are governed. [35] Ultimately, it argues why and how the notion of

---

[26] *See, e.g.*, Desai & Kroll, *supra* note 5 (describing transparency as the standard solution to algorthmic systems and engaging with its limits); Tal Z. Zarsky, *Transparent predictions*, 2013 U. ILL. L. REV. 1503 (2013); Sandra Wachter et al., *Transparent, Explainable, and Accountable AI for Robotics*, 2 SCI. ROBOTICS 1 (2017); Coglianese Cary & Lehr David, *Transparency and Algorithmic Governance*, 71 ADMIN. L. REV. 1 (2019); John Zerilli et al., *Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?*, 32 PHIL. & TECH. 661 (2019); Mike Ananny & Kate Crawford, *Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*, 20 NEW MEDIA & SOC'Y 973 (2018); Andrew Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085 (2018).

[27] *See, e.g.*, Wachter et al., *Right to Explanation*, *supra* note 5; Casey et al., *supra* note 5; Kaminski, *supra* note 5. *But see* Lilian Edwards & Michael Veale, *Slave to the Algorithm: Why a Right to an Explanation Is Probably Not the Remedy You Are Looking For*, 16 DUKE L. & TECH. REV. 18 (2017).

[28] *See, e.g.*, Matthias Leese, *The New Profiling: Algorithms, Black Boxes, and the Failure of Anti-Discriminatory Safeguards in the European Union,* 45 SEC. DIALOGUE 494 (2014); Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857 (2017); Sandra Wachter, *Affinity Profiling and Discrimination by Association in Online Behavioral Advertising*, 35 BERKELEY TECH. L.J. 367 (2020).

[29] *See, e.g.*, Sandra Wachter et al., *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, COMPUT. L. & SEC. REV., July 2021, at 1; Zarsky, *supra* note 5, at 129-30; Philipp Hacker, *Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law*, 55 COMMON MKT. L. REV. 1143 (2018).

[30] *See, e.g.*, Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI,* 2019 COLUM. BUS. L. REV. 494 (2019).

[31] *See, e.g.*, Hildebrandt & Koops, *supra* note 7, at 437; Citron & Pasquale, *supra* note 1.

[32] *See, e.g.*, Barocas & Nissenbaum, *supra* note 7; Maria Eduarda Gonçalves, *The EU Data Protection Reform and the Challenges of Big Data: Remaining Uncertainties and Ways Forward*, 26 INFO. & COMMC'N TECH. L. 90 (2017).

[33] *See, e.g.*, Schermer, *supra* note 5; Rubinstein, *supra* note 7; Paula Helm, *Group Privacy in Times of Big Data*: *A Literature Review*, 2 DIGIT. CULTURE & SOC'Y 137 (2016); Brent Mittelstadt, *From Individual to Group Privacy in Big Data Analytics*, 30 PHIL. & TECH. 475 (2017); Luciano Floridi, *Group Privacy: A Defence and an Interpretation*, *in* GROUP PRIVACY IN THE AGE OF BIG DATA 83 (Linnet Taylor et al. eds., 2017); Kammourieh et al., *supra* note 7. For an overview of the debate, see Mittelstadt et al., *supra* note 25.

[34] Such a right is established in Art. 22 of the General Data Protection Regulation. Council Regulation 2016/679 of Apr. 27, 2016, General Data Protection Regulation, art. 22, 2016 O.J. (L 119) 1, 46. Its scope, however, is narrow. *See* Wachter et al., *Right to Explanation*, *supra* note 5, at 92.

[35] Out of the vibrant scholarship working with the notion of governance to analyze

legitimacy can play a crucial role in assessing the adequacy of responses to MLB decision-making.[36]

The article proceeds in three steps. In Part II, it introduces prominent concepts to analyze corporate surveillance, namely *panopticism* and the *surveillance assemblage,* and develops an approach that allows us to exploit the analytical potential of both concepts for normative and legal assessments of MLB decision-making. It establishes criteria to distinguish between different settings of corporate surveillance and defines which concept is best suited to describe each setting.

In Part III, the article builds on the surveillance assemblage and panopticism to establish two perspectives on MLB decision-making. These perspectives provide diverging answers on a number of questions, including if and how MLB decision-making impacts the behavior of individuals, whether it provokes larger dynamics, and how it impacts society more generally.

The first perspective, building on the concept of the surveillance assemblage, highlights the opaqueness of algorithms and the fragmented and diverse nature of MLB decision-making. It depicts the individual who is judged by machine learning algorithms as lacking orientation and comprehension of her situation. In this perspective, MLB decision-making, however, does not have any larger effects on the environment in which individuals live or the way they behave. Its effects are best analyzed on a granular level, in the individual relationships between corporations and individuals.

The second perspective is based on the concept of panopticism. It argues that MLB decision-making submits human behavior to a system of sanctions and rewards, which is shaped by the correlative logic of machine learning algorithms. MLB decision-making determines the opportunities individuals have in life and the way they are perceived and treated by their surroundings. It defines how individuals need to behave to lead successful lives. It creates disciplining power, in reaction to which individuals have to self-police their lives. It is likely to influence our intuition, and with it one of

---

corporate power, see Balkin, *Three Laws of Robotics*, *supra* note 2, at 1226-28; Balkin, *Free Speech in the Algorithmic Society*, *supra* note 2; Berman, *supra* note 6; Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 Harv. L. Rev. 1598 (2018); Jonathan Zittrain, Three Eras of Digital Governance (Sept. 23, 2019) (unpublished manuscript) (available at https://perma.cc/68K7-FA65).

[36] For scholarship relying on the notion of legitimacy, see Ari Ezra Waldman, *Power, Process, and Automated Decision-Making*, 88 Fordham L. Rev. 613 (2019) (arguing that automated decision-making is an illegitimate source of authority in liberal democracies); Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, Geo. Wash. L. Rev. 1 (2019) (suggesting standards to reconcile automation with requirements of legitimacy); Pasquale, *supra* note 5 (viewing transparency requirements as necessary to legitimize algorithmic authority).

the key determinants of human behavior. The panoptic perspective provides a theoretical underpinning of the effects detected by Mayer-Schönberger, Balkin, and Pasquale. This perspective prompts us to reflect on how MLB decision-making transforms the environment in which individuals live. It allows us to leave a granular, individualistic focus behind and to engage with larger dynamics caused by corporate MLB decision-making.

Part IV builds on both perspectives and critically analyzes existing legal responses to algorithmic decision-systems. It develops a new approach to address MLB decision-making. It argues that the surveillance assemblage can be viewed as the implicit understanding of surveillance which underlies popular responses to algorithmic decision-making, including individual rights and transparency. When taken seriously as an explanatory model, however, the surveillance assemblage undermines the credibility of these responses and allows us to reveal inconsistencies. The increasing reliance on machine learning algorithms reinforces these inconsistencies and requires us to develop additional responses.

The panoptic perspective suggests that to respond to MLB decision-making adequately, we have to account for the way it transforms our environment. We have to acknowledge that MLB decision-making creates disciplining-effects and drastically increases the power of corporations to the detriment of individual freedom. The central question is no longer whether MLB systems are fair or transparent, but whether they are legitimate. Legal scholarship has begun to reconsider whether individual rights and procedural safeguards can provide adequate solutions to dramatic shifts in power caused by algorithmic systems.[37] It increasingly engages with questions of how to reconcile the impact of algorithmic systems on individuals and society with values of liberal democracies.[38] The analysis developed on the basis of the panoptic perspective adds to this scholarship, as well as to the debate focusing in particular on decision-making systems.[39] It explains why and how we should measure legal responses to MLB decision-making against standards of legitimacy. It is not individual rights, transparency, or explainability that effectively limit

---

[37] *See, e.g.*, Waldman, *supra* note 36, at 632 (challenging the "conventional wisdom that process and procedure" can prevent the harms of automated-decision-making and developing new strategies to tame algorithmic power).

[38] *See supra* notes 34 and 35; Christoph Winter, *The Challenges of Artificial Judicial Decision-Making for Liberal Democracy*, *in* JUDICIAL DECISION-MAKING: INTEGRATING EMPIRICAL AND THEORETICAL PERSPECTIVES (Bystranowski et al. eds., forthcoming); Berman, *supra* note 6 (discussing state uses of machine learning algorthms); Balkin, *Three Laws of Robotics*, *supra* note 2; Pak-Hang Wong, *Democratizing Algorithmic Fairness*, 33 PHIL. & TECH. 225 (2020); Iyad Rahwan, *Society-in-the-Loop: Programming the Algorithmic Social Contract*, 20 ETHICS & INFO. TECH. 5 (2018).

[39] *See supra* notes 25-34.

corporate power, but general principles such as data minimization and the limited purposes principle. The panoptic perspective and the suggested notion of legitimacy provide a foundation to restore the balance between the freedoms of corporations and the freedoms of individuals, which threatens to otherwise collapse in the face of MLB decision-making.

## II.   THEORETICAL FOUNDATIONS: ANALYZING SURVEILLANCE

Rarely does legal scholarship rely on surveillance theory to better understand the implications of algorithmic decision-making.[40] This Part presents one of the most prominent controversies in surveillance studies and introduces the concepts of panopticism (Part II.A, *infra*) and the surveillance assemblage (Part II.B, *infra*). Rather than merely summarizing this controversy, this Part frames the differences between concepts of surveillance in a novel way, which allows us to take advantage of the analytical strengths of both concepts to analyze MLB decision-making (Part II.C, *infra*).

### A.   Panopticism

For decades, the concept of *panopticism* has served scholars as a starting point to analyze the effects of surveillance.[41] The concept of panopticism, originally developed by Foucault, has been interpreted, modified in a variety of contexts, criticized, and defended innumerable times.[42]

The starting point for Foucault's panopticism concept is an architectural design by Bentham, the panopticon. Bentham was a utilitarian. As such, one of his objectives was to design buildings (most prominently, prisons) that served to accommodate individuals who were to be controlled by a superior authority in a way that made the use of physical

---

[40] For a useful introduction to the relationship between legal scholarship and surveillance studies, see Cohen, *supra* note 22.

[41] For Foucault's original concept of panopticism, see MICHAEL FOUCAULT, DISCIPLINE AND PUNISH: THE BIRTH OF THE PRISON 200-04 (Alan Sheridan trans., Vintage Books 2d ed., 1995) (1977).

[42] For modifications of the panopticism concept in different contexts, see Alexander Butchart, *The Industrial Panopticon: Mining and the Medical Construction of Migrant African Labour in South Africa, 1900–1950*, 42 SOC. SCI. & MED. 185 (1996); Hille Koskela, *'Cam Era' — The Contemporary Urban Panopticon*, 1 SURVEILLANCE & SOC'Y 292 (2003). For impactful critiques, see Thomas Mathiesen, *The Viewer Society: Michel Foucault's 'Panopticon' Revisited*, 1 THEORETICAL CRIMINOLOGY 215 (1997); Kirstie Ball, *Organization, Surveillance and the Body: Towards a Politics of Resistance*, in THEORIZING SURVEILLANCE: THE PANOPTICON AND BEYOND 296 (David Lyon ed., 2006); Kevin D. Haggerty, *Tear Down the Walls: On Demolishing the Panopticon*, in THEORIZING SURVEILLANCE: THE PANOPTICON AND BEYOND 23 (David Lyon ed., 2006). For a defense, see Gilbert Caluya, *The Post-Panoptic Society? Reassessing Foucault in Surveillance Studies*, 16 SOC. IDENTITIES 621 (2010).

force superfluous.[43] Bentham famously designed a prison in the shape of an octagon with a guard tower in its center. The central tower allowed a clear view into the cells of the inmates at all times. With the rules of the prison clearly laid out to the inmates, and the inmates within the gaze of the guards, the inmates had to comply with the rules or otherwise fear sanctions. The guard tower, conversely, was designed in a way that did not allow the inmates to see whether it was occupied. Consequently, it was impossible for the inmates to know when they were being observed. This provided the advantage that surveillance did not in fact have to be permanent to produce permanent effects. The panoptic prison, with inmates knowing that they could be observed at any time, created an incentive for them to self-police their behavior permanently and to comply with the rules of the prison at all times. Bentham's prison, through the ingenuity of the architectural design, converted physical power over the bodies of inmates into more discrete and effective power over their minds and their agency.[44]

Building on Bentham's architectural design, the panopticon, Foucault developed panopticism as an analytical model with which he famously described a form of exercising power, a way of governing people. [45] Foucault, in his genealogy of power, identified the dynamics Bentham created in his architectural design as defining features of the way power is exercised in what he described as "disciplinary societies."[46] In disciplinary societies, the goal of governance is not only to maintain physical control and define the outer limits of acceptable behavior, but to shape the way people lead their lives on a granular basis, to train their minds and souls. Individuals are incentivized to self-police their lives, to comply with the rules to which they are submitted at all times.

