



Texas A&M University School of Law
Texas A&M Law Scholarship

Faculty Scholarship


2020

Artificial Financial Intelligence

William Magnuson

Texas A&M University School of Law, magnuson@law.tamu.edu

Follow this and additional works at: <https://scholarship.law.tamu.edu/facscholar>

 Part of the [Civil Rights and Discrimination Commons](#), [Computer Law Commons](#), [Internet Law Commons](#), and the [Law and Society Commons](#)

Recommended Citation

William Magnuson, *Artificial Financial Intelligence*, 10 Harv. Bus. L. Rev. 337 (2020).

Available at: <https://scholarship.law.tamu.edu/facscholar/1435>

This Article is brought to you for free and open access by Texas A&M Law Scholarship. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Texas A&M Law Scholarship. For more information, please contact aretteen@law.tamu.edu.

ARTIFICIAL FINANCIAL INTELLIGENCE

WILLIAM MAGNUSON*

Recent advances in the field of artificial intelligence have revived long-standing debates about the interaction between humans and technology. These debates have tended to center around the ability of computers to exceed the capacities and understandings of human decisionmakers, and the resulting effects on the future of labor, inequality, and society more generally. These questions have found particular resonance in finance, where computers already play a dominant role. High-frequency traders, quantitative (or "quant") hedge funds, and robo-advisors all represent, to a greater or lesser degree, real-world instantiations of the impact that artificial intelligence is having on the field. This Article, however, takes a somewhat contrarian position. It argues that the primary danger of artificial intelligence in finance is not so much that it will surpass human intelligence, but rather that it will exacerbate human error. It will do so in three ways. First, because current artificial intelligence techniques rely heavily on identifying patterns in historical data, use of these techniques will tend to lead to results that perpetuate the status quo (a status quo that exhibits all the features and failings of the external market). Second, because some of the most "accurate" artificial intelligence strategies are the least transparent or explainable ones, decisionmakers may well give more weight to the results of these algorithms than they are due. Finally, because much of the financial industry depends not just on predicting what will happen in the world, but also on predicting what other people will predict will happen in the world, it is likely that small errors in applying artificial intelligence (either in data, programming, or execution) will have outsized effects on markets. This is not to say that artificial intelligence has no place in the financial industry, or even that it is bad for the industry. It clearly is here to stay, and, what is more, has much to offer in terms of efficiency, speed, and cost. But as governments and regulators begin to take stock of the technology, it is worthwhile to consider artificial intelligence's real-world limitations.

INTRODUCTION	338
I. ARTIFICIAL INTELLIGENCE IN THE FINANCIAL INDUSTRY	341
A. <i>The Technology of Artificial Intelligence</i>	342
B. <i>Artificial Intelligence Strategies in Finance</i>	348
C. <i>Existing Literature</i>	352
II. ARTIFICIAL FINANCIAL INTELLIGENCE	355
A. <i>Data Dependency</i>	355
B. <i>Overweighting</i>	359
C. <i>Echo Effects</i>	363
III. REGULATING ARTIFICIAL FINANCIAL INTELLIGENCE	365
A. <i>Artificial Fairness</i>	366
B. <i>Artificial Efficiency</i>	368
C. <i>Artificial Risk</i>	371
D. <i>The Role of Self-Regulation</i>	373

* Associate Professor, Texas A&M University School of Law; J.D., Harvard Law School; M.A., Università di Padova; A.B., Princeton University. The author wishes to thank Jack Goldsmith, Todd Henderson, Chris Brummer, Tom Vartanian, Bob Ledig, Sarah Jane Hughes, Angela Walch, Tom Lin, and Rory Van Loo for their suggestions and advice.

IV.	OBJECTIONS	375
	A. <i>The Mirror Critique</i>	375
	B. <i>The Precautionary Critique</i>	378
	C. <i>The Balance Critique</i>	380
	CONCLUSION	381

INTRODUCTION

In the novel *Tell the Machine Goodnight*, a super-powered computer called the Apricity (named after the now-defunct English word for “the feeling of sun on one’s skin in the winter”) uses advanced artificial intelligence to deliver lifestyle recommendations to users.¹ A simple cotton swab applied to the mouth, and then swiped across the computer’s reader, is all the computer needs to formulate a list of recommended changes for users to make in their lives. These changes, if implemented, promise to lead to happier, more fulfilled lives, or, as Apricity’s corporate legal department advises them to say, to improvements in one’s “overall life satisfaction.” The story opens with one unfortunate user receiving his own personalized set of recommendations from the Apricity:

The machine said the man should eat tangerines. It listed two other recommendations as well, so three in total. A modest number, Pearl assured the man as she read out the list that had appeared on the screen before her: one, he should eat tangerines on a regular basis; two, he should work at a desk that received morning light; three, he should amputate the uppermost section of his right index finger.²

Pearl, the technician tasked with operating the device, watches the man raise his right hand before his face and wonders if he is going to cry. She assures him that his recommendation is “modest” compared to others she has seen and reminds him that the system boasts a 99.97% approval rating. “The proof is borne out in the numbers,” she concludes.³

Writing about artificial intelligence requires a certain dose of imagination. This has always been the case. It is not just because artificial intelligence draws more than its fair share of attention in science fiction and fantasy circles, although this certainly plays a part. It is also, and perhaps more importantly, because the term *artificial intelligence* conjures up images of a world that does not currently exist. When most people think of artificial intelligence, they think of superpowered computers that act and think like human beings, with complicated motives, wide-ranging capacities, and often dangerous tendencies. This combination of qualities simply does not exist in

¹ KATIE WILLIAMS, *TELL THE MACHINE GOODNIGHT* (1st ed. 2018).

² *Id.* at 1.

³ *Id.* at 2.

the state of today's artificial intelligence research. As more than one artificial intelligence researcher has remarked, most of the people making the *doomsday* prognostications of artificial intelligence do not actually work in the field. Those that do have significantly more modest forecasts about the potential of artificial intelligence. This is not to say that they believe artificial intelligence is an ineffective technology. It is just that they understand its limitations.

At the same time, artificial intelligence algorithms have undeniably made significant advances in recent years. Improved algorithms, bigger data sets, and more powerful computers have combined to give machine learning strategies ever-growing capabilities. In the last decade alone, we have witnessed machine learning-based computers beat the world's best Jeopardy contestants,⁴ accurately identify breast cancer from pathology slides,⁵ and drive cars around the streets of Phoenix, Arizona.⁶ Governments around the world are researching autonomous weapons systems that could potentially mark a major shift in the nature of warfare and national security.⁷

Financial institutions, intrigued by the potential of machine learning, have begun to explore how to incorporate artificial intelligence into their own businesses. In recent years, large investment banks have hired away talented experts from academia and Silicon Valley to head up machine learning divisions.⁸ Fintech startups have created credit rating models and fraud detection algorithms based on artificial intelligence strategies.⁹ And high-frequency traders, quantitative hedge funds, and robo-advisors are implementing artificial intelligence into their businesses as well.¹⁰ "Artificial financial intelligence" will likely take on increasing importance in years to come.

But how should our legal structures respond to the rise of artificial financial intelligence? Should we attempt to encourage it, in the hopes of nur-

⁴ STEPHEN BAKER, *FINAL JEOPARDY: THE STORY OF WATSON, THE COMPUTER THAT WILL TRANSFORM OUR WORLD* 251 (1st ed. 2012).

⁵ See Jessica Kent, *Google Deep Learning Tool 99% Accurate at Breast Cancer Detection*, HEALTH IT ANALYTICS, (Oct. 22, 2018), <https://healthitanalytics.com/news/google-deep-learning-tool-99-accurate-at-breast-cancer-detection>.

⁶ See Tim Higgins & Jack Nicas, *Waymo's Self-Driving Milepost: Humans Take a Backseat*, WALL ST. J., (Nov. 7, 2017), <https://www.wsj.com/articles/waymos-self-driving-milepost-humans-take-a-backseat-1510070401>.

⁷ See generally Rebecca Crootoff, *War Torts: Accountability for Autonomous Weapons*, 164 U. PA. L. REV. 1347 (2016); Rebecca Crootoff, *Autonomous Weapons Systems and the Limits of Analogy*, 9 HARV. NAT'L SEC. J. 51 (2018); GREG ALLEN & TANIEL CHAN, BELFER CTR. FOR SCI. & INT'L AFFAIRS, *ARTIFICIAL INTELLIGENCE AND NATIONAL SECURITY* (2017).

⁸ See Sarah Butcher, *The Top Machine Learning Teams in Investment Banks*, eFINANCIAL-CAREERS, (May 23, 2018), <https://news.efinancialcareers.com/us-en/315969/top-machine-learning-teams-banks>.

⁹ See Rory Van Loo, *Making Innovation More Competitive: The Case of Fintech*, 65 UCLA L. REV. 232, 240 (2018).