The central means to exercise power in the disciplinary society are technocratic structures that enable surveillance. As Bentham demonstrated, pervasive surveillance, permanently influencing behavior rather than simply identifying, suppressing, and punishing disobedience, allows for the non-violent and "automatic functioning of power."[47] The concept of panopticism offers a detailed account of this functioning of power, describing the dynamics in which rules, established by an authority that has the power to surveil and sanction, translate into the agency of individuals who internalize those rules.

---

[43] *See* JEREMY BENTHAM, THE PANOPTICON WRITINGS (Miran Božovič ed., 1995).

[44] Galič et al., *supra* note 21, at 12.

[45] For Bentham's designs of panopticons, see BENTHAM, *supra* note 43, at 31. For Foucault's analytical model, see FOUCAULT, *supra* note 41, at 202-05.

[46] FOUCAULT, *supra* note 41, at 209.

[47] *Id.* at 201.

*B. Surveillance Assemblage*

Panopticism, especially in more recent surveillance theory, has been criticized as an oversimplified and outdated explanatory model.[48] The most important parts of the critique relate to (1) panopticism's narrow focus on a particular kind of surveillance; and (2) panopticism's disregard for the variety of ways in which information is collected and analyzed.

Regarding its narrow focus, it has been argued that panopticism only describes one very particular constellation in which surveillance happens and does not account for other forms of surveillance.[49] Panopticism only engages with surveillance carried out by a central agent who observes a target in a confined space, built on a hierarchical relationship between the observer and the target, and characterized by the objective of the observer to transform the behavior of the target. Post-panoptic scholars have argued that these characteristics are typical for state surveillance, but do not necessarily capture the nature of other forms of surveillance.

Post-panoptic scholars observe that the possibility to collect data remotely, digitally, and automatically leads to the expansion of surveillance beyond physical confinements and makes any physical relation between the observer and the observed redundant.[50] They describe how in liberal and capitalist societies novel forms of surveillance have emerged, namely corporate surveillance, which does not have these characteristics but which still has a profound impact on the lives of individuals.[51]

Post-panoptic scholars have demonstrated that corporate surveillance is fragmented and diverse, constituting an intangible assemblage which does not fit in the explanatory structures of panopticism.[52] In the corporate surveillance assemblage, a great plurality of observers coexists, collecting information for a variety of objectives and developing distinct strategies on the basis of the information. The observers are not necessarily in a hierarchical relationship with the observed, do not have the power to sanction them, and often do not have the goal to transform the behavior of the observed. Rather, they intend to optimize their own behavior and make

---

[48] Haggerty, *supra* note 42; David Lyon, *The Search for Surveillance Theories*, *in* THEORIZING SURVEILLANCE: THE PANOPTICON AND BEYOND 3 (David Lyon ed., 2006).

[49] Mark Andrejevic*, The Discipline of Watching: Detection, Risk, and Lateral Surveillance*, 23 CRTICIAL STUD. IN MEDIA COMMC'N 391 (2006).

[50] Roger A. Clarke, *Information Technology and Dataveillance*, 31 COMMC'NS OF THE ACM 498 (1988).

[51] Gilles Deleuze, *Postscript on the Societies of Control*, 59 JSTOR 3 (1992); Gary T. Marx, *What's New About the "New Surveillance"? Classifying for Change and Continuity*, 1 SURVEILLANCE & SOC'Y 9 (2002).

[52] Kevin D. Haggerty & Richard V. Ericson, *The Surveillant Assemblage*, 51 J. SOCIO. 605 (2000).

their business strategies more efficient.[53] Post-panoptic scholars have argued that it is impractical, if not impossible, to analyze corporate surveillance in the framework of panopticism.

The second regard in which panopticism has been significantly criticized relates to its understanding of how information is collected and analyzed.[54] In novel surveillance environments, individuals are not physically but digitally observed, and collected information is not analyzed in cognitive human processes but in statistical models and data analysis systems. Post-panoptic scholars have observed that panopticism does not attribute great attention to the way information is collected and analyzed. The collection and analysis of information is simply conceived as an act of human observation. However, in digital surveillance environments, this stage of the surveillance process has changed profoundly and has gained great significance. It is in this regard, in the way information is collected and analyzed, that analog and digital surveillance differs structurally. Post-panoptic scholars have argued that panopticism does not account for the importance and complexity of this stage of the surveillance process and thus, when confronted with novel surveillance practices, reaches its conceptual limits.[55]

Changes in the way information is collected and analyzed have prompted scholars to develop alternative concepts to theorize surveillance, most prominently the surveillance assemblage. The starting point for the concept of the surveillance assemblage is the observation that the object of new forms of surveillance is the "decorporealized body." [56] The decorporealized body is "more mobile and measurable than its physical counterpart" and, in the process of analyzing the collected data, is "re-assembled in the . . . 'data double.'" [57] "Data doubles" are not comprehensive representations of physical bodies.[58] Rather, they consist in a particular perspective on a part of a person's existence that is constructed out of dispersed, collected data points to serve a particular function.[59] Many of the reconstruction processes and purposes of today's economy are well known: Advertising companies analyze data about the behavior of individuals to identify consumer preferences and target consumers; banks calculate scores for their potential customers to identify trustworthy

---

[53] *See, e.g.*, Bernard Marr, *Starbucks: Using Big Data, Analytics and Artificial Intelligence to Boost Performance*, FORBES (May 28, 2018, 2:39 AM EDT), https://perma.cc/F7LZ-XFV4.

[54] *See, e.g.*, Haggerty & Ericson, *supra* note 52, at 610-15.

[55] *Id.*

[56] *Id.* at 611.

[57] Galič et al., *supra* note 21, at 22.

[58] Haggerty & Ericson, *supra* note 52, at 605.

[59] *Id.* at 611-14.

debtors and to minimize credit defaults; law enforcement analyzes data to decide on the probability of recidivism of criminal offenders. In each of these cases, it is not the totality of a person's existence which interests the surveilling entity, but a particular subset of that existence, and a particular perspective on that subset.

The concept of the surveillance assemblage has been regarded by many scholars as the more suitable model to account for the complexity and diversity of corporate surveillance and has thus become the predominant mean to describe its effects.

### C.    Differentiating Between Surveillance Settings

The debate on what constitutes an adequate analytical framework to analyze corporate surveillance is usually conducted on an abstract and theoretical level, and it all too often leads to a dead end. The debate is characterized by a categorical dichotomy, in which panopticism is either generally defended or rejected in favor of the alternative concept of the surveillance assemblage. A more constructive way to work with both concepts consists in differentiating between different surveillance settings and in exploring what concept is suitable in consideration of the particularities of each setting and the goals of the analysis. This approach pays tribute to the fact that the dynamics produced by surveillance are different depending on the setting and that different concepts bring out different aspects. It allows the application of each framework where it promises the greatest explanatory value. This section suggests a way to distinguish panoptic surveillance settings from settings resembling surveillance assemblages and explains for what kind of analyses we should rely on each concept.

A setting can be qualified as panoptic when surveillance prompts individuals to internalize a set of rules. Panopticism describes a dynamic in which individuals modify their behavior to avoid sanctions or achieve rewards. Whenever surveillance creates a self-policing effect, or what Foucault describes as "disciplining power," it can be conceived of as panoptic.

In conventional panopticons, such as Bentham's prison, disciplining power has been intentionally created. An authority establishes rules, constantly surveils individuals, and sanctions them if they do not comply with the rules. The threat of sanctions—or conversely a promise of rewards—creates incentives for individuals to modify their behavior and internalize the rules they are expected to comply with. The rules on the basis of which the inmates are judged are disclosed to the inmates so that

they can form stable expectations of what behavior is sanctioned and what behavior is rewarded and develop strategies that allow them to avoid detrimental results and achieve favorable outcomes. While Bentham's prison is the archetype of a panopticon, not all panoptic surveillance settings have to resemble a prison. It is not guards or a central tower that characterizes a panoptic setting, but a particular mechanic of social ordering in which the collection of information and the threat of sanctions incentivizes individuals to modify their behavior.

Foucault's concept of panopticism and of "disciplining power" departs from Bentham's architectural design because it illustrates how surveillance can lead to individuals changing their behavior. Panopticism, however, is not about the particularities of an architectural design, but about a certain kind of behavioral effect. The internalization of rules is what is at the core of panopticism. When qualifying settings as panoptic, we should not focus narrowly on the concrete modalities by which discipling power is achieved. Rather, we should focus on the question of whether a given setting, through the collection of information and the threat of sanctions, incentivizes individuals to internalize a set of rules.

The valid critique of panopticism, when formulated against this understanding, consists in the observation that while corporate surveillance increases in scale, it does not create one overarching system that causes disciplining effects. The diversity of corporate surveillance prevents us from looking at corporate surveillance as a whole and analyzing on what rules it is based, if and how it distributes sanctions or rewards, and what effects it has on the agency of individuals. Corporate surveillance is characterized by a plurality of players who engage in different practices and pursue a variety of goals. The environment of corporate surveillance constitutes an assemblage rather than a panopticon. This critique does not imply that corporate surveillance has never been panoptic and has never created disciplining effects. Singular corporate surveillance practices, such as the surveillance of workers or contractors, might (even intentionally) create the dynamics characteristic of panopticism. [60] The fact that panopticism can be a suitable concept to analyze some practices of corporate surveillance, but not others, shows that the concept of panopticism is neither generally flawed nor generally adequate for analyzing surveillance. Rather, it shows that whether panopticism is a suitable framework depends on the characteristics of the *setting* we analyze and how we delimit this setting.

---

[60] Amazon, for example, is relying on automated surveillance to control and boost the performance of its employees. *See* Colin Lecher, *How Amazon Automatically Tracks and Fires Warehouse Workers for 'Productivity'*, VERGE (Apr. 25, 2019, 12:06 PM EDT), https://perma.cc/FF7C-UVK4.

When contextualized in the distinction between panoptic and non-panoptic settings, the critique of the concept of panopticism no longer takes the shape of a general rejection. Instead, it provides the basis for a differentiated assessment, in which the suitability of concepts of surveillance can be judged based on the nature of the setting that is assessed and the goal that is pursued by the analysis. The question of which concept we should rely on to assess MLB decision-making does not depend on a general theoretical argument between panopticism and the surveillance assemblage. Rather, it depends on the particularities of the surveillance setting that is created by MLB decision-making and the goals which we pursue with our analysis.

Panopticism directs the analysis towards identifying disciplining effects. It is a useful analytical concept whenever we examine surveillance settings in which such effects occur or in which we believe they are likely to occur. If we are engaging with surveillance practices that impose sanctions, enforce rules, and incentivize individuals to behave in particular ways, the concept of panopticism allows us to describe the dynamics created by these practices. Whenever surveillance practices are too diverse and fragmented to incentivize particular behavior, when they are not based on rules and do not impose sanctions, alternative concepts, such as the surveillance assemblage, might prove more fruitful.

As MLB decision-making is a rather new form of surveillance, and we are lacking empirical information on its effects, it is still too early to decide which concept describes it adequately. Such a decision, however, is also not necessary. Both concepts allow us to develop a particular perspective on the nature and effects of MLB surveillance. Both can further our understanding of the existing and potential harms of MLB decision-making. Both allow us to better understand the normative problems that legal scholars face when engaging with MLB decision-making and provide a vocabulary that can guide us in future assessments.