¹⁰ See *Machine-Learning Promises to Shake Up Large Swathes of Finance*, THE ECONOMIST (May 25, 2017), <https://www.economist.com/finance-and-economics/2017/05/25/machine-learning-promises-to-shake-up-large-swathes-of-finance>.

turing a potentially transformative technology? Or should we try to restrict it, out of concerns about its broader risks for society? Is it possible to do both? In order to answer these questions, we must first identify the problems that artificial financial intelligence poses, and the sources from which these problems spring. Many scholars argue that artificial intelligence poses existential questions about the nature of discrimination, the future of work, and the repercussions of super-human intelligence.¹¹ This Article, however, will take a somewhat contrarian position. It will argue that the primary danger of artificial financial intelligence is not so much that it will surpass human intelligence, but rather that it will exacerbate human error.¹²

Artificial financial intelligence can magnify the effects of human error in three ways. First, because artificial intelligence techniques, and in particular machine learning algorithms, rely heavily on identifying patterns in historical data, use of these techniques will tend to lead to results that perpetuate the status quo in the data. Thus, if the data used to train artificial intelligence algorithms is flawed, either through poor selection methods or unavoidable problems in the external market itself, the resulting outputs of artificial intelligence will reflect (and, indeed, strengthen) those flaws. Second, because many of the most powerful artificial intelligence strategies are the least transparent or explainable, decisionmakers within financial institutions may give more weight to the results of the algorithms than they are due. It is a well-known problem in machine learning, and in particular in the sub-fields of deep learning and neural networks, that the complexity of the methods used to reach a given result makes identifying a particular *reason*,

¹¹ See Tom C. W. Lin, *The New Market Manipulation*, 66 EMORY L. J. 1253, 1254 (2017) (arguing that artificial intelligence risks widespread market manipulation); Rory Van Loo, *Digital Market Perfection*, 117 MICH. L. REV. 815 (2019) (arguing that AI-based automated assistants “could make some large markets more volatile, raising unemployment costs or financial stability concerns as more firms fail”); Rory Van Loo, *Technology Regulation by Default: Platforms, Privacy, and the CFPB*, 2 GEO. L. TECH. REV. 531, 544-45 (2018) (arguing that the Consumer Financial Protection Bureau should exert additional authority to inspect financial algorithms); Chris Brummer & Yesha Yadav, *Fintech and the Innovation Trilemma*, 107 GEO. L. J. 235, 275 (2019) (arguing that “[g]iven the potential for [AI and machine learning] to result in widespread, cascading costs, mapping the likely performance of sophisticated algorithms becomes especially necessary”); Gregory Scopino, *Preparing Financial Regulation for the Second Machine Age: The Need for Oversight of Digital Intermediaries in the Futures Markets*, 2015 COLUM. BUS. L. REV. 439, 440 (2015) (“[H]umans who are operating as futures market intermediaries . . . are likely to be displaced by digital intermediaries, that is, artificial agents that perform critical roles related to enabling customers to access the futures and derivatives markets.”); Luca Enriques & Dirk A. Zetzsche, *Corporate Technologies and the Tech Nirvana Fallacy*, Eur. Corp. Governance Inst. Working Paper No. 457, 2019) (analyzing the ways in which modern algorithms exacerbate agency problems in corporate law).

¹² Other scholars have made related points in other contexts, such as discrimination. See Solon Barocas & Andrew D. Selbst, *Big Data’s Disparate Impact*, 104 CAL. L. REV. 671, 671 (2016) (“[A]n algorithm is only as good as the data it works with. Data is frequently imperfect in ways that allow these algorithms to inherit the prejudices of prior decision makers.”); More recently, financial regulation scholars have turned to this problem as well. See Tom C.W. Lin, *Artificial Intelligence, Finance, and the Law*, 99 FORDHAM L. REV. 531, 531 (2019) (examining the “perils and pitfalls of artificial codes, data bias, virtual threats, and systemic risks relating to financial artificial intelligence.”).

or even several reasons, for the result difficult to do. Combined with cognitive decision biases related to availability and herding effects, we can expect that artificial financial intelligence will be hard to resist in financial decisionmaking processes. Finally, because much of the financial industry depends not just on predicting what will happen in the world, but also on predicting what other people will predict will happen in the world, it is likely that small errors in artificial intelligence (either in data, programming, or execution) will have outsized effects on markets. The artificial intelligence strategies of one firm may interact unpredictably with the artificial intelligence strategies of other firms.¹³ The resulting echo effects could harm both the efficiency of markets and the stability of financial systems.¹⁴

What does this mean for financial regulation? In a way, it is a cause for optimism. Financial regulators are well-placed to deal with artificial financial intelligence because they have a wide array of laws and regulations covering the relevant behaviors—ensuring fairness, promoting efficiency, and protecting stability. While accomplishing these goals will not be easy, the general categories of problems are ones to which financial regulators are accustomed. At the same time, regulating artificial financial intelligence may require slightly different regulatory mechanisms, or at least targets. Understanding and monitoring artificial financial intelligence may, and indeed likely will, require specialized expertise. It may also call for more intrusive and hands-on inspections of sensitive business data. We may not need better laws, but we do need better information.

I. ARTIFICIAL INTELLIGENCE IN THE FINANCIAL INDUSTRY

The computer science field of artificial intelligence has exploded in recent years. New advances in technical approaches, computing power, and the availability of data have led to notable successes in a variety of applications, from image recognition¹⁵ to natural language processing¹⁶ to medical diagnoses¹⁷ to game strategy.¹⁸ These advances have brought renewed atten-

¹³ See Rory Van Loo, *The Rise of the Digital Regulator*, 66 DUKE L.J. 1267, 1294 (2017) (discussing the unpredictable ways that different firms' algorithmic pricing tools may interact).

¹⁴ See Hilary J. Allen, *Driverless Finance*, 10 HARV. BUS. L. REV. 158, 158 (2019) (arguing that financial regulators should adopt a precautionary approach to regulating financial algorithms due to their potential to create systemic risk for the broader financial system).

¹⁵ See Parry Olson, *Image-Recognition Technology May Not Be as Secure as We Think*, WALL ST. J., (June 4, 2019), <https://www.wsj.com/articles/image-recognition-technology-may-not-be-as-secure-as-we-think-11559700300>.

¹⁶ See Tom Young, et al., *Recent Trends in Deep Learning Based Natural Language Processing*, ARXIV 1708.02709 (2018), <https://arxiv.org/pdf/1708.02709.pdf>.

¹⁷ See A. Michael Froomkin, et al., *When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning*, 61 ARIZ. L. REV. 33 (2019).

¹⁸ See Nick Statt, *How Artificial Intelligence Will Revolutionize the Way Video Games Are Developed and Played*, THE VERGE (Mar. 6, 2019), <https://www.theverge.com/2019/3/6/18222203/video-game-ai-future-procedural-generation-deep-learning>; George Anadiotis, *The State of AI in 2019: Breakthroughs in Machine Learning, Natural Language Processing, Games, and Knowledge Graphs*, ZDNET (July 8, 2019), <https://www.zdnet.com/article/the->

tion to artificial intelligence in the business world and, in particular, in the financial industry. Financial firms, who have always relied heavily on statistics and quantitative analysis in their services, are a natural fit for artificial intelligence strategies. Artificial intelligence, after all, attempts to identify patterns and probabilities from historical data in order to improve results, such as predicting future statistics or identifying complicated associations. Increasingly, financial firms, such as quant hedge funds, high-frequency traders, and robo-advisors, are turning to artificial intelligence to improve their own results. This new *artificial financial intelligence* ecosystem has, in turn, generated its fair share of hand-wringing in policy and academic circles, where scholars have worried about cyber-systems that could potentially exceed human capabilities. This Part will describe the rise of artificial intelligence, its use in the financial industry, and the existing literature on how policymakers should respond to it.

A. *The Technology of Artificial Intelligence*

What is artificial intelligence? The term has many meanings, a problem derived from its long history in computer science and popular fiction.¹⁹ The term itself is generally believed to have originated in 1956, when a Dartmouth mathematics professor, John McCarthy, used it in connection with a conference he was organizing on the topic of *thinking machines*.²⁰ But the general outlines of the idea go back much further. Edgar Allan Poe wrote an essay in 1836 about a mechanical chess player called The Turk that was purportedly better than any human.²¹ Alan Turing, at least as early as 1950, was directly tackling the question of whether computers could think.²² Turing famously concluded that the question itself was ill-formed and instead needed to be reconceptualized as whether machines could act in such a way that they would be indistinguishable from humans.²³ In order to determine this, he developed a test, now called the Turing test, in which a human, corresponding through text with two players, one of whom was a human and one of whom was a computer, attempted to determine which one was human.²⁴ If he could not, then, under the Turing test, the computer could be said to be intelligent.

state-of-ai-in-2019-breakthroughs-in-machine-learning-natural-language-processing-games-and-knowledge-graphs/.

¹⁹ See NAT'L SCI. & TECH. COUNCIL, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE 5 (2016).

²⁰ See NILS J. NILSSON, THE QUEST FOR ARTIFICIAL INTELLIGENCE 52–56 (2009).

²¹ See EDGAR ALLAN POE, *Maelzel's Chess-Player*, in Edgar Allan Poe: Complete Tales and Poems 373 (2003).

²² Alan M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433, 433 (1950).

²³ *Id.* at 433–34.

²⁴ It turns out that the Turing test has generated its fair share of controversy, both with respect to whether individual computers have in fact passed it and with respect to whether the test itself is a good indicator of true intelligence on the part of machines.

More modern definitions of artificial intelligence tend to be somewhat more inclusive. Berkeley computer scientist Stuart Russell defines artificial intelligence as:

the study of methods for making computers behave intelligently. Roughly speaking, a computer is intelligent to the extent that it does the right thing rather than the wrong thing. The right thing is whatever action is most likely to achieve the goal, or, in more technical terms, the action that maximizes expected utility.²⁵

As an even cursory perusal of this selection will reveal, this definition runs into its own set of problems. For one, we might say that a calculator does the right thing when it says that $1 + 1$ is 2, but most observers would not say that it is behaving *intelligently* in any ordinary sense of the term. For another, once we define intelligence to include an element of morality (does a computer's action maximize expected utility?), we open the door to contentious, unanswerable questions that distract from the technology itself. To borrow from Socrates, if a user asks a computer for a sword, and the computer receives the question, finds a sword for him in the real world, purchases it, and then delivers it to him, most observers would conclude that the computer is behaving intelligently.²⁶ But if that same user requesting the sword happens to be insane and plans to use it to harm himself or others, then it would (normally) be wrong to give him the sword. Does this mean that the computer is now *not* displaying artificial intelligence, because it is doing the *wrong* thing, ethically speaking? These types of questions could be raised for any conceivable use of a computer, and they are not closely connected to the problems unique to artificial intelligence itself.

On the other side of the spectrum, narrower definitions of artificial intelligence tend to circumscribe the range of the field unnecessarily. Many artificial intelligence researchers use the term *artificial intelligence* to refer to a particular set of techniques or algorithms that computers can adopt to achieve better results. So, in the 1980s, artificial intelligence became closely associated with an approach to problem-solving known as "expert system" design.²⁷ Under expert systems, computers would be given a long set of if-then rules that they could then apply to problems in a given field.²⁸ The logic

²⁵ STUART RUSSELL, Q&A: THE FUTURE OF ARTIFICIAL INTELLIGENCE, UNIV. OF CAL. BERKELEY, <http://people.eecs.berkeley.edu/~russell/temp/q-and-a.html>.

²⁶ The discussion of the madman and the sword takes place in *The Republic*, Book I. Appropriately for this article, it involves the question of whether justice requires you to pay your debts. See PLATO, REPUBLIC 6 (C. D. C. Reeve ed., G.M.A. Grube trans., Hackett Pub. Company, Inc. 2d ed. 1992) (c. 375 B.C.E.) ("Everyone would surely agree that if a sane man lends weapons to a friend and then asks for them back when he is out of his mind, the friend shouldn't return them, and wouldn't be acting justly if he did.").

²⁷ See generally PAUL HARMON & DAVID KING, EXPERT SYSTEMS: ARTIFICIAL INTELLIGENCE IN BUSINESS (1985); Bruce G. Buchanan & Reid G. Smith, *Fundamentals of Expert Systems*, in 4 HANDBOOK OF ARTIFICIAL INTELLIGENCE 149 (Avron Barr, et al. eds. 1981-89).

²⁸ See ROBERT A. EDMUNDS, THE PRENTICE HALL GUIDE TO EXPERT SYSTEMS 28-31 (1988).

of expert systems was that humans who are specialists in a field tend to behave according to a number of heuristics, and if those heuristics could simply be identified and then hard-coded into a set of simple rules for computers, then computers could achieve similar or even superior results to specialists.²⁹ It turned out that expert systems did not live up to the high expectations of researchers, and their limitations eventually led to a decline in interest in artificial intelligence generally.³⁰

More recently, however, artificial intelligence has had something of a renaissance, becoming associated with a new and promising technique known as deep learning.³¹ Deep learning is a sophisticated, data-driven model from a field of computer science known as machine learning. Machine learning, rather than relying on a set of rules established by a programmer, as in expert systems, instead starts with a data set and then attempts to derive rules on its own.³² For example, a computer might be given a set of images with labels attached—this is a person, this is a tree, this is a dog—and be asked to use the images (the “training set”) to establish rules for identifying when an image contains those objects.³³ The machine learning algorithm might then check the accuracy of these rules against another set of images (the “test set”) to see if the rules work well outside of the original context.³⁴ Many different variants of machine learning exist, and one particular variant, known as deep learning, has had notable success in recent years.³⁵ In deep learning algorithms, neural networks or units identify patterns in data, transform those patterns into outputs, and then pass those outputs along to additional units.³⁶ The first “layer” of units might identify patterns in the data, and then the next layer might identify patterns of patterns in the data, and so on and so forth.³⁷ The algorithms associated with deep learning techniques have proven remarkably accurate at improving accuracy and predictive power in machines, and they have become synonymous with artificial intelligence in many circles.

²⁹ See Andrea Roth, *Machine Testimony*, 126 YALE L. J. 1972, 1998–99 (2017).

³⁰ See DANIEL CREVIER, *AI: THE TUMULTUOUS SEARCH FOR ARTIFICIAL INTELLIGENCE* 204–8 (1993).

³¹ See Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. DAVIS L. REV. 399, 402 (2017); Jack M. Balkin, *The Three Laws of Robotics*, 78 OHIO ST. L. J. 1217, 1219–20 (2016); Emily Berman, *A Government of Laws and not of Machines*, 98 B.U. L. REV. 1277, 1284–90 (2018).

³² See KEVIN P. MURPHY, *MACHINE LEARNING: A PROBABILISTIC PERSPECTIVE* 1 (2012) (“[W]e define machine learning as a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty . . .”).

³³ See David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 684–88 (2017).

³⁴ *Id.*

³⁵ *Id.* at 669–70.

³⁶ See Curtis E.A. Karnow, *The Opinion of Machines*, 19 COLUM. SCI. & TECH. L. REV. 136, 141–50 (2017).

³⁷ *Id.*

But defining artificial intelligence in such a narrow way, to encompass a small set of algorithmic techniques and to exclude all others, has its problems as well. The explosion of interest in deep learning has only occurred in the last few years, and it may well be that other, newer strategies will emerge that tackle the same issues but use different methods. These newer strategies may well be more effective than older ones. There would not appear to be any categorical reasons for excluding new algorithmic strategies from the definition of artificial intelligence simply because they are unorthodox or new.

It is hard to regulate what we cannot define.³⁸ Hopefully this discussion clarifies some of the confusion in the field about what precisely we are talking about when we talk about artificial intelligence. Some people use it to refer generally to any instance when a computer performs an action correctly, while others use it to refer to a specific set of algorithmic strategies that are currently en vogue in the industry. When computer scientists talk about the *AI revolution* of recent years, they tend to be talking about the machine learning and deep learning strategies that researchers have deployed to such great success in certain fields. But policymakers, news media and commentators often use *artificial intelligence* to refer more generally to *thinking machines*, regardless of the actual techniques that the machines are using to accomplish their feats. It may be most useful to conceive of artificial intelligence as a spectrum, with the capacity to perform simple tasks (such as arithmetic and other basic mathematical calculations) at the low end, the capacity to match human performance somewhere in the middle, and the capacity to outperform humans at the high end. This conception is itself simplistic (computers are already better than humans at many tasks, while simultaneously being worse than humans at many others), but it helps clarify the range of potential artificial intelligence applications in the world. And while this Article will focus in particular on machine learning, it is worthwhile to keep in mind that some other technique may come along that replaces it.

Now that we have settled on a workable definition of artificial intelligence, we can now turn to its real-world uses. In particular, it is useful for our purposes (that is, understanding how artificial intelligence is affecting the financial industry, and how it might do so in the future) to know the types of things that current instantiations of artificial intelligence are good at doing. It is also important to know *why* they are good at doing these things.

A list of the accomplishments of artificial intelligence in the last decade is breathtaking. Here is just a short list of some of the most noteworthy achievements:

³⁸ For an excellent analysis of the regulatory issues raised by the uncertain scope of “artificial intelligence” as a concept, see Mark A. Lemley & Bryan Casey, *You Might be a Robot*, 105 CORNELL L. REV. (forthcoming 2019). Lemley and Casey conclude that, in fact, there is no correct definition of artificial intelligence for regulators, and that regulators should instead focus on behaviors and actions.

- In 2011, IBM's computing system, Watson, won a game of Jeopardy against two of the game show's most successful contestants, Ken Jennings and Brad Rutter, with a final tally of \$77,147 to Jennings' \$24,000 and Rutter's \$21,600;³⁹
- Between 2010 and 2017, in a popular image recognition contest, deep learning algorithms lowered their error rate from 28% to 2.3% (surpassing humans, who, on average have a 5% error rate);⁴⁰
- In 2016, by reviewing and analyzing all published literature on ALS, or Lou Gehrig's disease, IBM's Watson system was able to identify five previously unknown genes related to the disease;⁴¹
- In 2017, a computer program known as AlphaGo defeated the world's best player of the board game Go with a score of three matches to zero;⁴²
- In 2017, countries deployed forty-nine different automated weapons systems that could detect and attack targets without human intervention;⁴³
- In 2018, Waymo launched the world's first commercial self-driving car service, a technology based heavily on machine learning algorithms.⁴⁴

These remarkable advances, displayed across a wide and diverse array of fields, have been made possible by three related developments: better algorithms, deployed by more powerful computers, applied to ever greater amounts of data.⁴⁵ On the algorithm side, the major advances in recent years have been driven by machine learning, which, as described above, refer to a computer's ability to analyze data and develop models for recognizing patterns and predicting future data—in other words, its ability to learn and im-

³⁹ See STEPHEN BAKER, *FINAL JEOPARDY: THE STORY OF WATSON, THE COMPUTER THAT WILL TRANSFORM OUR WORLD* 251 (2012).

⁴⁰ See Imagenet, *Large Scale Visual Recognition Challenge*, <http://image-net.org/challenges/LSVRC/2017/>.

⁴¹ Emma Hinchliffe, *IBM's Watson Supercomputer Discovers 5 New Genes Linked to ALS*, MASHABLE, (Dec. 14, 2016), <http://mashable.com/2016/12/14/ibm-watson-als-research/#oKfRVPg3C8qI>.

⁴² A later version of the AlphaGo algorithm, known as AlphaGo Zero, defeated the prior version of AlphaGo 100 games to zero. See Larry Greenemeier, *AI versus AI: Self-Taught AlphaGo Zero Vanquishes Its Predecessor*, SCIENTIFIC AMERICAN, (Oct. 18, 2017).

⁴³ See VINCENT BOULANIN & MAAIKE VERBRUGGEN, *MAPPING THE DEVELOPMENT OF AUTONOMY IN WEAPON SYSTEMS* 26 (2017).

⁴⁴ See Shuyang Cheng & Gabriel Bender, *Autonomous Driving*, MEDIUM, (Jan. 15, 2019), <https://medium.com/waymo/automi-automating-the-design-of-machine-learning-models-for-autonomous-driving-141a5583ec2a> ("At Waymo, machine learning plays a key role in nearly every part of our self-driving system. It helps our cars see their surroundings, make sense of the world, predict how others will behave, and decide their next best move.").

⁴⁵ See NAT'L SCI. & TECH. COUNCIL, *supra* note 19, at 6; STANFORD UNIVERSITY ONE HUNDRED YEAR STUDY, *ARTIFICIAL INTELLIGENCE AND LIFE IN 2030* 51 (2016).

prove its performance over time.⁴⁶ Advancements in the sub-field of deep learning have played particularly important roles in this success.⁴⁷ These advancements have, in turn, been supported by enormous leaps in computing power.⁴⁸ Moore's Law—the famous proposition that the number of components per integrated circuit, and thus processing power, doubles roughly every eighteen to twenty-four months—demonstrates the magnitude of performance improvement over time.⁴⁹ If human populations followed the same trend, a village of 100 people in 1970 would have grown to 1.7 billion people by 2018. Even more relevant for artificial intelligence is the vast improvement in the speed of graphical processing units (GPUs) used in the field.⁵⁰ In 2000, Nvidia's top-of-the-line desktop GPU, the GeForce 256 DDR, had a processing power of 480 million operations per second. By 2018, Nvidia's top desktop GPU, the Titan RTX, had reached a processing power of 16,312 gigaflops, a speed more than 30,000 times faster.⁵¹ The vastly greater computational capacity of modern-day computers, combined with effective machine learning algorithms, have been unleashed on the massive amounts of data made available by the internet age.⁵² With the rise of the internet and mobile computing, people are generating more data, and more sensitive and personalized data, than they have ever done before. This information has proven a fertile source of training data for artificial intelligence.⁵³ As the computer scientist Andy Hertzfeld has stated, the artificial intelligence revolution of recent years

couldn't have happened no matter how brilliant your algorithm was twenty years ago, because billions were just out of reach. But what they found was that by just increasing the scale, suddenly things started working incredibly well. It kind of shocked the computer science world, that they started working so good.⁵⁴

So artificial intelligence has made enormous leaps in ability in recent years, and these leaps have made it both practical and efficient for firms and

⁴⁶ See Calo, *supra* note 31, at 402.