### III. TWO PERSPECTIVES ON MACHINE LEARNING-BASED DECISION-MAKING

This Part develops two perspectives on MLB decision-making. It builds on the concepts of the surveillance assemblage and panopticism to develop two ways of analyzing the effects of MLB decision-making. Both perspectives paint distinct, even opposing pictures of MLB decision-making and its effects. Rather than advocating for one perspective over the other, this Part aspires to develop both in their strongest shape. Both perspectives provide a unique understanding of the harms that MLB decision-making creates and that legal scholarship must address. Each perspective conceives

the challenges that MLB decision-making brings with it differently, and each raises different normative concerns.

The first perspective describes MLB decision-making on the basis of the surveillance assemblage. It argues that the reliance on machine learning algorithms does not substantially transform the effects and implications of conventional corporate surveillance and decision-making—and that the surveillance assemblage is a suitable model to describe MLB decision-making. This perspective conceives the problem of surveillance on a granular level. It highlights the opaqueness of MLB processes and the fragmented and diverse nature of MLB surveillance. In this perspective, MLB decision-making is not viewed as incentivizing individuals to modify their behavior or as provoking larger social and societal dynamics.

The second perspective, which builds on the concept of panopticism, argues that MLB decision-making transforms the environment in which individuals live and incentivizes them to self-police their behavior. It engages with the overarching effects of MLB decision-making and the social and societal dynamics it creates. This perspective argues that while conventional forms of corporate surveillance might not have significantly affected the behavior of individuals, MLB decision-making is fundamentally different. It suggests that the reliance on machine learning algorithms in decision-making has the effect of creating a coherent set of rules on the basis of which individuals are sanctioned or rewarded. MLB decision-making incentivizes individuals to alter their behavior, thus transforming the assemblage of corporate surveillance into a panoptic surveillance setting.

### A.   Corporate Surveillance on Steroids

When developing a perspective on MLB decision-making through the surveillance assemblage framework, it is helpful to first engage with automated decision-systems more generally and then explore how the increasing reliance on machine learning algorithms affects our analysis of these systems. To develop that analysis, it is important to clearly define and distinguish between conventional automated decision-systems and MLB decision-systems. Like automated decision-systems, MLB decision-systems function on the basis of automation and algorithms. MLB decision systems can thus be conceived as a sub-category of automated decision-systems– and this article is interested in that increasingly popular and influential subcategory.

When referring to automated decision-making rather than MLB decision-making, this article refers to practices in which corporations use decisions-systems that are automated but do not rely on information

generated by machine learning algorithms. Automated decision-systems are systems that analyze relatively small data sets on the basis of a predefined set of factors. They are programmed by a human, and they strictly execute the commands the human has defined.

MLB decision-systems, in contrast, rely on algorithms that analyze a sheer, unlimited amount of data and, in the process of analyzing this data, refine the strategies based on which they infer information. MLB decision-systems are particular in that they analyze information according to machine learning algorithms and the results they generate are shaped by the logic underlying machine learning. The notion of 'Big Data' often refers to this exact particularity.

Qualifying automated decision-making as corporate surveillance does not provide any difficulties. Automated decision-making is based on information that has been collected about individuals and can thus be conceived as a form of surveillance. The argument that it should be conceived as a surveillance assemblage also applies. Post-panoptic surveillance scholars have insisted that players involved in corporate surveillance are numerous and diverse, that the goals they pursue and the way they pursue them differ, and that there is no consistent system or set of rules connecting behavior to consequences. They have argued that corporate surveillance, as a whole, does not create a stable set of incentives and that it is therefore impossible for individuals to comprehend the workings of corporate surveillance, let alone to develop strategies to optimize outcomes.

All of these objections to conceptualizing corporate surveillance as a panopticon can be raised in a similar form to conceptualizing automated decision-making as a panoptic setting. Automated decision-making only possesses one of the three characteristics—the existence of sanctions, of rules, and of incentives to modify behavior—that make panoptic settings and that produce disciplining effects. It imposes sanctions, or rewards, but these sanctions are not based on rules and thus do not create incentives for individuals to modify their behavior.

One could already challenge that automated decision-systems distribute sanctions and rewards. It can be argued that corporations who rely on automated decision-making do not have the intention of sanctioning. Rather, they aim at minimizing risks and optimizing business strategies. As these corporations do not intend to sanction individuals, results produced by their decision-making systems should not be conceived as sanctions.

This argument, however, is not convincing. It disregards that the practical effect that automated decision-systems have on the lives of

individuals does not depend on the intention with which such a system is employed. Rather, what is decisive when exploring the effects on individuals is how these individuals experience the decisions. The experience of an individual receiving a good evaluation and thus a low interest or insurance rate, or receiving a bad evaluation and thus no access to a loan or insurance at all, is similar to being rewarded or sanctioned.

The results of MLB decisions, such as not having access to credit, to housing and employment, or being unable to find insurance for oneself and one's family, can have a devastating impact on a person's life. It can practically prevent the pursuit of dreams and the free development of personality. Financing college or building up a business can become an impossible task if access to credit is refused. The dignity of work may be denied to those who receive poor MLB evaluations and unfavorable decisions. They may no longer be able to meet their most basic needs, such as caring for health or providing their children with education. MLB decision systems can make building a life worth living impossible.

What matters to the individual is the negative impact of such events on her life–whether that impact is experienced as the result of a formal sanction or a corporate decision is of peripheral importance. Even if one regards automated decision-making as a surveillance assemblage, it seems reasonable that the decisions these systems make can be both detrimental or favorable, and thus can be conceived of as sanctions or rewards. Pasquale recognizes this, too, and astutely formulates that "[g]overnments and firms in capitalist democracies increasingly use automated processes to allocate punishments and rewards."[61]

The reason why the surveillance assemblage rather than panopticism might be the more suitable concept to analyze automated decision-making is not because automated decisions cannot be conceived as sanctions, but because these sanctions are not based on a consistent set of rules and thus do not create incentives for individuals to modify their behavior. The way corporations collect data about individuals, the way they analyze it, and the way they make decisions differs from corporation to corporation. Corporations might rely on different databases and algorithms, causing their automated systems to produce different results. As there is no one singular system of automated decision-making on which all corporations rely, there is no singular way in which favorable or detrimental treatments are distributed.

As all corporations develop their own distinct strategies, there is no consistent logic, no general set of rules on the basis of which individuals' behavior    is    connected    to    consequences.    The    diversity    and

---

[61] Pasquale, *supra* note 5, at 235.

incomprehensibility of automated systems make it impossible for individuals to develop stable expectations and to anticipate what behavior leads to what outcomes. There are no graspable rules that individuals could try to comply with. Thus, automated decision-making, even though it creates a system of sanctions and rewards, does not provoke the self-policing dynamics characteristic for panoptic surveillance settings. The lack of a coherent set of rules translates into a lack of incentives for individuals to modify their behavior and, hence, prevents the emergence of disciplining effects.

If automated decision-making does not create the dynamics characteristic for panoptic settings, making the argument that MLB decision-making creates no such dynamics is rather straightforward. To make that argument, one does not have to claim that machine learning algorithms do not transform automated decision-making at all. One can acknowledge that machine learning algorithms might make automated decision-making more effective, more broadly deployable, and more popular. One only has to defend the position that gains in effectiveness and scale do not transform the dynamics at play, and that MLB decision-making is still diverse, is not based on rules, does not create incentives, and thus does not cause disciplining effects.

Machine learning algorithms can be viewed as tools for corporations to ameliorate their conventional forms of decision-making. Machine learning algorithms allow corporations to analyze large data sets and to detect correlations and patterns in these data sets, which previously could not be detected.[62] Processing large amounts of information at relatively low costs, machine learning algorithms promise to make automated decision systems cheaper, which, in turn, allows corporations to rely on them not only in sectors such as finance, housing, or employment, but in practically all spheres of life.[63] These algorithms can be used to infer information about individuals. Inferred information includes predictions relating to previously unknown characteristics or behaviors of an individual. Inferred information can then be fed into automated decision systems, making decision-making more opaque, as it is neither obvious what data points the decisions systems account for, nor how they have been generated. Inferred information also makes the systems more efficient, as it adds to the data that can be analyzed and often reveals important characteristics and patterns of behavior.[64] MLB decision-making, from this perspective, is

---

[62] MAYER-SCHÖNBERGER, *supra* note 6.

[63] For more detailed analysis of the functionality of machine learning and the notion of Big Data, see *supra* note 6.

[64] PASQUALE, *supra* note 3.

automated decision-making on steroids, not different in kind, but only different in scale.

If machine learning algorithms merely contribute to the spread of automated decisions-systems and make their use cheaper, more efficient, and opaque, there is no reason to believe that they transform the nature and effects of the surveillance setting that they create. The adaptation of machine learning algorithms might empower corporations to engage in more comprehensive surveillance, more precise data analysis, and the large-scale rollout of algorithmic decision-systems, while saving costs.

Nonetheless, the landscape of corporate decision-making still remains diverse and fragmented, with many players striving for different goals in different ways. It remains impossible for individuals to comprehend the workings of MLB decision-making and unlikely that individuals modify their behavior in reaction to it. The opacity and complexity of machine learning algorithms, in fact, are likely to further reduce the already limited understandings of individuals of the working of automated systems. Thus, MLB decision-making is even more unlikely to provoke behavioral changes among individuals. In the absence of disciplining effects, the surveillance assemblage remains an adequate framework to analyze the effects of MLB decision-making.

### B.   A New Form Of Corporate Surveillance

MLB decision-making can also be conceived as a genuinely new form of corporate surveillance. Developing a second perspective on MLB decision-making, this section argues that the large-scale employment of machine learning algorithms has a unifying effect on the previously diverse landscape of corporate surveillance and decision-making. MLB decision-making establishes a system of sanctions and rewards that incentivizes individuals to behave in certain ways. It transforms the decision-making environment of individuals and is likely to create discipling effects—what Balkin has referred to as the harm of "normalization (or regimentation)." MLB decision-making creates a new panoptic surveillance setting.

### 1.   The Unifying Effect of MLB Analysis

The notion of the "unifying effect" stands for the idea that, as corporations begin to rely on machine learning algorithms to make decisions, their decisions are shaped by the particularities of the results produced by these algorithms. Thus, even if there is a great variety of corporate actors, they begin to perceive the world and make decisions in similar ways. The idea of the unifying effect stems from the observation that corporations increasingly rely on knowledge generated by machine

learning algorithms. To understand if and how machine learning algorithms transform decision-making and whether that has a unifying effect, we have to understand the particularities of the results that such algorithms create.

Understanding what characterizes the results produced by machine learning algorithms does not presuppose in-depth expertise regarding the technological functionality of such algorithms. For the purpose of this article, it suffices to describe the basic traits of the way machine learning algorithms—or what is often described as Big Data—operate and to identify how these traits shape the results that algorithmic systems produce.[65]

Although the designs of machine learning algorithms vary, the foundational logic on which they operate is relatively similar, and its basic traits can be described in a general manner.[66] Before machine learning algorithms are used to infer information about particular individuals, they are trained on large data sets. Data sets will typically contain all kinds of information about individuals, such as GPS locations, social media activity, and shopping habits, as well as information about particular outcomes that are relevant for the purpose for which the decision-making system is employed, such as financial reliability or health. In the process of training, algorithms identify correlations and patterns within the data set, showing how data relates to relevant outcomes in the future. The algorithm identifies data points as predictors for future outcomes. Ultimately, on the basis of the knowledge previously generated, algorithms evaluate and make decisions over particular cases. What characterizes the process of MLB systems is the fact that they analyze data sets comprehensively, require no prior hypotheses, and require minimum human intervention. They are "autonomous" in the sense that once set up, they infer information and produce results independently from human involvement. Conventionally, corporations decided on what data to collect and analyze, how to interpret it, and what consequences to draw from its interpretation on the basis of distinct strategies, in processes shaped by human decisions.

The reason why machine learning algorithms are likely to have a unifying effect on decision-making systems is that by reducing human involvement in decision processes and replacing it with an algorithmic assessment, they also diminish the particularities of such processes. All

---

[65] For a concise and accessible account of how machine learning algorithms create knowledge, see Hildebrandt, *supra* note 6. For a focus on the particularities of knowledge created by machine learning algorithms, see boyd & Crawford, *supra* note 6; MAYER-SCHÖNBERGER, *supra* note 6.