⁴⁷ *Id.* at 401–02.

⁴⁸ See Alan Stevens, *GPU Computing: Accelerating the Deep Learning Curve*, ZDNET, (July 2, 2019), <https://www.zdnet.com/article/gpu-computing-accelerating-the-deep-learning-curve/>.

⁴⁹ For a broader discussion of the effects of Moore's law on artificial intelligence, see John O. McGinnis, *Accelerating AI*, 104 NW. U. L. REV. 1253 (2010).

⁵⁰ See Michael Byrne, *Why Machine Learning Needs GPUs*, MOTHERBOARD, (Jan. 9, 2018), https://motherboard.vice.com/en_us/article/kznnnn/why-machine-learning-needs-gpus.

⁵¹ See CPUREVIEW, *nVidia GeForce 256 DDR*, <http://www.gpureview.com/GeForce256-DDR-card-114.html>; SKUEZTECH, *Nvidia Titan RTX*, <http://www.skueztech.com/db/gpu/nvidia-titan-rtx>.

⁵² See Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016); Neil M. Richards & Jonathan H. King, *Big Data Ethics*, 49 WAKE FOREST L. REV. 393 (2014).

⁵³ See Calo, *supra* note 31, at 420–25.

⁵⁴ ADAM FISHER, *VALLEY OF GENIUS: THE UNCENSORED HISTORY OF SILICON VALLEY* 423 (2018).

governments to use artificial intelligence to improve their day-to-day operations. And at least for now, there are no signs of this progress slowing down.

B. Artificial Intelligence Strategies in Finance

The remarkable power of artificial intelligence to identify patterns in complicated data and use these patterns to predict future behavior has made the field a major area of interest for financial firms. This is true both for the large Wall Street banks that have traditionally dominated finance and for the wave of smaller, nimbler fintech firms that have sprung up to challenge them in recent years. On the one hand, large financial institutions like JP Morgan, Goldman Sachs, and Morgan Stanley are increasingly hiring machine learning specialists and creating entire divisions devoted to artificial intelligence.⁵⁵ On the other, a handful of smaller, more focused financial firms have turned to artificial intelligence as a primary strategy for their businesses.⁵⁶ These trends suggest that artificial intelligence will likely play a large role in determining the future of the financial sector.⁵⁷

The potential uses of artificial intelligence in finance are nearly limitless. As Charles Elkan, the head of machine learning at Goldman Sachs, has said, “within every business area there are opportunities to apply machine learning.”⁵⁸ That said, most financial firms to date have focused on a few key areas where machine learning has obvious applications. First, they are using artificial intelligence to improve credit risk assessments for counterparties. Second, they are using it to protect themselves from fraud and wrongdoing, either external or internal to the firm. And third, they are using it to devise better trading strategies. It may be helpful to look at each of these areas in turn to get a better sense of how artificial intelligence can affect the financial world. One should keep in mind that most financial firms are relatively tight-lipped about how, precisely, they are using artificial intelligence in their businesses. It is, after all, competitively valuable information. At the same time, broad categories may be identified.⁵⁹

⁵⁵ See Sarah Butcher, *The Top Machine Learning Teams in Investment Banks*, eFINANCIAL-CAREERS, May 23, 2018.

⁵⁶ See Sameer Maskey, *How Artificial Intelligence is Helping Financial Institutions*, FORBES, (Dec. 5, 2018), <https://www.forbes.com/sites/forbestechcouncil/2018/12/05/how-artificial-intelligence-is-helping-financial-institutions/#441cd632460a>.

⁵⁷ See Thomas Vartanian, *Nefarious Nations Will Take the Lead on AI If the US Doesn't*, THE HILL, (Dec. 12, 2018) (“As part of the country’s critical infrastructure, the highest priority should be given to the development of a financial services and capital markets strategy to foster AI innovations while at the same time protecting against its unprecedented threats.”).

⁵⁸ Charles Elkan, *Will Machine Learning Transform Finance*, YALE INSIGHTS, (Jan. 3, 2019), <https://insights.som.yale.edu/insights/will-machine-learning-transform-finance>.

⁵⁹ For a comprehensive industry survey of bankers’ opinion on artificial intelligence in finance, see Laura Noonan, *AI in Banking: The Reality Behind the Hype*, FINANCIAL TIMES, (Apr. 11, 2018), <https://www.ft.com/content/b497a134-2d21-11e8-a34a-7e7563b0b0f>.

Perhaps no area of artificial financial intelligence has received as much attention (or criticism) than its use in credit ratings.⁶⁰ Banks are in the business of loaning money to people—bank loan volume is a carefully watched sign, not just of bank profits, but also of economic health more generally.⁶¹ One of the primary concerns that banks have when they extend a loan, of course, is that the borrower may not pay the money back. This means that the ability to accurately predict the credit risk of borrowers is tremendously important (and valuable) to financial institutions as a whole. They often outsource this process to credit scoring agencies, such as Equifax, Experian, and TransUnion. But as many observers have noted, the credit rating agencies' process itself is far from perfect.⁶² Joe Nocera of the *New York Times* has written that a "credit score is derived after an information-gathering process that is anything but rigorous" but, nevertheless, "essentially . . . has become the only thing that matters anymore to the banks and other institutions that underwrite mortgages."⁶³ Many financial firms believe that machine learning algorithms could do better. The theoretical benefits are clear. By sifting through massive amounts of prior data about both defaulting and non-defaulting borrowers, machine learning algorithms might be able to identify hidden or unexpected variables that affect a borrower's likelihood of repaying a loan. Rather than relying on simple and obvious relations, such as whether a borrower has defaulted on past loans or has large amounts of credit card debt, machine learning algorithms might find that other, less noticeable information, such as a borrower's purchase history, friend group, or Twitter posts, provides meaningful information about his ability to repay a loan. And indeed, financial firms are increasingly using artificial intelligence to make these difficult assessments. For example, Zest Finance, a fintech startup based in California, offers a machine learning-based credit modeling product for mortgages, auto loans, business loans, and consumer loans.⁶⁴ It claims that its algorithm, on average, leads to a 15% increase in loan approval rates and a 30% decrease in charge-off rates, and that it only requires

⁶⁰ See Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 34 (2014); Christopher K. Odinet, *Consumer Bit-Credit and Fintech Lending*, 69 ALA. L. REV. 781, 858 (2018); Matthew A. Bruckner, *Regulating Fintech Lending*, 37 BANKING & FIN. SERVICES POL'Y REP. 1, 7 (2018); Matthew Adam Bruckner, *The Promise and Perils of Algorithmic Lenders' Use of Big Data*, 93 CHIKENT L. REV. 3, 60 (2018).

⁶¹ See Peter Eavis, *Is a Slowdown in Bank Lending a Bad Sign for the Economy*, N.Y. TIMES, (Oct. 12, 2018), <https://www.nytimes.com/2018/10/12/business/dealbook/bank-lending-slowdown-economy.html>; Rachel Louise Ensign, *Business-Loan Drought Ends for Banks*, WALL ST. J., (July 8, 2018), <https://www.wsj.com/articles/business-loan-drought-ends-for-banks-1531058400>.

⁶² See Citron & Pasquale, *supra* note 60, at 10–16.

⁶³ Joe Nocera, *Credit Score Is the Tyrant in Lending*, N.Y. TIMES, (July 23, 2010), <https://www.nytimes.com/2010/07/24/business/24nocera.html>.

⁶⁴ See Steve Lohr, *Banking Start-Ups Adopt New Tools for Lending*, N.Y. TIMES, (Jan. 18, 2015), <https://www.nytimes.com/2015/01/19/technology/banking-start-ups-adopt-new-tools-for-lending.html>.

three months to implement.⁶⁵ The German startup Kreditech uses big data and machine learning to assess borrowers' creditworthiness and even extends loans, both domestically and internationally, in real time using its automated system.⁶⁶ Their model is so attractive and effective that some of the credit scoring agencies are themselves getting into the game, with Equifax recently adopting machine learning in its models and Experian similarly offering a machine learning product to their clients.⁶⁷ And, to be clear, the effects of AI-enhanced credit rating are not simply negative, providing a rationale for declining a loan to someone who might otherwise have received one. AI may also show that certain groups have been systematically *underrated* by current systems, and thus more deserving of loans than traditionally believed.

Artificial intelligence also plays an important role in financial institutions' identification and prevention of fraud and wrongdoing. Finding effective ways to prevent fraudulent payments or claims has been a constant thorn in the side of the financial industry. And despite considerable efforts, fraud is rising—in 2016, 15.4 million U.S. consumers were victims of identity fraud, an all-time high, leading to \$16 billion in losses.⁶⁸ Most of these losses stem from credit card fraud.⁶⁹ In order to better identify these fraudulent behaviors and, ideally, prevent them before they occur, financial institutions are increasingly turning to machine learning algorithms.⁷⁰ Machine learning algorithms, after all, can analyze and assess the tremendous amounts of transaction data constantly flowing across the sector much more quickly than a human ever could. Bank of America is exploring how to incorporate artificial intelligence into its fraud detection systems.⁷¹ The startup Feedzai specializes in using machine learning to prevent fraudulent payments in the financial industry, and its success has led Citigroup to incorporate its product into its own fraud prevention systems.⁷² Monzo, the popular British credit card company, deployed a machine learning model for preventing fraud in its prepaid debit cards, allowing it to reduce fraud rates from 0.84%

⁶⁵ See ZESTAI, *What is ZAML?*, <https://www.zest.ai/zaml>.

⁶⁶ See WALL ST. J., *Germany's Kreditech Scores \$40M for Automated Lending*, (June 25, 2014).

⁶⁷ See Alex Hickey, *Equifax Debuts Machine Learning-Based Credit Scoring System*, CIO DIVE (Mar. 28, 2018), <https://www.ciodive.com/news/equifax-debuts-machine-learning-based-credit-scoring-system/520095/>.

⁶⁸ See AnnaMaria Andriotis & Peter Rudegeair, *Credit-Card Fraud Keeps Rising, Despite New Security Chips—Study*, WALL ST. J., (Feb. 1, 2017), <https://www.wsj.com/articles/credit-card-fraud-keeps-rising-despite-new-security-chipsstudy-1485954000>.

⁶⁹ *Id.*

⁷⁰ See Sara Castellanos & Kim S. Nash, *Bank of America confronts AI's "Black Box" With Fraud Detection Effort*, WALL ST. J., (May 11, 2018), <https://blogs.wsj.com/cio/2018/05/11/bank-of-america-confronts-ais-black-box-with-fraud-detection-effort/>.

⁷¹ *Id.*

⁷² See BUS. WIRE, *Citi Partners with Feedzai to Provide Machine Learning Payment Solutions*, (Dec. 19, 2018), <https://www.businesswire.com/news/home/20181219005388/en/Citi-Partners-Feedzai-Provide-Machine-Learning-Payment>.

to less than 0.01% in just six months.⁷³ Other financial institutions are experimenting with using machine learning more broadly to prevent bad behavior among their own employees. In 2018, ING Bank partnered with a machine learning startup specializing in speech recognition to spot and stop insider trading and collusion within the firm.⁷⁴

Finally, financial firms are also using artificial intelligence to improve investment returns. Again, the change here is more a matter of degree than of kind. Investment firms, both large and small, have long used computers to analyze stock information, identify price discrepancies, and make investments.⁷⁵ The artificial intelligence revolution of recent years, however, has opened up new avenues for incorporating algorithms more deeply into investment strategies. One clear use case is for quant hedge funds, which use algorithms to determine investment decisions.⁷⁶ Already familiar (and comfortable) with computerized, automated investment decisions, quant hedge funds might be expected to be early adopters of machine learning strategies. And indeed, it appears that they are. In 2017, one study found that 20% of hedge funds make use of machine learning or artificial intelligence in their investment strategies.⁷⁷ Just a year later, that number had risen to 56%.⁷⁸ Large investment firms are investing heavily in machine learning as well. BlackRock, the asset management company, uses artificial intelligence to spot trends in financial data, including finding relationships between securities or market indicators, scanning social media for insights into employee sentiment, and analyzing search engines for popular search terms.⁷⁹ Regular investors can now take advantage of machine learning-based investment strategies as well. In 2017, EquBot launched an exchange traded fund, called AI Powered Equity, that uses machine learning algorithms to decide on its investment components.⁸⁰ EquBot states that its mission is to “give everyone

⁷³ See MONZO, FIGHTING FRAUD WITH MACHINE LEARNING, (Feb. 3, 2017), <https://monzo.com/blog/2017/02/03/fighting-fraud-with-machine-learning/>; see also THE ECONOMIST, *Machine-Learning Promises to Shake Up Large Swathes of Finance*, (May 25, 2017), <https://www.economist.com/finance-and-economics/2017/05/25/machine-learning-promises-to-shake-up-large-swathes-of-finance>.

⁷⁴ See Intelligent Voice, *Compliance Monitoring Enters AI Age at ING with Intelligence Voice and Relativity Trace*, (Dec. 3, 2018), <https://www.intelligentvoice.com/2018/12/03/compliance-monitoring-enters-ai-age-at-ing-with-intelligent-voice-and-relativity-trace>.

⁷⁵ See, e.g., Edward L. Pittman, *Quantitative Investment Models, Errors, and the Federal Securities Laws*, 13 N.Y.U. J. L. & BUS. 633, 643–63 (2017).

⁷⁶ See generally Thomas J. Brennan & Andrew W. Lo, *Dynamic Loss Probabilities and Implications for Financial Regulation*, 31 YALE J. ON REG. 667, 692–95 (2014); SCOTT PATTERSON, *THE QUANTS: HOW A NEW BREED OF MATH WHIZZES CONQUERED WALL STREET AND NEARLY DESTROYED IT* (2011).

⁷⁷ See Amy Whyte, *More Hedge Funds Using AI, Machine Learning*, INSTITUTIONAL INVESTOR (July 19, 2018), <https://www.institutionalinvestor.com/article/b194hmlkjbvd37/More-Hedge-Funds-Using-AI-Machine-Learning>.

⁷⁸ *Id.*

⁷⁹ See Conrad De Aenlle, *A.I. Has Arrived in Investing. Humans Are Still Dominating.*, N.Y. TIMES, (Jan. 12, 2018).

⁸⁰ See Bailey McCann, *The Artificial-Intelligent Investor: AI Funds Beckon*, WALL ST. J., (Nov. 5, 2017).

access to investment opportunities that artificial intelligence can uncover.⁷⁸¹ Another potential use for artificial financial intelligence is in investment banking, where firms might use artificial intelligence algorithms to identify opportunities for acquisitions or divestments that might not be readily apparent to insiders.⁸²

All of these developments suggest that artificial financial intelligence is poised to become an integral part of the financial industry. Financial firms are actively researching, and in many cases adopting, artificial intelligence strategies. These strategies have been particularly useful in assessing risk, preventing fraud, and investing capital. And the trend is clear—artificial financial intelligence is on the rise. As further proof, one need only look as far as the chartered financial analyst (CFA) exam, the prized certification for future investment professionals. In 2019, the exam added computer science, with a focus on artificial intelligence and data mining, as a new subject that all aspiring CFAs must have familiarity with by 2020.⁸³ Future bankers will know much more about artificial intelligence than they do today.

C. Existing Literature

The stunning success of artificial intelligence in recent years has sparked renewed interest in the field from academia. Scholars have begun to ask pointed questions about the economic, social, and moral consequences of turning over more and more of our lives to the control of algorithms. Legal scholars, as they are wont to do, have proposed new laws to constrain and guide the use of artificial intelligence. And while the literature is broad and continually growing, much like artificial intelligence itself, a few common threads can be identified in the scholarship. First, a group of scholars has argued that the progress of artificial intelligence risks widespread negative effects on labor markets. Second, a number of scholars have argued that the results of artificial intelligence are so difficult to explain, and, thus, to scrutinize and monitor, that artificial intelligence will end up undermining the regulatory state. Finally, a group of scholars has argued that artificial intelligence methods might perpetuate racial or other discrimination in broader society.

On the question of employment, a number of scholars have raised concerns about artificial intelligence's capacity to supplant and replace human workers.⁸⁴ Indeed, President Barack Obama issued a report on the topic, con-

⁸¹ EquBot Website, <https://equbot.com>. The ETF has, however, underperformed—from its inception in October 2017 to the end of 2018, it rose 3.1%, while the S&P 500 rose 5.1%. See De Aenlle, *supra* note 79.

⁸² See Elkan, *supra* note 58.

⁸³ See Trevor Hunnicutt, *CFA Exam to Add Artificial Intelligence, "Big Data" Questions*, REUTERS, (May 23, 2017).

⁸⁴ See MARTIN FORD, *THE RISE OF THE ROBOTS: TECHNOLOGY AND THE THREAT OF A JOBLESS FUTURE* (2016); Erik Brynjolfsson & Andrew McAfee, *Human Work in the Robotic*

cluding that we need new policies and laws to “educate and prepare new workers to enter the workforce, cushion workers who lose jobs, keep them attached to the labor force, and combat inequality.”⁸⁵ The idea of machines replacing humans in the labor force is not new—they have been doing so since the Industrial Revolution. But artificial intelligence’s vastly greater power, combined with its speed of improvement and breadth of application, present these concerns on a grander scale. Some scholars argue that artificial intelligence will lead to job losses as artificial intelligence simply eliminates the need for human workers in many fields.⁸⁶ Others argue that even if artificial intelligence does not lead to job losses, it will still lead to increased wealth disparities (or “income polarization”) as the techno-savvy reap the benefits of cheaper labor and the techno-ignorant are relegated to low wages and long hours.⁸⁷ In either case, the artificial intelligence revolution could cause major labor force disruptions.

Another group of scholars has argued that the problem with artificial intelligence is explainability.⁸⁸ One of the issues with machine learning-produced results is that they are often difficult to interpret in easily understood language. Imagine, for example, if you asked a bank why they turned you down for a loan, and, in reply, they handed you the source code of their machine learning algorithm and the values of millions of automatically tuned algorithm parameters—this might be an accurate description of what they had in fact done to reach their result, but it would not be very helpful for understanding the motivation for the decision, or what you might do to change it. This is less of a problem in some fields than in others. For example, if a machine learning algorithm is better at spotting cancer than human experts, we probably do not care much about *how* it is doing it. If it turns out that the machine learning algorithm spotted a complex array of patterns that have no intuitive counterpart in human reason, we probably would not care, as long as it *works*. But when we turn from the world of facts (do I have cancer?) to a world of judgment (should I loan money to this person?), then the process and rationale of the decisionmaker suddenly takes on greater importance. Regulators need to understand algorithms in order to ensure

Future: Policy for the Age of Automation, FOREIGN AFFAIRS, (July/Aug., 2016); Calo, *supra* note 31, at 425–27.

⁸⁵ EXEC. OFFICE OF THE PRESIDENT, ARTIFICIAL INTELLIGENCE, AUTOMATION, AND THE ECONOMY (2016), <https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.pdf>.

⁸⁶ See Michael Guihot, et al., *Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence*, 20 VAND. J. ENT. & TECH. L. 385 (2017).

⁸⁷ See Cynthia Estlund, *What Should We Do After Work? Automation and Employment Law*, 128 YALE L. J. 254 (2018).