[66] For more detailed account of how machine learning algorithms operate, see MAYER-SCHÖNBERGER, *supra* note 6; Berman, *supra* note 6; Hildebrandt & Koops, *supra* note 7, at 431-33; Wachter & Mittelstadt, *supra* note 30, at 505-11.; Kammourieh et al., *supra* note 7, at 40-43.

processes are increasingly shaped by inferences produced and correlations detected by machine learning algorithms. The phenomenon that different algorithmic systems produce relatively similar results, that the use of algorithms to make decisions eliminates diversity and fosters uniform outcomes, has been widely recognized among scholars examining biases and discriminatory effects of algorithmic systems. [67] Scholars have convincingly argued that machine learning algorithms, generating evaluations and predicting the future on the basis of data about the past, replicate existing social structures that themselves have been the result of social injustices and discrimination. This scholarship is based on the observation that with the large-scale use of MLB systems to form corporate strategies, a plurality of players no longer translates into a diversity of results. This observation is vital to understanding the great discriminatory potential of algorithmic processes and automated decision-making. Understanding shared effects of algorithmic systems is also vital to appreciating the nature and effects of the surveillance setting that they create.

The unifying effect of MLB data analysis can also be described with the vocabulary that has originally been used to demonstrate the diverse and fragmented nature of corporate surveillance processes. Scholars have established that in digital surveillance environments, observation does not consist in the watching of the physical body, but in the recording and measuring of the decorporealized body. [68] In modern surveillance processes, knowledge about individuals is not simply collected, but is generated through the re-assembling of granular information to data doubles.[69] Scholars criticizing the concept of panopticism have argued that the process of re-assembling collected data for particular purposes is the central feature of modern digital surveillance processes—and that panopticism does not account for the complexities of this stage of the surveillance process. Post-panoptic scholars demonstrate that the process of reassembling data, as it is carried out by different corporations in different ways, introduces diversity into surveillance practices. [70] It is this process of reassembling data, today more commonly referred to as data analysis or processing, which is increasingly carried out by machine learning

---

[67] *See, e.g.*, Barocas & Selbst, *supra* note 7; Tal Z. Zarsky, *Understanding Discrimination in the Scored Society*, 89 WASH. L. REV. 1375 (2014); Berman, *supra* note 6, at 1326-28*;* Leese, *supra* note 28; boyd & Crawford, *supra* note 6, at 666-68; Matthias Spielkamp, *Inspecting Algorithms for Bia*s, 120 MIT TECH. REV. 96 (2017); Zarsky, *supra* note 5, at 126-27; Hildebrandt & Koops, *supra* note 7, at 436-37; Ignacio N. Cofone, *Algorithmic Discrimination Is an Information Problem*, 70 HASTINGS L.J. 1389 (2019); Kim, *supra* note 28.

[68] Haggerty & Ericson, *supra* note 52, at 611.

[69] *Id.* at 605.

[70] *See supra* Part II.B*.*

algorithms. The process of data analysis, as described by post-panoptic scholars with regard to conventional forms of corporate surveillance, originally had a fragmenting and diversifying effect. With the growing employment of machine learning algorithms to carry out the analysis of data, reassembling data doubles on the basis of a similar logic, this stage of the process may have a unifying effect on corporate surveillance and MLB decision-making.

The emergence of large, specialized companies that calculate scores and profiles and sell these to other corporations is likely to further fuel unifying effects of MLB decision-making. In the field of credit ratings, in which scores and profiles are most prominently used today, a few specialized companies dominate the market. Credit scores in the U.S. are calculated primarily by three companies: Experian, TransUnion, and Equifax.[71] These companies provide information on the basis of which the financial services market operates. Their ratings determine whether individuals get access to loans to open their businesses, can get a mortgage to buy a house, or take out a loan to finance the studies of their children. As scores and ratings are used in more fields, we must expect similar market structures in these fields, too.

As machine learning algorithms and Big Data further complicate the way scores and ratings are calculated, centralization tendencies are likely to further intensify. Small companies have neither the data nor the resources to train machine learning algorithms. To take advantage of these new technologies, they will have to rely on the services of large, specialized companies, who have the resources to develop sophisticated machine learning algorithms and gather large data sets. It is therefore likely that a few large companies will dominate the market of MLB data analysis in other commercial sectors as well. Thus, the data and algorithms of these few companies will shape the way individuals are treated by the majority of corporations, relying on the services of these companies. For their automated decision-making to be efficient, all corporations would then have to rely on the extraordinary data analysis capacities of a few, and the algorithms of those few corporations will have a tremendous impact on the lives of individuals.

---

[71] For a brief introduction to American Credit Scores, see Joshua Rosenblat, *What You Should Know About Credit Scores*, N.Y. TIMES (July 23, 2017), https://perma.cc/Z57W-UHDF.

*2.   A Panoptic Surveillance Setting*

A panoptic surveillance setting has been defined as one in which sanctions and rewards are distributed on the basis of a uniform set of rules, creating incentives for individuals to alter their behavior. If we accept the claim that machine learning algorithms have a unifying effect on the decision-systems of corporations, this prompts us to revisit how we conceptualize the harms of MLB decision-making. This section explores arguments for conceiving the environment created by MLB decision-making as a panoptic surveillance setting.

The idea that the ubiquitous collection, analysis, and use of information by corporations is likely to influence the behavior of individuals increasingly surfaces in academic engagements with Big Data. Philip Agre, an early surveillance studies scholar, argues that "people whose activities are being captured will probably adjust their conduct based on their understanding of what will become of the data and what this entails for their own lives."[72] More recently, Viktor Mayer-Schönberger, engaging with MLB data analysis, anticipates that "subsequent generations may develop a 'big data consciousness'—the presumption that there is a quantitative component to all that we do."[73] Jack Balkin, as explained earlier, explores how individuals might self-police their behavior in response to digital and algorithmic surveillance, relying on the notion of "normalization (or regimentation)."[74] He argues that algorithmic systems cause people to alter their "identity, behavior, or other aspects of personal self-presentation in order to appear less risky to the algorithm, or to fall into a different category," or to "engage in behavior that the algorithm does not pay attention to."[75] The "normalizing" or "regimenting" effects Balkin describes are the very effects that the concept of panopticism engages with and desires as disciplining effects of surveillance.

The dynamics that Agre, Mayer-Schönberger, and Balkin describe engage with behavioral changes. Behavioral changes are at the center of another debate that prominently engages with manipulative practices based on machine learning algorithms. While both debates engage with behavioral changes, they examine fundamentally different practices and are concerned with different normative implications; it is therefore important to clearly distinguish between them.

---

[72] Philip E. Agre, *Surveillance and Capture: Two Models of Privacy*, 10 THE INFO. SOC'Y 101, 111 (1994).

[73] MAYER-SCHÖNBERGER, *supra* note 6, at 97.

[74] Balkin, *Three Laws of Robotics*, *supra* note 2, at 1238.

[75] *Id.* at 1238-39. Scholars have also focused on discovering behavioral changes online. *See, e.g.*, sources cited *supra* note 18 (examining the emergence of chilling effects online).

The debate on manipulation engages with targeted advertisement, practices in which MLB data analysis and psychological profiles are employed to nudge, seduce, and deceive individuals into particular behavior.[76] In this debate, manipulative practices are critically assessed because they jeopardize conditions vital for the formation of free will, obstruct unbiased decisions, and intentionally and subconsciously manipulate the individual.

The debate on algorithmic decision-making, conversely, engages with corporations individualizing the conditions on the basis of which they enter into contracts with consumers. The goal of such decision-making is not to manipulate individuals but to minimize risks resulting from unreliable contract partners. Algorithmic decision-making is not problematic because it manipulates individuals, but because it can limit the opportunities available to them. Behavioral changes are not intentionally provoked; they result from individuals policing their own behavior in order to avoid detrimental conditions and treatments.

While research on behavioral changes caused by manipulation is vast and advanced, research on self-policing effects provoked by algorithmic decision-making is rare. Scholarship has begun to identify the dynamics provoked by algorithm-based decision-making, but the origin, nature, and consequences of these dynamics, especially in the context of corporate surveillance, have yet to be theorized in detail. The categories developed in this article to identify panoptic surveillance settings provide a framework to develop a more comprehensive account of these dynamics and can help to explore behavioral changes beyond manipulation.

Surveillance scholars are engaging with the effects described by Mayer-Schönberger and Balkin in greater detail. They have argued that with the omnipresence of information technology and profiling practices, panopticism has become a ubiquitous reality.[77] Enabled by novel means of gathering information about the lives of individuals, the shift of life into the digital sphere, and the "datafication" of analog life, panopticism is no longer limited to confined spaces.[78]

While this work is promising and substantiates observations made by Mayer-Schönberger and Balkin, it mainly engages with the expansion of surveillance that is hierarchical and characterized by one observer. While

---

[76] *See supra* note 8.

[77] Caluya, *supra* note 42. With the rise of digital data processing, scholars have already engaged with the emergence of an "electronic panopticon." David Lyon, *An Electronic Panopticon? A Sociological Critique of Surveillance Theory*, 41 THE SOC. REV. 653 (1993). *See also* Thomas McCullan, *What Does the Panopticon Mean in the Age of Digital Surveillance?*, GUARDIAN (July 23, 2015), https://perma.cc/FNR7-QAYY.

[78] On the notion of "datafication," see MAYER-SCHÖNBERGER, *supra* note 6, at 76-99.

such accounts are insightful with regard to electronically empowered state surveillance, they do not explain when surveillance, such as corporate surveillance, could be perceived as panoptic. To understand if the reliance of corporations on MLB decision-making creates a panoptic setting, we need to engage with the objections raised against panopticism—and assess more carefully whether in an environment chartered by MLB decision-making a system of rules and sanctions emerges, which incentivizes individuals to modify their behavior.

There is little doubt that decisions taken by corporate algorithmic system can have either favorable or detrimental effects on individuals—and as such, from the point of view of the individuals, can be experienced as rewards or sanctions.[79] Hard questions regarding MLB decision-making as a panoptic setting consider the existence of a coherent set of rules underlying these sanctions—and the comprehensibility of these rules for individuals.

It has already been argued that MLB algorithms can have a unifying effect on automated systems. Generating knowledge in similar ways, such systems create relatively similar decisions. To understand how these similarities translate into a shared set of rules underlying MLB systems, it is helpful to conceive of MLB data analysis as a practice of grouping.[80] MLB generation of knowledge can be described as the practice of grouping individuals on the basis of their behavior and characteristics and deducing probabilities from the fact that an individual belongs to a particular group.[81] Groups, in this context, are not stable and enclosed entities. Rather, a great number of groups coexist, they partly overlap, and their composition changes depending on the behavior of individuals.[82] Furthermore, each individual is part of a large number of groups, as every data point reflecting behavior clusters her with different individuals. The machine learning algorithms then identify correlations between groups and outcomes.[83] Subsequent to the training, the identified patterns and correlations between groups and outcomes are used to predict relevant future outcomes in individual cases.

---

[79] Pasquale, *supra* note 5, at 235.

[80] Some scholars focus on the category of groups when describing and responding to MLB data analysis. *See, e.g.*, Kammourieh et al., *supra* note 7; Luciano Floridi, *Open Data, Data Protection, and Group Privacy*, 27 PHIL. & TECH 1 (2014); Mittelstadt, *supra* note 33. Bart van der Sloot, *Do Groups Have a Right to Protect Their Group Interest in Privacy and Should They? Peeling the Onion of Rights and Interests Protected Under Article 8 ECHR*, *in* GROUP PRIVACY IN THE AGE OF BIG DATA 197 (Linnet Taylor et al. eds., 2017).

[81] Kammourieh et al., *supra* note 7, at 40-43; Mittelstadt*, supra* note 33, at 477-81.

[82] Linnet Taylor et al., *Introduction: A New Perspective on Privacy*, *in* GROUP PRIVACY IN THE AGE OF BIG DATA 1, 6-8 (Linnet Taylor et al. eds., 2017).