⁸⁸ See Cary Coglianese & David Lehr, *Transparency and Algorithmic Governance*, 71 ADMIN. L. REV. 1 (2019); Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085 (2018); David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653 (2017); Joshua A. Kroll, et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633 (2017).

they comply with the law.⁸⁹ Consumers need to understand the algorithms to trust that they have been dealt with fairly, and to be able to change their behaviors.⁹⁰ Indeed, the European Union has gone so far as to require companies that employ automated decisionmaking to provide “meaningful information about the logic involved.”⁹¹ As one group of scholars has put it, “[b]ecause automated decision systems can return potentially incorrect, unjustified, or unfair results, additional approaches are needed to make such systems accountable and governable.”⁹²

In a similar vein, some scholars worry that the deployment of artificial intelligence could lead to disguised discrimination.⁹³ The logic is as follows. Machine learning algorithms attempt to identify patterns in large data sets, and then draw conclusions from those patterns. In order to do their work, then, machine learning algorithms must be fed data, and large amounts of it. But that data itself is not necessarily neutral. It may (in fact, it likely will) incorporate some form of bias or discrimination from the outside world. And if the data is biased or discriminatory, then the machine learning algorithm itself may become biased or discriminatory, too. This could happen in one of two ways: first, a nefarious actor might intentionally feed the machine learning algorithm bad data;⁹⁴ and second, a perfectly innocent, but naïve, actor might unintentionally use data that has unrecognized bias in it.⁹⁵ As an example of the first problem, intentional discrimination, if a firm (or a rogue employee at a firm) wanted to refrain from doing business with individuals of certain races or religions, it might construct a machine learning algorithm that *looked* unbiased but in fact encoded certain nefarious factors. After all, software engineers are the ones who have to make the hard decisions about what data to use, how to structure the data, and how to interpret it. Bias could creep in at any of these levels, and it might be difficult for outside observers to detect. But even if there is no intentional discrimination in the machine learning algorithm or the data set, it may still reflect the effects of past discrimination in the world. For example, if minorities received poor credit scores in the past due to discrimination, then machine learning algorithms might *learn* that these minorities should be scored lower than their

⁸⁹ See Selbst & Barocas, *supra* note 88, at 1089–90.

⁹⁰ *Id.* at 1119–22.

⁹¹ Commission Regulation 2016/679, 2016 O.J. (L 119) 1.

⁹² Kroll et al., *supra* note 88, at 633.

⁹³ See *id.* at 678–82; Citron & Pasquale, *supra* note 60, at 10–16; Margaret Hu, *Algorithmic Jim Crow*, 86 FORDHAM L. REV. 633 (2017); Barocas & Selbst, *supra* note 88; Anupam Chander, *The Racist Algorithm?*, 115 MICH. L. REV. 1023 (2017); Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 54 (2019) (arguing that “[t]he issue of algorithmic bias represents a crucial new world of civil rights concerns, one that is distinct in nature from the ones that preceded it”); Kate Crawford, *Think Again: Big Data*, FOR. POL’Y, (May 10, 2013); Talia B. Gillis & Jann Spiess, *Big Data and Discrimination*, 86 U. CHI. L. REV. 458 (2019) (arguing that machine learning algorithms hinder the application of discrimination laws).

⁹⁴ See Lehr & Ohm, *supra* note 33, at 703–05.

⁹⁵ *Id.*

the problem of “non-stationary” behavior.¹⁰³ As noted before, scholars have pointed out that if a data set reflects a biased world (such as low credit scores for racial minorities), then machine learning algorithms based on those data sets may well reinforce those biases. This is a major problem, and has slightly different valences in the financial sector, which has a major influence on the direction of economic development. Imagine, for example, a machine learning algorithm that was trained on a data set from the period of 1995 to 2000. Basing its recommendations on the relative success of various industries during the dotcom boom, it might have concluded that, as a general rule, technology startups vastly overperformed other sectors. It might well have recommended that internet-focused companies such as Pets.com (which raised \$82 million in an IPO in 2000 before going bankrupt just nine months later)¹⁰⁴ or GeoCities (which was bought by Yahoo for \$3.6 billion in 1999 but quickly collapsed)¹⁰⁵ were excellent investments. The machine learning algorithm, by basing its recommendations on prior data from a market that was fundamentally different from the market post-2000, was simply ineffective. The data was incurably biased. Similarly, the use of machine learning might lead to the prolongation or even perpetuation of short-term trends in markets, to the detriment of economic efficiency. Without data from the period after 2000, when the dotcom bubble burst, the machine learning algorithm might well have recommended investments in the tech sector and, by doing so, reinforced preexisting patterns of industry focus. One potential consequence of this is that bubbles could become more dramatic—machine learning algorithms could magnify momentum in particular sectors or trends, leading to eventual and catastrophic collapse. Another potential consequence, and one that is perhaps even more troubling, is that artificial financial intelligence might *prevent* bubbles from bursting. In other words, if a large section of financial institutions uses machine learning algorithms to decide on their investment strategies, and those machine learning algorithms conclude that, say, the technology sector outperforms other sectors, then it may be difficult for the market to self-correct. Other industries might simply wither away from lack of investment or access to capital.

Third, the dependence of machine learning algorithms on data creates strong incentives for financial institutions not only to put the data they currently have to better use, but also to gather more of it.¹⁰⁶ If the quality of artificial intelligence as a financial tool grows as the size of its data set increases, then financial firms may search for new ways to get increasingly

¹⁰³ See MASASHI SUGIYAMA & MOTOAKI KAWANABE, MACHINE LEARNING IN NON-STATIONARY ENVIRONMENTS: INTRODUCTION TO COVARIATE SHIFT ADAPTATION (2012).

¹⁰⁴ See Pui-Wing Tam & Mylene Mangalindan, *Pets.com Will Shut Down, Citing Insufficient Funding*, WALL ST. J. (Nov. 8, 2000), <https://www.wsj.com/articles/SB973617475136917228>.

¹⁰⁵ See Robert Cyran, *Yahoo Is a Case Study in Poor Timing*, N.Y. TIMES, (July 26, 2016), <https://www.nytimes.com/2016/07/27/business/dealbook/yahoo-is-a-case-study-in-poor-timing.html>.

¹⁰⁶ See Calo, *supra* note 31, at 420–23.

intrusive and sensitive data on citizens. Their practices could quickly raise privacy concerns. Part of the appeal of artificial intelligence is that it aims to find patterns and relations that humans would never consider. But it cannot find these patterns if it does not have data to work on. Engineers or investment advisers that decide what information to include in data sets therefore have a potentially unlimited world of information to scour. Perhaps your Facebook friend list contains useful information about your credit score. Perhaps the prior personal relationships of a CEO have correlations with company performance. Perhaps the email patterns, mobile phone conversations, or locations of consumers contain relevant information for mortgage default rates. We have already seen a backlash against the tech giants for their gathering, storage, and use of user data.¹⁰⁷ But as the stories of privacy violations proliferate, the value of data to firms is becoming increasingly clear. Artificial intelligence may accelerate that trend.

Finally, and on a related note, the data dependency of artificial intelligence gives significant advantages to firms that have access to large data sets.¹⁰⁸ Large Wall Street firms are sprawling, multi-faceted companies that operate in hundreds of different markets and sectors. This means that they can see and track more information than anyone else. And at the risk of repetition, access to data in a machine learning environment is a competitive advantage. At the extreme, this might lead to large financial institutions taking an insurmountable dominant position that no smaller startups could challenge. This worry about artificial intelligence's benefits to the powerful has been raised in the national security world. A recent study by the Belfer Center at Harvard University concluded that, at least in the near term, "bringing AI technology applications into the cyber domain will benefit powerful nation-state actors."¹⁰⁹ The most powerful nation states are the actors that have the scientific knowledge, research budgets, and access to information that are necessary to develop artificial intelligence systems. The same might be said of financial institutions: large financial institutions can hire away artificial intelligence experts from academia, industry, and their competitors, and they have the capital to fund long-term research into artificial intelligence's applications for finance. Small startups simply do not have this capacity. Over time, this trend might exacerbate the already sizeable amount of concentration in the financial sector. To be sure, the long-term consequences of artificial intelligence might be to level the playing field, not tilt it. Machine learning algorithms are, after all, algorithms, which can be copied and disseminated much more easily than hard assets. To the extent that prominent artificial intelligence strategies become widely accessible and usable by new financial firms, they may serve as a powerful equalizing fac-

¹⁰⁷ See Anita L. Allen, *Protecting One's Privacy in a Big Data Economy*, 130 HARV. L. REV. F. 71, 71–72 (2016).

¹⁰⁸ See Calo, *supra* note 31, at 424.

¹⁰⁹ See ALLEN & CHAN, *supra* note 7, at 20.

tor in financial markets. But until then, the concern about concentration in the financial sector will persist, and may, at the extreme, create antitrust problems.¹¹⁰

The data dependency problem that artificial intelligence struggles with, and that raises issues for the financial sector, highlights a broader feature of our financial system and the potential limitations of artificial financial intelligence. Capital markets simply are not like the kinds of problems that artificial intelligence has, to date, excelled at. In cancer screening, image recognition, and language cognition, there is a *fact* that the artificial intelligence algorithm is attempting to discern—whether someone has cancer, what that image depicts, what that person said. But in capital markets, there is no similar factual counterpart. To be sure, there are factual *elements* of markets, from EBITDA to market capitalization to employee numbers and product sales (although even these can be fiddled with). But the fundamental feature of capital markets—the allocation of value to companies—is not so much a fact as a belief. In order to outperform the market, one needs to understand not only the facts about a company, but also how other people will view those facts. This is important because the very fact that a trend has been discovered often means that the trend disappears. Momentum, small capitalization, low volatility: all of these are market trends that have, at one time or another, delivered market superior returns.¹¹¹ But once others learn of the trend, they can quickly adopt strategies that take advantage of it, and the very fact of their adopting those strategies eliminates the market superior returns. This is a fundamental feature of capital markets. Indeed, it is the underlying basis of efficient capital markets theory. Artificial intelligence is not well-equipped to deal with these sorts of problems.