[83] For more detailed descriptions of how machine learning algorithms create knowledge, see Hildebrandt, *supra* note 6; Hildebrandt & Koops, *supra* note 7, at 431-33.

In light of this description of MLB systems, we can identify a rule that defines how individuals are evaluated and treated by these systems. The way individuals are evaluated and treated—whether they are sanctioned or rewarded—depends on what groups they are clustered into and how these groups correlate with future outcomes. Hence, if a person wants to control the way she is perceived, evaluated, and treated in an environment of MLB decision-making, she has to police with whom she is clustered by a machine learning algorithm. She has to control everything she does that leaves a data trace, wondering what group that data point puts her into, what outcomes correlate with her membership in that group, and what consequences being classed into that group has for the perception of herself.

While this rule appears trivial and indeterminate at first, its implications are, at a closer look, profound and tangible. They can be acutely illustrated with an example: consider the detrimental financial consequences that a couple can suffer when going to marriage counselling.[84] Going to marriage counselling puts a couple into a particular group—a group that correlates with the outcome of financial distress. Couples who go to marriage counselling have a higher than usual rate of getting divorced, and individuals who are divorced are more likely than others to suffer financial hardship.[85] Thus, a credit card company that uses an algorithmic system that discovers this correlation automatically increases the interest rates of that couple—thus further contributing to the risk of financial hardship.

What's more, an algorithm that identifies the correlation does not even need to have direct access to databases that include information on what couples go to marriage counselling. It can deduce such information, too, from data that is much more accessible, such as the location data of the couple. This effectively means that individuals who use apps that track their location data must expect that inferences such as the one described here shape the ways in which they are encountered by corporations.

Pasquale describes the increase in interest rate due to the credit company's algorithm as a "penalty" for the couple: "When statistics imply that couples in counseling are more likely to divorce than couples who aren't, counseling becomes a 'signal' that marital discord may be about to spill over into financial distress. This is effectively a 'marriage counseling penalty.'" Furthermore, "[o]nce disclosed, it could discourage a couple

---

[84] PASQUALE, *supra* note 3, at 31.

[85] Duhigg describes how credit card companies rely on marriage therapy as an indicator for potential depression, which again brings with it risks of job loss. Charles Duhigg, *What Does Your Credit-Card Company Know About You?*, N.Y. TIMES (May 12, 2009), https://perma.cc/DE3J-M3PA.

from seeking the counseling they need to save their relationship."[86] To avoid a "sanction" in the form of higher interest rates, couples have to make sure they do not fall into the group "couple that requires marriage counselling"—which entails behaving in one way and not in another.

What this example shows is that in an environment of pervasive data collection and MLB decision-making, everything that a person does produces echoes in the form of MLB decisions. Every movement, every habit, every hobby, every friendship, and every purchase has consequences that are defined by the correlative logic of machine learning algorithms. MLB systems sanction individuals for behavior that correlates with unfavorable outcomes and reward them for behavior that correlates with success.

Thus, MLB decision-making establishes a system that evaluates the social interactions of individuals and is likely to influence how individuals interact. It has the potential of creating dynamics that alter society more generally. Individuals striving for advantageous results in an environment of MLB decision-making have to avoid being part of groups that, according to MLB analyses, indicate a lack of discipline, financial unreliability, bad health, or any other unfavorable outcome relevant in MLB decision-making. MLB decision-making creates incentives for those who can afford it to retreat into environments that indicate success.

If we accept that MLB decisions have the character of sanctions or rewards, and that there is a general rule underlying the way in which such sanctions and rewards are distributed, the question remains whether individuals are likely to comprehend this rule. For the rule to translate into incentives and ultimately result in behavioral changes of individuals—which are characteristic for panoptic settings—individuals must develop a basic understanding of that rule.

For good reasons, we may doubt that this is the case. Scholars accurately describe the algorithmic processes that create scores and digital reputations and thus ultimately decide on how an individual is treated as complex, opaque, and non-transparent.[87] Consequently, it is hard for individuals to develop an understanding of these rules and to develop norms that help them to successfully navigate through MLB-based decision-making environments. However, there are also good reasons to believe that individuals might adapt to the environment of MLB decision-making faster than one might expect.

First, technical expertise is not necessary to adapt one's behavior to the rewarding and sanctioning regime of MLB decision-making. For the couple,

---

[86] PASQUALE, *supra* note 3, at 31.

[87] *Id*.

the decision to go to marriage counselling does not have to revolve around the complexities of machine learning algorithms for them to understand that their decision can provoke financial burdens. A suspicion, a news report, or a talk by scholars and public figures can be enough to instigate them to modify their behavior.

Furthermore, changes to behavior do not presuppose a clear understanding of the rules to which that behavior is submitted. Daniel Kahneman has famously argued that human behavior is only partly driven by deliberative and logical reasoning.[88] More important is often human intuition. Ubiquitous MLB personalization is likely to influence our intuition, and with that one of the key determinants of human behavior.

Humans might indeed be unlikely to understand the technicalities of MLB decision-making on an intellectual level. They are, however, likely to learn to comply with indicators of success by intuition. Individuals are likely to learn to play by the rules, avoid sanctions, and achieve rewards, despite the fact that they do not fully understand the technological processes through which those sanctions and rewards are generated.

Secondly, an abundance of policy initiatives and intense media coverage of "artificial intelligence" contribute to the awareness of the relevance and effects of MLB data analysis. In the sector of credit scores, there already are plenty of guides online on how to behave to improve one's score.[89] Similar guides are likely to emerge in all relevant fields in which MLB decisions have tangible consequences for individuals, educating them on how to comply with indicators of success. Calls to make algorithmic evaluations more transparent and make mandatory explanations on how an algorithm makes decisions, which aim at empowering individuals to learn about the processes shaping their opportunities, contribute to the layperson's consciousness of algorithmic processes.

Thirdly, MLB personalization, now penetrating the offline environment, is not unprecedented. Permanent data collection, analysis, and decision-making already characterizes the online environment, in which many individuals spend large parts of their time—and many individuals are well aware of these practices. Many people are well aware of the fact that their browsing behavior, their GPS locations, and their shopping habits determine the advertisements that will be displayed to them in the future.

---

[88] KAHNEMAN, *supra* note 9.

[89] *See, e.g.*, Rebecca Lake, *How to Improve Your Credit Score*, INVESTOPEDIA (Feb. 12 2021), https://perma.cc/V28Q-3BAN; Bev O'Shea, *How to Improve Credit Fast,* NERDWALLET (Oct. 18, 2021), https://perma.cc/MG94-EXH9. Experian even provides guidance on how to improve the score one receives from it itself. *How to Improve Your Credit Score,* EXPERIAN, https://perma.cc/4SJ4-8PVW. The issue of credit scores has been examined by Citron & Pasquale, *supra* note 1, at 11.

They understand that the selection of advertisements displayed to them is not only based on data about themselves, but also on data about people who have similar interests and hence similar behavioral patterns. "Customers who bought this item also bought . . ." is a familiar notice to all online shoppers. The logic underlying machine learning algorithms governing targeted advertisement is similar to the logic of MLB decision-making. Awareness for MLB decision-making outside the online space is likely to develop in similar ways and will build on individuals' already existing familiarity with similar practices online.

While it might, at a first look, seem far-fetched that individuals comprehend the logic on the basis of which machine learning algorithms distribute sanctions and rewards, a closer look shows that an understanding is likely to develop over time. Although it is still too early to ascertain with certainty—and assess empirically—how MLB decision-making affects the behavior of individuals, a panoptic perspective offers starting points to discuss trends that are taking shape and that might have a profound impact in the future. Conceiving MLB decision-making as a panoptic surveillance setting allows us to engage with behavioral dynamics and harms it might provoke over time. While these effects might be negligible at the moment, they might play a defining role in how MLB decision-making impacts individuals' lives and society in the future.[90]

## IV.    REVISITING THE LEGAL DEBATE

Both the panoptic perspective and the perspective of the surveillance assemblage draw distinct pictures of MLB decision-making. Both develop different understandings of the problems with which MLB decision-making confronts us. Both provide a basis to discuss the adequacy of existing regulatory approaches—and can provide foundations for developing new ones.

In Part IV.A, existing legal responses are revisited against the background of the perspective of the surveillance assemblage. Then, Part IV.B uses the panoptic perspective as a foundation to suggest a new approach to respond to MLB decision-making.

---

[90] On the merit of long-term perspectives (and how to reconcile them with short term perspectives) on AI, see Seth D. Baum, *Reconciliation Between Factions Focused on Near-Term and Long-Term Artificial Intelligence,* 33 AI & SOC'Y 565 (2018); Carina Prunkl & Jess Whittlestone, *Beyond Near- and Long-Term: Towards a Clearer Account of Research Priorities in AI Ethics and Society*, ARXIV 1 (Jan. 21, 2020); Stephen Cave & Seán S. ÓhÉigeartaigh, *Bridging Near- and Long-Term Concerns About AI*, 1 NATURE MACH. INTEL. 5 (2019).

### A.  Responding to the Surveillance Assemblage

The concept of surveillance on which large parts of legal scholarship on automated decision-making are implicitly based, is that of the surveillance assemblage. When made explicit, the understanding of surveillance underlying these responses can both advance such scholarship and expose limitations.

In the following, popular responses to automated decision-making are briefly presented (Part IV.A.1, *infra*). It is then shown that the surveillance assemblage can be viewed as the implicit understanding of surveillance that underlies these responses (Part IV.A.2, *infra*). However, when taken seriously as an explanatory model, the surveillance assemblage undermines the credibility of existing responses and reveals inconsistencies (Part IV.A.3, *infra*). The increasing reliance on MLB algorithms reinforces these inconsistencies. Revisiting legal responses on the basis of the concept of the surveillance assemblage suggests that these responses are increasingly inadequate, even if judged on the basis of the premises on which they are based.

### 1.  Rights and Transparency

A large part of existing responses to MLB decision-making focus on two elements: guaranteeing transparency and explainability and granting individual rights. [91] In these approaches, informing the individual and empowering her through rights is viewed as being key to combating the harmful effects of MLB decision-making.

One of the main objectives of transparency[92] and explainability[93]—and the function that this article focuses on—is to increase the individual's understanding of the ways in which she is surveilled. Transparency and explainability are also viewed as important preconditions for regulators to analyze the functionality of MLB systems and to intervene whenever the operations of such systems are viewed as problematic, and for the public to engage with the implications of such systems. This article does not dispute the virtue of transparency and explainability with regard to these purposes. Rather, it engages with the limits of transparency and explainability with regard to their function of informing and empowering

---

[91] For overviews on responses to automated decision-making, see FJELD ET AL., *supra* note 25 and Mittelstadt et al., *supra* note 25. For a variety of detailed suggestions, see Kroll et al., *supra* note 25.

[92] *See supra* note 26.

[93] *See* Wachter et al., *Right to Explanation*, *supra* note 5; Casey et al., *supra* note 5; Kaminski, *supra* note 5. *But see* Edwards & Veale, *supra* note 27.

the individual. As far as standards of transparency and explainability are criticized, it is not because they do not serve the function of allowing the public and regulators to engage with MLB decision-making, but because they fail to sufficiently protect and empower the individual.

The individual rights approach consists of rights to non-discrimination, [94] fairness, [95] reasonableness, [96] due process, [97] data protection,[98] privacy,[99] and to not be subjected to automated decision-making.[100] These rights aim at empowering the individual to take action against specific surveillance practices by specific corporations. Rights that have previously served the individual to defend her interests in the workplace or when interacting with the state are viewed as being equally suitable to empower the individual to challenge corporate decision-making practices.

What all the suggested responses—explainability and transparency and the array of individual rights—have in common is that they aim at furthering the individual's understanding of her situation and at empowering the individual to defend her interests vis-à-vis those corporations who surveil and make decisions over her.[101] The individual is viewed as an agent who, if we give her the right tools, will defend her own interests. All of these responses construct the battle line of corporate surveillance as existing between the individual and a particular corporation who has data and makes decisions about the individual. They are characterized by an atomistic and individualistic focus and aim at furthering the situation of the individual by increasing her understanding and granting her rights to defend her interests against those corporations that she believes treat her unfairly.

## 2.   Implicit Foundations of the Transparency and Rights Approach

Transparency and rights-based legal responses are based on a particular understanding of the subject matter that they address, even though this understanding remains largely unconscious and implicit. The

---

[94] *See supra* note 28.

[95] *See* Zarsky, *supra* note 5; Hacker, *supra* note 29.

[96] *See* Wachter & Mittelstadt, *supra* note 30.

[97] *See* Hildebrandt, *supra* note 6; Citron & Pasquale, *supra* note 1.

[98] *See* Hildebrandt & Koops, *supra* note 7; Barocas & Nissenbaum, *supra* note 7; Gonçalves, *supra* note 32.

[99] *See* Schermer, *supra* note 5; Rubinstein, *supra* note 7; Helm, *supra* note 33; Mittelstadt, *supra* note 33; Floridi, *supra* note 33; Kammourieh et al., *supra* note 7. For an overview of the debate, see Mittelstadt et al., *supra* note 25.

[100] *See supra* note 34.

[101] Waldman, *supra* note 36, at 622-24 (arguing that a variety of legal responses have commonality in being "process oriented").

understanding of surveillance that implicitly underlies transparency and rights-based legal responses is the concept of the surveillance assemblage.

The concept of the surveillance assemblage does not view surveillance as a collective phenomenon that has shared social and societal consequences, but conceives the problem of surveillance on a granular level. The surveillance assemblage describes corporate surveillance as consisting in billions of singular relationships, and it rejects accounts of overarching dynamics and effects provoked by surveillance.

As the problem is perceived on a granular level, responses are constructed on a granular level, too. Suggesting individual rights such as rights to non-discrimination, fairness, reasonableness and due process as a remedy to automated decision-making makes sense, if we believe that the most important effects of surveillance occur on an individual level, as effects of singular surveillance relationships. The individualistic understanding of the problem is mirrored in an individualistic response.

The perspective we have developed building on the surveillance assemblage also highlights the opaqueness of MLB processes and the fragmented and diverse nature of MLB surveillance. It depicts the individual in the context of such surveillance as lacking orientation and comprehension of her situation. Transparency and explainability, thus, appear as consistent solutions to better the situation of the individuals, too. It provides the individual with the possibility to inform herself on how the decision systems operate, which impact her life.

### 3.  Inconsistencies and Limitations

While the concept of the surveillance assemblage allows for articulating the implicit understandings on which the transparency and rights-based approach rests, it also allows us to point out inconsistencies and limitations. The approach can be challenged on two grounds.

The first ground consists in exposing inconsistencies. Even if MLB decision-making was best conceived as a surveillance assemblage, the suggested responses (informing and granting rights to the individual) are insufficient. Informing and entitling the individual to rights can only be viewed as an adequate response to large scale corporate surveillance and MLB decision-making if one assumes that information and rights effectively empower the individual to protect and realize her interests. This assumption, however, can be challenged both on a general level as well as in the context of MLB decision-making when relying on the concept of the surveillance assemblage.

On a general level, the idea that the information and individual rights are adequate solutions to corporate power confronts larger objections regarding the neo-liberal construction of the individual as an optimizing, rational actor.[102] Formal rights and the theoretical possibility of taking advantage of these rights are equated with guaranteeing protection, although they have little or no practical effect on the situation of those whom they are supposed to empower.

The objections that are raised here against rights-based and individual-focused responses to MLB decision-making have been raised in a related context by Julie Cohen, criticizing the foundations of privacy theory. Cohen writes that privacy theory assumes the existence of a "liberal self" that is presumed to possess both "the capacity for rational deliberation and choice" and is "capable of exercising its capacities."[103] The liberal individualist model believes the individual to be an "autonomous, precultural island"[104]—and insofar as individuals are neither autonomous nor resemble an independent island, privacy strategies building on that presumption are fundamentally flawed. Cohen explains that "[i]n fact, the liberal self who is the subject of privacy theory and privacy policymaking does not exist."[105] It is this same "liberal self," however, that constitutes the basis of individual-focused and transparency and rights-based responses to algorithmic decision-making. The transparency and rights-based responses to automated decision-making build on the same assumptions on which policy responses such as the GDPR's inform and consent mechanism build—and which have been widely criticized and portrayed as ineffective in that context.[106]

Relying on the agency of the well-informed and entitled individual seems like a particularly misplaced response when accounting for the complexities of MLB decision-making. The perspective developed on the basis of the surveillance assemblage stressed that it was impossible to paint a consistent picture of MLB decision-making. MLB decision-making was described as a fragmented environment, in which the individual is entrenched in innumerable and ungraspable surveillance relationships. If the individual is incapable of comprehending the environment in which she is embedded, how can she seriously and at scale challenge corporate surveillance? The picture we drew of MLB decision-making, based on the surveillance assemblage, can be viewed as contradicting central claims on

---

[102] Waldman, *supra* note 36, at 624-29.

[103] Julie E. Cohen, *What Privacy Is For*, 126 HARV. L.R. 1904, 1907 (2013).

[104] *Id.* at 1906.

[105] *Id.* at 1905.

[106] On the limits of consent in times of Big Data, see Barocas & Nissenbaum, *supra* note 7.

which narratives of individual empowerment build. It suggests that the individual is incapable of comprehending and maneuvering through the chaos of MLB decision-making in her environment. Although the transparency and rights-based approach builds on an understanding of surveillance that can be best described as surveillance assemblage, the responses that are suggested are, at a closer look, inconsistent with this understanding.

The inadequacy of transparency and rights-based responses become more salient as machine learning-algorithms further complicate algorithmic systems. The information and rights-based approach only provides an adequate solution if individuals take advantage of their rights. As MLB decision-making situates individuals in an ever more complex assemblage of surveillance, and the workings of surveillance become more opaque and complex due to machine learning algorithms, it becomes more unlikely that individuals will be able to effectively take advantage of their rights.

The second ground on the basis of which the transparency and rights approach can be criticized consists in the understanding of surveillance on which it is based. Not only are the suggested solutions inconsistent with the premise of the surveillance assemblage, but the very premise is flawed, too.

The individualistic focus of the surveillance assemblage framework ignores overarching dynamics and neglects the broader and impactful societal effects of corporate surveillance and automated decision-making. Understanding corporate surveillance as an assemblage of billions of individual relationships without accounting for overarching effects and dynamics deriving from corporate surveillance and decision-making more generally ignores some of the most important effects of corporate surveillance. These overarching effects and dynamics are caused by the entirety of corporate surveillance practices and cannot be identified when surveillance relationships are viewed as individualistic phenomena.

The limitations of the transparency and rights approach ultimately derive from a limited understanding of the effects of corporate surveillance—which originate in the limitations of the concept of the surveillance assemblage. With its individualistic focus and its disregard for systematic effects of surveillance, the surveillance assemblage provides legal scholarship a poor basis to build on. The transparency and rights-based response reproduces the conceptual shortcomings of the surveillance assemblage.

To sum up, explainability and transparency, as well as individual rights such as non-discrimination, fairness, reasonableness, and due process, are important but insufficient responses to MLB decision-making. The increasing impact of machine learning algorithms prompts us to re-evaluate

the effectiveness of these responses, even if we rely on the concept of the surveillance assemblage to describe the effects of MLB decision-making.

### B.   Accounting for a Panoptic Perspective: Rethinking Legal Responses to MLB Decision-Making

The panoptic perspective on MLB decision-making can enhance legal scholarship by providing an alternative understanding of the effects of such decision-making. It provides not only theoretical starting points to criticize the concept of the surveillance assemblage on which much legal scholarship is implicitly based, but also offers a constructive starting point to assess the effects of MLB decision-making more comprehensively.

As the panoptic perspective alters our understanding of corporate surveillance, it also prompts us to develop new approaches to respond to it. The panoptic perspective suggests that legal scholarship on algorithms and decision-making must develop approaches to analyze broader dynamics and transformations that MLB decision-making provokes.

In this section, it is first argued why it is important to account for the larger dynamics and transformations caused by MLB decision-making when developing legal responses (Part IV.B.1, *infra*). Secondly, starting points are developed to analyze the normative implications of these larger transformations. It is argued that a debate on the legitimacy of MLB decision-making can help develop standards against which legal responses can be measured (Part IV.B.2, *infra*). Finally, this section discusses implications of the panoptic perspectives and revisits existing responses in the light of standards of legitimacy (Part IV.B.3, *infra*).

### 1.   A Panoptic Perspective as a Foundation for Legal Responses

On the basis of the panoptic framework, this article has developed a new perspective on the effects of MLB decision-making. This perspective has prompted us to reflect on how MLB decision-making transforms the environment in which individuals live. These transformations do not only concern the relationship between individuals and particular corporations, but the lives of individuals more generally.

Relying on the panoptic framework, it has been argued that MLB decision-making establishes a system of sanctions and rewards, which generates and enforces rules. MLB decision-making regulates human behavior by attaching predicted outcomes as immediate costs or awards. It submits human behavior to a correlative logic that is defined by the epistemological particularities of machine learning algorithms. It determines the opportunities individuals have and the way they are

perceived and treated by corporations. It defines how individuals need to behave to have access to important services.

The panoptic perspective suggests that to understand the dynamics and effects on our environments caused by surveillance, we must not only examine surveillance relationships on a granular level, but we have to be mindful of the larger system of which singular surveillance and decision-making mechanisms are a part and what the effects of this larger system are. The argument is that while corporate surveillance practices are diverse, they have shared effects. We have to account for those shared effects, too, to properly understand the implications of the MLB surveillance to which we try to respond.

If we look at singular snowflakes, we do not understand the impact of a snowstorm. It is only once snowflakes transform into a greater whole, into a thick layer of snow, that they have a significant effect—that they block highways, make houses tremble under their weight, transform our environment, and change our perception of the world.

MLB decision-making has similar effects that only become visible when viewed as a whole. Any legal assessment that is based on an individualistic understanding, and only looks at MLB decisions in their individuality, misses the most profound effects. Earlier, this article explored arguments that it is the logic underlying machine learning and centralization tendencies in the data market that connect all individual decisions to a greater whole with shared effects. [107] The existing focus on individuals, rights, and singular relationships between corporations and individuals views MLB decisions as individual snowflakes, and thus does not address some of their most impactful consequences.

An atomistic understanding of surveillance ignores the experience of being permanently subject to data collection and exposed to decisions that are based on a shared logic and the tremendous impact of that experience. While automated decisions could perhaps be conceived as singular snowflakes, MLB decisions must be conceived as part of a greater whole that transforms our environment and our perception thereof. Legal scholarship must account for the shared effects of MLB decision-making to develop adequate responses.

If we acknowledge that MLB decision-making has effects that are not limited to individual relationships between people and corporations, we have to develop new starting points to engage with its normative implications. The panoptic framework suggests that the main problem with MLB decision-making is not opacity or unfairness. The main problem is that

---

[107] *See supra* Part III.B.1.

it transforms the environment in which we live and allows corporations to exert power over us.

The panopticon was designed as a utilitarian model of perfect governance, of achieving compliance and obedience without having to rely on force. Analyzing MLB decision-making on the basis of panopticism suggests that it creates a system of governance—a system that influences the behavior of individuals by embedding it in an invisible net of sanctions and rewards.[108] If we accept that there is some truth, or plausibility, to this portrayal of MLB decision-making and what it might become in the future, the most pressing questions do not concern the individual relationships between people and specific corporations, but the systemic effects. Crucial questions then center on the identification of power relations. The panoptic perspective prompts us to ask questions such as: Who governs and who is governed? What are the rules enforced by the governance? And what notions can we rely on to challenge this system of governance?

The analysis we have developed on the basis of panopticism suggests some answers to these questions, and these answers must take a nuanced shape to account for the particularities of MLB decision-making. The governors are corporations, whose decision-making systems rule over the fate of individuals. The governed are individuals, who seek access to the fruits of society, including services, insurance, and housing. The rules that are enforced by the governors have been designed by machine learning algorithms, which again have been created to optimize the business of those who rely on them.