B. Overweighting

Another issue with the use of artificial financial intelligence is the potential for human decisionmakers to rely too heavily on its recommendations. Given the difficulty of explaining and understanding machine learning algorithms and the outputs they generate, financial decisionmakers, including consumers, might simply default, without deliberation or debate, to accepting the conclusions or recommendations that the machine learning algorithms make. Even if decisionmakers are aware of the limitations of machine learning, in the absence of clear methods for refuting or disproving the artificial intelligence outcome, artificial intelligence may take on undue weight in the structure of the financial industry.¹¹²

¹¹⁰ See Van Loo, *supra* note 9, at 251.

¹¹¹ See *Is Efficient-Market Theory Becoming More Efficient?*, THE ECONOMIST, May 27, 2017, <https://www.economist.com/finance-and-economics/2017/05/27/is-efficient-market-theory-becoming-more-efficient>.

¹¹² Of course, this may be desirable if we believe that financial decisionmakers are already biased in problematic ways. Leaning on less-biased, or less harmfully biased, machines may

We have already discussed one way in which machine learning algorithms fail: they often struggle with problems that possess non-stationary qualities.¹¹³ When the fundamental nature of the market studied (for example, the credit risk of borrowers, the stock prices of companies, or the methods of fraudsters) changes from one period to the next, the patterns and trends that a machine learning algorithm can identify from one period's data will not apply to the next. The non-stationary problem is particularly pronounced where the mere knowledge of the trend may lead to the trend itself disappearing, as might be the case if everyone knew that, say, the stock market always rose 10% every year. Arbitrageurs could simply immediately push the price of the stock up at the beginning of the year or even earlier. The trend would then disappear.

But machine learning algorithms also struggle with other known problems. One is "overfitting."¹¹⁴ The goal of machine learning algorithms is to find statistical trends in data and then generalize them so that they work for new data. But if the algorithm generalizes in a way that incorporates trends that are based on random factors that do not apply outside the particular data set, then it is said to *overfit* the data, that is, it finds trends that match the training data but do not match the outside world. To take a simplified version of this, imagine if a machine learning algorithm only had data on two companies, Apple and Kodak. It might conclude that all companies that start with "A" outperform the market, and all companies that start with "K" go bankrupt. This would obviously fit the training data perfectly, but it likely would not generalize to the wider world of non-Apple and non-Kodak companies. The process of finding trends that do not just fit the available data, but will also generalize well to future or outside data, is a difficult one, and one that machine learning algorithms have a hard time doing.

Another known problem for artificial intelligence algorithms is low-prevalence data.¹¹⁵ The idea here is that machine learning algorithms need a certain threshold amount of data on a given topic before being able to make reliable predictions about it. But where certain subpopulations within a data set are rare and different from the broader data set, machine learning algorithms may simply fail to draw accurate conclusions about them. This has been a serious problem in facial recognition algorithms: one study found that they incorrectly classified only 1% of light-skinned men, but more than one

then be preferable. See Holger Spamann & Lars Klöhn, *Justice is Less Blind, and Less Legalistic, Than We Thought: Evidence from an Experiment with Real Judges*, 45 J. LEG. STUD. 255, 255 (2016) (finding that judges' decisions were affected by irrelevant characteristics of defendants but failed to mention these characteristics in their written reasoning).

¹¹³ See SUGIYAMA & KAWANABE, *supra* note 103, at xi.

¹¹⁴ See STUART J. RUSSELL & PETER NORVIG, *ARTIFICIAL INTELLIGENCE: A MODERN APPROACH* 705 (3rd ed. 2010); RICHARD BERK, *STATISTICAL LEARNING FROM A REGRESSION PERSPECTIVE* 142 (2008); Lehr & Ohm, *supra* note 33, at 684.

¹¹⁵ See Lehr & Ohm, *supra* note 33, at 678–79.

peers. Even if the algorithm explicitly could not take race into account, it might find that other factors that correlate with race (such as names, geography, or other information) are just as effective.⁹⁶ While scholars do not have easy solutions to these problems, they have pointed out their deeply harmful effects.

II. ARTIFICIAL FINANCIAL INTELLIGENCE

Artificial intelligence has grown exponentially in recent years as more powerful computers apply more advanced algorithms to analyze more expansive data sets. Financial institutions have increasingly deployed these artificial intelligence strategies to improve their own results. Scholars have, in turn, begun to grapple with the thorny legal and ethical issues raised by the incorporation of artificial intelligence into the financial world. They have raised questions about its effects on labor, its inscrutable outputs, and its capacity for harmful discrimination. All of these are important problems, and ones that apply to the financial world. But at the heart of these critiques is the assumption that artificial intelligence's dangers are rooted in its *effectiveness*. In other words, artificial intelligence may lead to widespread job losses, opaque financial decisions, or even discriminatory results, because it is so capable at what it does. It is simply better at reading data and finding patterns than humans are, and in doing so, it may create risks. But this Part will argue that artificial financial intelligence's primary risks are not the product of its potential to surpass human intelligence. To date, prognostications about artificial intelligence's "Cambrian explosion" are exaggerated. Instead, the real risk from artificial financial intelligence is its capacity to magnify human failings. If artificial intelligence ends up being incorporated widely in the financial world, its harms will likely stem from inadequate data input, improper usage, and feedback effects between algorithms. These problems are dangerous and must be contained, but they are also, in a sense, less radical than some of the concerns popular in the media today.

A. Data Dependency

A common critique of artificial intelligence is that it relies heavily on prior data sets.⁹⁷ As mentioned earlier, the machine learning algorithms that have become the focus of current artificial intelligence research depend on receiving and analyzing large amounts of data. This leads to several related problems, some of which are specific to the financial sector, and others of

⁹⁶ See Citron & Pasquale, *supra* note 60, at 14; Rory Van Loo, *The Corporation as Court-house*, 33 YALE J. ON REG. 547, 579–80 (2016) (observing that businesses' algorithms may lead to racial discrimination by using socioeconomic factors as a proxy); Anya Prince & Daniel Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data*, 105 IOWA L. REV. 1257 (2020).

⁹⁷ See Brummer & Yadav, *supra* note 11, at 274–75.

which are more general. As *The Economist* has written, “the world’s most valuable resource is no longer oil, but data.”⁹⁸

First, the use of artificial financial intelligence will cause important financial decisions to increasingly turn not on human judgments about the wisdom or soundness of the decision, but rather on the quality and nature of the data that are fed into machine learning algorithms. This information is, necessarily, determined by humans. Its effectiveness as a training tool will thus depend heavily on the knowledge and sophistication of the employees who decide which data to use and how to appropriately code it for machine learning algorithms.⁹⁹ IBM’s Watson supercomputer, which uses machine learning techniques and was widely billed as a potentially lifesaving cancer detecting tool, was recently found to be issuing “unsafe and incorrect” treatment recommendations. While the causes are unclear, some observers believe it is due to engineers training the software on small numbers of hypothetical cancer cases, rather than real patient data.¹⁰⁰ Others believe that the machine learning algorithm simply does not do a good job at handling rare or recurring cancers because it simply lacks data on them.¹⁰¹ And the financial world has seen similar mistakes. McKinsey reported that one Asia Pacific bank lost \$4 billion when it applied interest rate models based on improperly entered data.¹⁰² Thus, the use of machine learning in financial decisions raises concerns about potentially harmful inaccuracies contained in data sets, as well as the possibility that these data sets will tend to miss out on rare but important financial occurrences (the *black swan* types of events that many investors focus on).

Second, even if the data that is used in financial applications of machine learning is not flawed, its use may create a series of related problems in market efficiency. For one, the data may poorly reflect the nature of financial markets. Machine learning algorithms, by their very nature, aim to find statistical trends in data and then generalize them so they work for new data. This can be quite useful in a number of contexts, but it tends to run into problems when the nature of the field changes over time. This is known as

⁹⁸ See *The World’s Most Valuable Resource Is No Longer Oil, But Data*, THE ECONOMIST (May 6, 2017), <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>.

⁹⁹ See Brummer & Yadav, *supra* note 11, at 274–75.

¹⁰⁰ See Julie Spitzer, *IBM’s Watson Recommended “Unsafe and Incorrect” Cancer Treatments, STAT Report Finds*, BECKER’S HEALTH IT & CIO REPORT (July 25, 2018), <https://www.beckershospitalreview.com/artificial-intelligence/ibm-s-watson-recommended-unsafe-and-incorrect-cancer-treatments-stat-report-finds.html>.

¹⁰¹ See Daniela Hernandez & Ted Greenwald, *IBM Has a Watson Dilemma*, WALL ST. J., (Aug. 11, 2018), <https://www.wsj.com/articles/ibm-bet-billions-that-watson-could-improve-cancer-treatment-it-hasnt-worked-1533961147>.

¹⁰² See Philipp Harle, et al., *The Future of Bank Risk Management*, MCKINSEY & COMPANY, (July 2016), <https://www.mckinsey.com/business-functions/risk/our-insights/the-future-of-bank-risk-management>.

third of dark-skinned women.¹¹⁶ In the financial world, this might lead to similar problems, such as incorrectly risk-scoring minorities, but it might also lead to quite different ones, such as failing to identify uncommon but significant subpopulations of a market, such as *unicorn* startups that have provided enormous returns to venture capitalists.¹¹⁷

To be sure, machine learning experts are well aware of these problems. They have devised ways to deal with nonstationary markets, overfitting algorithms, and low-prevalence data. But these methods are imperfect and, at most, can limit the problems, not resolve them. One researcher I spoke with has a rule of thumb for assessing the claims of new artificial intelligence products, something he calls the “87% rule.” If an engineer or entrepreneur claims that his artificial intelligence software can achieve accuracy rates of higher than 87% purely by fitting data, then the researcher assumes that the software is either overfitting or unrealistically constraining the prediction problem in a way that does not accurately reflect real-world conditions. If, instead, data-driven predictions provide accuracy around 87% and domain expertise or insight are layered in to boost accuracy above that number, then he deems it consistent with systems that have been deployed successfully in practice. In other words, even with cutting-edge machine learning algorithms deployed by experts in the field, the best-case result for most artificial intelligence strategies still fails a significant portion of the time.