The system of governance established by corporate MLB decision-making is, however, different from conventional governance systems, such as the ones within states. Corporations do not possess overtly coercive powers and, when relying on MLB decision-making, aspire to optimize their own behavior rather than to shape the behavior of individuals. Unlike practices of targeted advertisement, MLB decision-making, by individualizing the conditions under which individuals can access services, does not have the explicit goal to manipulate and nudge individuals, but to increase profits by minimizing risks.

Furthermore, corporations, when using MLB decision-making, do not govern intentionally. However, from the point of view of the individual, it is not the intentions behind but the effects of MLB decision-making that matter. How individuals experience being subjected to these systems is what matters. This experience is shaped by sanctions and rewards, which

---

[108] The notion of governance has been widely used to engage with algorithmic power in the context of Big Data and free speech. Balkin, *Three Laws of Robotics*, *supra* note 2, at 1226-28; Balkin, *Free Speech in the Algorithmic Society*, *supra* note 2, at 1182-201; Klonick, *supra* note 35; Zittrain, *supra* note 35.

are distributed by the MLB decision-systems of corporations. Sanctions do not have to be coercive, as state sanctions typically are, to have a tremendous impact on individuals' lives and prompt them to modify their behavior. Again, it is not the modality but the impact of an action that is decisive from the point of view of an individual. The kind of governance that corporate MLB decision-making establishes is thus different from conventional state governance—but it is equally impactful.

As MLB decision-making establishes a system of sanctions and rewards—a certain kind of governance—legal scholars do not only have to wonder whether this system is fair (whatever that may mean), explainable and transparent. Rather, the central question becomes whether such a system and the power relations it constructs are *legitimate*.[109] To develop adequate regulatory responses to MLB decision-making, we need to operationalize the notion of legitimacy in a context in which it is not, at least in the constitutional and philosophical debate, usually used: the context of corporate power.

## 2. The Why and How of Legitimacy

The notion of legitimacy is becoming increasingly popular among scholars engaging with algorithmic systems.[110] Yet, the debate over *why* and *how* to apply standards of legitimacy to MLB systems is still in its infancy.

There is no one correct way in which the notion of legitimacy is to be employed with respect to corporate surveillance and MLB decision-making.[111] The notion of legitimacy is used in different ways and with different meanings across constitutional law and political philosophy—and it allows for a variety of starting points to engage with the problem of MLB decision-making.

---

[109] While the notion of legitimacy has no clearly defined meaning and can mean different things in different contexts, it has provided legal philosophers of different generations with starting points to tackle some of the most challenging questions of their time. For a history of thought regarding political legitimacy, see Fabienne Peter, *Political Legitimacy*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Summer 2017 ed.), available at https://perma.cc/N6JC-5B76.

[110] For first explanations on the notion of legitimacy in the context of algorithmic systems, see Waldman, *supra* note 36, at 614-15 (arguing that "algorithmic decision-making (is) an illegitimate source of authority in a liberal democracy"). PASQUALE, *supra* note 3, at 193 (describing how "the mater of legitimation was tabled"); Pasquale, *supra* note 5 (viewing transparency requirements as necessary to legitimize algorithmic authority); Berman, *supra* note 6 (engaging with the legitimacy of governmental uses of machine learning algorithms).

[111] For a related discussion of the notion of legitimacy, see Waldman, *supra* note 36, at 614 n.12.

Conventionally, the notion of legitimacy is used to engage with the question of whether state power is justified.[112] The notion of legitimacy is closely related to the notion of political freedom. The problem of political freedom has been poignantly summarized by Kelsen: "How is it possible to be subject to a social order and still be free?"[113] The "social order" that Kelsen and other constitutional lawyers and philosophers have been interested in is the state. By Kelsen's understanding, the state can be understood as "a social order by which individuals are bound to a certain behavior,"[114] and because this social order interferes with individual liberty, it requires legitimacy.

This article builds on Kelsen's definition of political freedom to develop a broader understanding of legitimacy—one that still centers on the idea that legitimacy engages with freedom in a social order, but one that does not restrict the kind of social order in which it plays a role to the state. Here, the notion of legitimacy is understood more broadly and taken to refer to the requirement to justify power in a social order and to reconcile it with the liberty of people.[115]

The analysis that we have developed on the basis of panopticism suggests that machine learning algorithms, too, create a system in which individuals are required to behave in certain ways; it creates a social order in which individuals suffer consequences for certain behavior and enjoy rewards for others. In the system of the state, laws constitute the rules on the basis of which individuals are judged. In the system of MLB decision-making, algorithms decide on how individuals are perceived and treated. If we acknowledge that both state and MLB decision-making impact the social order in which individuals live—and that the reason why standards of legitimacy apply to the state is because of its effects on the social order—it is no big leap to apply similar standards of legitimacy to MLB decision-making.[116]

The notion of legitimacy invites us to raise the same questions with which we assess the legitimacy of state power when assessing the systemic effects of MLB decision-making.[117] When state laws require individuals to behave in a certain way, we require these laws to satisfy standards of legitimacy. In liberal democracies, constitutions establish mechanisms to

---

[112] Citron and Pasquale have already drawn parallels between the power of the sovereign and algorithms. Citron & Pasquale, *supra* note 1, at 19. On the notion of legitimacy, see WOJCIECH SADURKSI ET AL., LEGITIMACY: THE STATE AND BEYOND (2019); Allen, *Political Legitimacy and Democracy*, 112 ETHICS 689 (2002); David Copp, *The Idea of a Legitimate State*, 28 PHIL. & PUB. AFFS. 3 (1999).

[113] HANS KELSEN, GENERAL THEORY OF LAW AND STATE 285 (2009).

[114] *Id.* at 284.

[115] Peter, *supra* note 109.

[116] A step also taken by Waldman, *supra* note 36.

[117] *See* Citron & Pasquale, s*upra* note 1, at 19.

legitimize laws and to reconcile individual freedom with the duty of individuals to comply with laws. They guarantee participation, individual rights and accountability, while limiting state power. It is only legitimate for the state to demand compliance with its laws when those laws have been created according to constitutional processes with the participation of the people and when those laws respect constitutional principles and fundamental rights in ultimate furtherance of the public interest. Critically, these laws can be changed if representatives of the people choose to do so. Ideally, in liberal democracies, the laws of the state increase social order while respecting and realizing the political self-determination of its citizens.

When corporate MLB decision-making creates systems of sanctions and rewards that incentivize and ultimately pressure individuals to behave in particular ways, these systems interfere with human freedom without similar safeguards. When individuals are submitted to rules created by MLB algorithms and self-police their lives in order to comply with standards of success in corporate surveillance systems, their lives are shaped by rules over which they have no influence. Individuals are not involved in the creation of these rules, and the rules do not pursue the wellbeing of citizens but the realization of company interests. There are no constitutional structures to reconcile the freedom of individuals with the power of corporations. If we measure MLB decision-making's influence against the same standards as we measure state power, it becomes clear that sophisticated mechanisms have been developed with regard to the latter but are largely absent with regard to the former.

Articulating the influence of MLB decision-making on individuals' opportunities and lives, penetrating its impact on behavior, acknowledging its coercive power, and generally describing it with language developed to describe the power of the state creates a perspective that adds to the already rich but still emerging critical literature on the power of algorithms.[118] It enables us to see that what is threatened by MLB decision-making is the capacity of an individual to lead her life freely, without having to comply with rules she has no influence over.

Take the example of the couple. Imagine that the state-imposed fines for individuals going to marriage counselling. There can be no doubt that this would interfere with the freedom and fundamental rights of that couple. Such an interference would have to be based on law, to assure democratic legitimacy and compliance with the standard of legality. However, even if there was such a law, a constitutional court would likely

---

[118] *See, e.g.*, JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM (2019); ZUBOFF, *supra* note 8.

strike it down, because it unreasonably interfered with the freedom of the couple. If a law cannot legally impose a fee for individuals going to marriage counselling, why then should MLB decision-systems?

Scholarship is beginning to acknowledge that we need to reconsider how corporate power is addressed—and whether the responses suggested thus far are appropriate.[119] The panoptic perspective on MLB decision-making can further critical legal scholarship that aspires to develop new remedies against new forms of corporate power. One goal of such scholarship is to understand how MLB decision-making changes our environment, creates a system of sanctions and rewards, affects the social order more generally, and limits freedom. The goal would be to find means to legitimize these effects—or, if legitimacy cannot be achieved, to minimize them by restricting the ways and fields in which MLB decision-making can be employed.

### 3.  Implications

A debate on the legitimacy of corporate MLB decision-making can provide standards against which to measure regulatory responses. Responses that mitigate the power of corporations and help maintain an environment in which individuals can freely decide how they want to act, without having to fear detrimental consequences by MLB systems, would contribute to the legitimacy of such systems. Explainability and transparency and rights-based responses, but also other standards such as data minimization and limited purpose principles, can be tested against the larger standard of legitimacy.

Individual rights, which empower individuals to take actions against corporations whose systems violate standards of non-discrimination, fairness, reasonableness, and due process, can contribute to limiting the power of corporations. We can now contextualize scholarship for rights against decision-making as being part of a greater endeavor that aims at legitimizing corporate power. In the context of the state, we see that fundamental rights, which protect individual freedom, contribute to the legitimacy of power. What the broad support for the establishment of rights against algorithms shows is that the idea that algorithmic power must be submitted to standards of legitimacy already enjoys great support—even if the endeavor to guarantee individual rights is usually not explicitly embedded in larger reflections on the matter of legitimacy.

However, as we have discussed earlier, a first set of limitations of the individual rights approach results from the limited practical ability of

---

[119] *See, e.g.*, Waldman, *supra* note 36.

individuals to take advantage of these rights. Individuals neither have the resources nor the expertise to defend their rights against each individual corporation they engage with.

A second set of limitations results from the limited reach of these rights. Decisions taken by corporations can be non-discriminatory, fair, and reasonable and still attach detrimental consequences to individual behavior in ways that limit the ability of individuals to freely maneuver through their environment. Furthermore, rights against automated-decision-making, which only apply to fully automated processes but not to processes in which decisions are suggested by machine learning algorithms and immediately adopted and executed by humans, can be easily circumvented by corporations.[120] The couple who receives higher interest rates because they went to marriage counselling has not been discriminated against and, if the increase in interest rates is based on accurate statistics, has not been treated unfairly or unreasonably. Still, the increase in interest rates contributes to transforming the environment of the couple in a panoptic setting, in which behavior is constantly rewarded or sanctioned. Also, it makes no difference to the couple if the final decision is made by a human on the basis of information generated by machine learning algorithms or by a machine directly.

Thirdly, the individual rights approach does not effectively respond to larger transformations that are not caused by individual decisions but by the fact that MLB decision-making transforms our environment more generally. Entitling the individual with rights she may attempt to enforce against individual corporations does little or nothing to change the environment through which a person navigates more generally. The greatest threats of MLB decision-making, from the panoptic perspective, are not singular violations of rights, but structural transformations affecting the environment in which we live and dictating how we have to behave to succeed. Allowing the couple to challenge the decision of the credit card company to increase their interest rates does not alter the fact that innumerable corporations are still making decisions over them. There is no one single player—as it is the case in the context of state surveillance—whose actions the couple could challenge to significantly improve their situation and free themselves from permanently being subject to corporate decisions. There is no single gaze that they could try to evade, and no single guard whose sanctioning power they could challenge.

Thus, providing individuals with individual rights to respond to the way their environment transforms because of MLB decision-making is like only

---

[120] Wachter et al., *Right to Explanation*, *supra* note 5, at 92.

enabling them to fight off individual snowflakes when they find themselves in a snowstorm. Individual rights constitute an important, though limited, means to legitimize the power of MLB decision-making.

From a panoptic perspective, responses such as explainability and transparency, widely praised as important and effective responses to MLB data analysis and decision-making, also do not address the core of the problem. Explainability and transparency contribute to a better understanding by individuals of the rules by which they are judged and thus enable them to internalize these rules more easily.[121] The internalization of rules, however, is not a sign of emancipation, but of submission to foreign standards. As the context in which panopticism has been originally developed—the prison—illustrates, establishing rules over people and subsequently surveilling them is a tool of suppression. Creating conditions in which individuals internalize rules is a way to govern them, not to set them free.[122]

Furthermore, providing individuals with the information necessary to adapt to a system of decision-making does not empower them; it only facilitates compliance with rules—rules that are shaped by the economic rationale that underlies corporate MLB decision-making. The problem of the couple is not solved when someone explains to them that it was their visits to marriage counselling that led to the increase in interest rates. To the contrary, it might discourage them from seeking the help that they require, which in turn can cause personal distress and ultimately the financial hardship predicted in the first place.

Despite their shortcoming, there is no doubt that transparency and explainability have their virtues. They are especially important with regard to manipulative practices, in which individuals are nudged to adopt certain behaviors without their knowledge. Explaining to individuals how psychological profiles and targeted advertisement can be used to manipulate them empowers individuals to resist such manipulation. In the context of MLB decision-making, however, explainability and transparency as such do not empower the individual—they merely expose the rules to which they are submitted, thus contributing to the emergence of a panoptic setting.

Besides the popular approaches of individual rights and transparency, regulators and scholars have engaged with a third set of approaches, which, when measured against standards of legitimacy in a panoptic setting, are more promising. These approaches consist in establishing general principles that data processors, or those developing or relying on MLB systems, have

---

[121] A related critique is developed by Waldman, *supra* note 36, at 622-31.
[122] For a variety of further objections against transparency as a solution, see Ananny & Crawford, *supra* note 26.

to comply with. Some important principles are explicitly laid down in laws, most prominently in the General Data Protection Regulation (GDPR) of the European Union. [123] These principles include transparency, consent, accuracy, relevance, purpose limitation, and data minimization.[124] In an attempt to shape the global landscape of regulation of MLB systems, the European Commission has also recently published a draft of an "Artificial Intelligence Act" [125] that prohibits public social scoring entirely [126] and classifies some forms of scoring practices by private parties as "high risk,"[127] which entails certain compliance requirements. [128]

This article does not aspire to analyze the concrete content of the GDPR's general principles or the Commission's draft of the "Artificial Intelligence Act." Rather, it is interested in the idea of principles more generally and the potential they have to provide effective responses to MLB decision-making. Looking at the GDPR, however, allows critical scholars to engage with the most influential manifestations of the principles and can further overarching reflections.

Upon examination, some of the general principles do little to respond to the harmful implications of MLB decision-making exposed by the panoptic perspective. The limitations of consent and transparency have been discussed earlier. Nor do standards of relevancy and accuracy protect interests like those of the couple going to marriage counselling. If they were actually at the counselling, the data that has been processed was accurate. If marriage counselling correlates with higher risks of financial insolvency, the data was relevant to determine their credit rate. The decision regarding their interest rate was not only non-discriminatory, but also based on relevant data and an accurate assessment of that data. Thus, the principles of relevancy and accuracy do not protect the couple's interest in seeking help when they require it without having to fear detrimental consequences.

Other principles, such as data minimization and limited purposes principles, can, however, contribute to improving the situation of the

---

[123] For an introduction to the GDPR, see Christina Tikkinen-Piri et al., *EU General Data Protection Regulation: Changes and Implications for Personal Data Collecting Companies*, 34 COMPUT. L. & SEC. REV. (2018). For a comparison between the GDPR and the CCPA, see Anupam Chander et al., *Catalyzing Privacy Law*, 105 MINN. L. REV. 1733, (2021).

[124] Rubinstein, *supra* note 7 (engaging with principles laid down in the predecessor of the GDPR, the European Data Protection Directive, from a perspective of Big Data challenges).

[125] *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artifical Intelligence (Artifical Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206 final (Apr. 21, 2021).

[126] *Id.* art. 5 § 1(c).

[127] *Id.* arts. 6-7.

[128] *Id.* arts. 8-15.

couple and point towards ways to respond to MLB decision-making. The GDPR's data minimization principle requires corporations to process personal data only where it "is necessary in relation to the purposes for which they are processed."[129] When enforced strictly, data minimization that limits the availability of data about individuals can prevent individuals from being—and feeling—observed in all spheres of life, and it can limit the data available for MLB decision-making. If individuals are not permanently and excessively observed, they do not have to second guess their behavior and wonder whether it will provoke detrimental or favorable MLB decisions. The environment of individuals can thus not be personalized as easily and prevents their environment from turning into a panoptic setting.

The GDPR's limited purpose principle requires corporations to collect data only for "specified, explicit and legitimate purposes" and not to process data in ways that are incompatible with those purposes.[130] What derives from this principle, and how it interplays with the principle of consent, is subject to debate. The important question for the purpose of this article is not what the concrete content of the GDPR's principle is, but what can be derived from a limited purpose principle to effectively respond to MLB decision-making.

What shape would a limited purpose principle, which could be established in any jurisdiction that chooses to do so, have to adopt to limit the discipling effects of MLB decision-making? A limited purposes principle could require that information collected for one purpose not be used to infer information unrelated to that purpose. More importantly, it could require that such information—and inferences on the basis of this information—not be used to personalize the way an individual is treated in unrelated fields of life. Such a limited purposes principle would protect a couple going to marriage counselling. It would prevent a navigation app provider from identifying what people the couple meets with and whether they go to marriage counselling, and it would prevent it from using or selling that information to adjust the credit ratings of the couple. Such a limited purpose principle could apply generally, to all data processing, and would not have to rely on individuals attempting to enforce their rights. It could prevent MLB decision-making from creating a ubiquitous system of rewards and sanctions, which requires individuals to behave in certain ways if they do not want to suffer detrimental consequences—and thus significantly contributes to limit harmful effects of MLB decision-making.

---

[129] Council Regulation 2016/679 of Apr. 27, 2016, General Data Protection Regulation, art. 5 § 1(c), 2016 O.J. (L 119) 1, 35.
[130] § 1(b).

While we can design general principles that mitigate the harms of MLB decision-making effectively, such principles bring with them major challenges. Some challenges are more technical and relate to questions of how to define them concretely, how to reconcile them with other principles, and how to enforce them. Such rather technical challenges, however, are not the most significant. The most significant challenges consist of objections that reject such principles for economic reasons and are backed by constitutional arguments.

These objections are backed by an ideology that advocates for unconstrained economic freedom and the unrestricted economic exploitation of new technologies. Principles such as data minimization and purpose limitation fundamentally limit how corporations can extract value from data. They significantly hamper the growth of what has become one of the largest sectors of the economy. These principles entail severe constraints on the business models of corporations and can conflict with the constitutionally protected interests of corporations. The debate on establishing such principles inevitably leads to a larger ideological and constitutional conflict on how to reconcile the freedom of individuals with the freedom of corporations. The panoptic perspective and the suggested notion of legitimacy are so important when engaging with MLB decision-making because they demonstrate how legal responses are embedded in this larger conflict and provide foundations to resolve it.

The contribution that the panoptic perspective and the notion of legitimacy can make in the debate on the regulation of MLB decision-making is to establish a background necessary to adequately balance the interests at stake. In particular, the panoptic perspective demonstrates that limitations of corporate freedoms can be justified—in fact, they are necessary to prevent the illegitimate demise of human freedom and self-determination.

While it is fairly obvious that principles governing the processing and use of data processing constrain the freedom of corporations, it is not as obvious what may justify these constraints. The panoptic perspective shows that MLB decision-making transforms the environment in which we live and defines the opportunities available to individuals. It attaches significant costs to behavior and establishes rules on the basis of which individuals are sanctioned or rewarded. The panoptic perspective demonstrates that MLB decision-making can create dynamics that require individuals to self-police their lives, that it creates disciplining effects, and that it substantially expands corporate power. It prompts us to carefully consider whether MLB decision-making, as it becomes more popular, unacceptably interferes with individual freedom.

Finally, as the dynamics caused by MLB decision-making become more graspable, we have to consider recalibrating the balance between the freedoms of corporations and the freedoms of individuals. To maintain conditions of self-determination and freedom, which are central to the endeavor of liberal democracies and aspired to by the notion of legitimacy, we have to assess whether the interest of individuals not to be subjected to MLB decision-making outweighs the interests of corporations to rely on it. Panoptic and legitimacy focused perspectives on MLB decision-making can present the debate on how to balance corporate freedom with individual freedom with depth, and arguments in favor of protecting individual freedom with additional weight.

## V.    CONCLUSION

The adequacy of legal responses to MLB decision-making will depend on how we conceive of its effects on individuals and society. The way we conceive of the effects of MLB decision-making is, in turn, shaped by the concepts we choose to describe them. The question of how to conceptualize the effects of MLB decision-making should thus be at the center rather than at the periphery of legal scholarship. This article has provided an account of how surveillance theory can provide different perspectives on MLB decision-making and how it can thus advance legal scholarship.

There is still great uncertainty regarding the question of what concrete effects MLB decision-making is going to trigger. The point this article has made is not that there is one correct way to analyze the effects of MLB decision-making, but that that there are different, even contradictory ways to analyze the effects of MLB decision-making, and that each way suggests different conclusions for the legal debate. The article has shown that what concepts we rely on to describe the effects of MLB decision-making matters and that we should have a discussion on the merits and flaws of relying on different concepts.

Each perspective on MLB decision-making has its merits and can contribute to a better understanding of its implications. The perspective that is based on the concept of the surveillance assemblage conceives the problem of surveillance on a granular level. It emphasizes the opaqueness of MLB decision-making and the fragmented and diverse nature of corporate surveillance. The panoptic perspective, conversely, allows us to reflect on how MLB decision-making transforms the environment in which individuals live more generally, and how it might affect their intuition and behavior. MLB decision-making, from this perspective, establishes a system of sanctions and rewards, which is shaped by the correlative logic of

machine learning algorithms and incentivizes individuals to behave in certain ways. This perspective enables us to leave a granular, individualistic focus behind and to engage with larger transformations caused by corporate MLB decision-making.

While the surveillance assemblage might be the more suitable framework to describe automated decision-making practices at present and in the near future, the panoptic framework might be the more suitable framework to assess the effects of MLB decision-making at a later stage. The panoptic perspective exposes trends that are likely to become stronger over time, and which ultimately might alter the mainstream understanding of corporate surveillance and the problems with which MLB decision-making confronts legal scholars. Scholars have already identified certain behavioral dynamics, under notions such as "normalization (or regimentation)," and a panoptic perspective allows for substantiation of such observations. Even though we might not be able to empirically assess these dynamics yet, considering the fast development of the technological progress in the field of data processing, we can begin to theorize them— and to explore what adequate legal responses would be.

Both perspectives developed in this article have served as a foundation to revisit the legal debate on MLB decision-making. It has been argued that the surveillance assemblage constitutes the implicit understanding that underlies popular responses to MLB decision-making, including individual rights and transparency. While these responses are important, they are not sufficient—even when judged on the premises on which they are based. The opacity of MLB decision-making and the complexity of corporate surveillance environments make it unrealistic for individuals to comprehend and challenge decisions that are made over them. Transparency and individual rights are unlikely to efficiently empower the individual to defend her interests and thus have to be supplemented by other kinds of protections.

On the basis of the panoptic perspective, it has been argued that one of the most important harms of MLB decision-making is the emergence of a novel system of sanctions and rewards that stretches over all spheres of life. Legal scholarship on MLB decision-making must respond to this harm. MLB decision-making drastically increases the power of corporations to the detriment of individual freedom. To effectively respond to MLB decision-making, legal scholarship must not only engage with surveillance relationships on a granular level, but it must account for the larger system of which singular relationships are only a part. Legal scholars should not only be concerned with questions of whether individual decisions are fair, non-discriminatory, and explainable. Rather, they must engage with the

question of whether the impact of MLB decision-making on the lives of individuals is legitimate. The notion of legitimacy provides starting points to assess the adequacy of legal responses, the purpose of which must be to restore the balance of power between individuals and corporations. It provides depth for the debate on how to reconcile the freedoms of corporations with the freedom of individuals—and can thus contribute to resolving one of the great challenges with which steadily improving data processing capabilities confront liberal democracies.