However, there are a number of reasons to believe that, once financial firms have incorporated artificial intelligence into their systems, it will be hard to resist its persuasiveness.¹¹⁸ The ready availability of a sophisticated recommendation based on millions of data points, and the relative difficulty of scrutinizing the reasoning or rationale of the recommendation, may lead banks and firms to overweight artificial intelligence relative to other decisionmaking processes. As one doctor has said about the use of artificial intelligence in hospitals, “In my practice, I’ve often seen how any tool can quickly become a crutch—an excuse to outsource decision making to someone or something else.”¹¹⁹ This anecdotal evidence is only bolstered by our growing knowledge of common cognitive biases in decisionmaking.

One well-known bias in human decisionmaking is the availability bias.¹²⁰ Availability bias refers to the tendency of decisionmakers, in the face

¹¹⁶ Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROC. MACHINE LEARNING RES. 1 (2018), <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>.

¹¹⁷ See Jennifer S. Fan, *Regulating Unicorns: Disclosure and the New Private Economy*, 57 B.C. L. REV. 583 (2016).

¹¹⁸ See Nizan Geslevich Packin, *Consumer Finance and AI: The Death of Second Opinions?*, N.Y.U. J. LEGIS. & LEG. PUB. POL’Y (forthcoming 2020).

¹¹⁹ Dhruv Khullar, *A.I. Could Worsen Health Disparities*, N.Y. TIMES, (Jan. 31, 2019), <https://www.nytimes.com/2019/01/31/opinion/ai-bias-healthcare.html>.

¹²⁰ See Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1477 (1998).

of uncertainty, to use readily available facts to inform their beliefs.¹²¹ This seems like an obvious, and even rational, response to uncertainty—when we do not know something, we use the facts we have at hand. But the result of this natural instinct is that, where the availability bias is present, decisionmakers tend to overweight the representativeness of easily recallable information, and underweight more complex or distant information.¹²² This bias is amplified when individuals lack sufficient information to form independent conclusions about an issue.¹²³ In the financial world, where uncertainty is an unavoidable feature of the industry, the availability of an easily accessible recommendation, provided by an algorithm and not susceptible to ready refutation, might well cause people to rely excessively on the algorithm over independently reasoned deliberation. They might simply default to the algorithmically generated output.

Another well-studied cognitive bias is the *anchoring* effect.¹²⁴ The anchoring effect refers to people's tendency to be swayed by initial reference points, or *anchors*.¹²⁵ The classic example is in housing prices: when people are asked how much they believe a house is worth, their answer will depend heavily on how much they are told the seller is asking for it.¹²⁶ The result of setting an anchor is that decisionmakers will tend to use it as a reference point, and then make adjustments from that point, typically departing from it in only incremental and minimal ways.¹²⁷ The application to artificial financial intelligence is clear. If an algorithm provides a recommended value for an asset, or a risk score for a potential borrower, it may serve as a strong anchor, informing and biasing subsequent discussion. Even if the ultimate decisionmaker is aware of the flaws and limitations of machine learning, the very process of providing a number or calculation will affect future decisions.

Finally, and perhaps most powerfully, artificial financial intelligence may take on undue weight in financial decisions because of a phenomenon known as "herd behavior."¹²⁸ Herd behavior is a relatively intuitive concept:

¹²¹ See Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, 90 PSYCHOL. REV. 293, 295 (1983); Jolls et al., *supra* note 120, at 1477.

¹²² Jolls et al., *supra* note 120, at 1477.

¹²³ See Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 STAN. L. REV. 683, 685–87 (1999).

¹²⁴ See Marcel Kahan & Michael Klausner, *Path Dependence in Corporate Contracting: Increasing Returns, Herd Behavior and Cognitive Biases*, 74 WASH. U. L. Q. 347, 362 (1996); Gregory B. Northcraft & Margaret A. Neale, *Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions*, 39 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 84 (1987); Eric A. Zacks, *Contract Review: Cognitive Bias, Moral Hazard, and Situational Pressure*, 9 OHIO ST. ENTREPRENEURIAL BUS. L.J. 379, 394 (2015).

¹²⁵ See Kahan & Klausner, *supra* note 124, at 362.

¹²⁶ See Northcraft & Neale, *supra* note 124, at 87–90.

¹²⁷ *Id.*

¹²⁸ See David S. Scharfstein & Jeremy C. Stein, *Herd Behavior and Investment*, 80 AM. ECON. REV. 465, 465 (1990); Kahan & Klausner, *supra* note 124, at 356.

when people see other people taking an action, or refraining from doing so, they often imitate it, particularly if they have incentives to avoid diverging too much from the status quo.¹²⁹ It is not, strictly speaking, irrational to do so. Individuals in management positions are often more likely to be punished for taking an *unusual* risk or action than for taking a *common* one, even if they both turn out equally poorly.¹³⁰ After all, if they are simply doing what everyone else is doing, it is hard to argue that they misbehaved. But if they are underperforming their peers in material ways, then their reputation will likely suffer. In the world of artificial financial intelligence, it is easy to imagine that following the recommendation of a machine learning algorithm would quickly become the low-risk, reputation-protecting action. If a decisionmaker follows an artificial intelligence-based recommendation, they can deflect blame to the algorithm itself. On the other hand, if they intentionally *refuse* to accept the artificial intelligence-based recommendation, then they will have some explaining to do if their divergent opinion turns out to be incorrect.

To be sure, all of these biases are avoidable with sufficient care, diligence, and monitoring. Much of cognitive psychology is concerned with finding ways to debias individuals and correct for these errors. But the biases highlight another important feature of artificial financial intelligence: however much we may want to eliminate human interaction or interference with artificial intelligence, it ultimately creeps back in, often in unexpected places. Financial decisionmakers need to review artificial intelligence recommendations based on their own knowledge about fundamental changes in markets, changes that might not be well reflected in prior data sets. Financial engineers must review artificial intelligence for overfitting problems. But in the absence of easily understandable results, decisionmakers may simply defer to the artificial intelligence algorithm, either out of ignorance, risk aversion, or simply inertia. The problem here is not so much the algorithms themselves, but our trust in them.

C. Echo Effects

Finally, artificial financial intelligence may lead to unexpected feedback effects between competing artificial intelligence systems. This is a well-known problem in the world of high-frequency trading.¹³¹ After the so-called “Flash Crash” of 2010, in which the Dow Jones plummeted 1000 points in a matter of minutes, many observers blamed high-frequency traders

¹²⁹ See Scharfstein & Stein, *supra* note 128, at 465.

¹³⁰ See Kahan & Klausner, *supra* note 124, at 356.

¹³¹ See Yesha Yadav, *How Algorithmic Trading Undermines Efficiency in Capital Markets*, 68 VAND. L. REV. 1607, 1617–31 (2015); Pankaj K. Jain et al., *Does High-Frequency Trading Increase Systemic Risk?*, 31 J. FIN. MKTS. 1, 1 (2016); Frank J. Fabozzi et al., *High-Frequency Trading: Methodologies and Market Impact*, 19 REV. FUTURES MKTS. 7, 9–10 (2011).

for increasing the speed and volatility of markets.¹³² And while the likelihood of *echo effects* between artificial intelligence systems may be strongest in capital markets, where stock prices depend heavily on the decisions of other players within the market, they may also appear in other areas of finance, such as credit scoring, fraud prevention, and strategic advising.

One concern is that decisions made by artificial financial intelligence systems may not truly be independent of one another. In other words, if financial institutions are all deploying similar machine learning algorithms on similar data, they may reach similar results. When humans make decisions, they at least nominally are doing so on their own (even if their decisions may be affected by the decisions of their peers). Not so for artificial intelligence. It is much simpler to copy an algorithm than it is to copy a human brain. As one scholar has written in the national security context:

When China stole the blueprints and R&D data for America's F-35 fighter aircraft . . . it likely shaved years off the development timeline for a Chinese F-35 competitor—but China didn't actually acquire a modern jet fighter or the immediate capability to make one But when a country steals the code for a cyberweapons [sic], it has stolen not only the blueprints, but also the tool itself—and it can reproduce that tool at near zero-marginal cost.¹³³

If two competitors base their artificial intelligence strategies on the same algorithms, they will likely reach the same conclusions about relevant problems. This may well amplify both the speed and size of market swings, as financial institutions increasingly adopt broadly consistent views of the market.

Another concern is that, even if financial institutions do not adopt the same artificial intelligence strategies, their separate strategies may still lead to worrying feedback effects. One way this might occur is by machine learning algorithms incorporating the results of other machine learning algorithms into their data sets. For example, if one financial institution's artificial intelligence system concludes that a borrower should not be given a loan, or should only be given a loan at a certain interest rate, then that result may filter down into the data sets that are used to train other financial institutions' artificial intelligence systems. A loan rejection might, then, lead to more loan rejections. This could happen even more quickly in liquid markets like stock exchanges, where artificial intelligence strategies depend on quickly identifying and acting on changes in the market. So even if investment com-

¹³² See Yadav, *supra* note 131, at 1628–29; Andrei Kirilenko et al., *The Flash Crash: The Impact of High Frequency Trading on an Electronic Market* (May 5, 2014), https://www.cftc.gov/sites/default/files/idc/groups/public/@economicanalysis/documents/file/oce_flashcrash0314.pdf.

¹³³ See Greg Allen, *America's Plan for Stopping Cyberattacks Is Dangerously Weak*, Vox (Mar. 27, 2017), <https://www.vox.com/the-big-idea/2017/3/27/15052422/cyber-war-diplomacy-russia-us-wikileaks>.

panies are using different artificial intelligence systems, the results of each artificial intelligence system will affect the others. The overall outcome of these interactions is uncertain, but, as before, might lead to more sudden or volatile market movements.¹³⁴

Finally, another concern, and one that is written about widely in the artificial intelligence industry, is the possibility of *adversarial artificial intelligence*.¹³⁵ Adversarial artificial intelligence (or machine learning) refers to the intentional manipulation of input data in order to fool artificial intelligence systems or lead them to unintended results.¹³⁶ If an adversary or competitor knows how a given artificial intelligence system works, it can often devise strategies that lead the system to malfunction. One way to do this is by inserting *poisoned* data into the system. A good example of this is image recognition in self-driving cars. Researchers have shown that by simply putting a few black and white stickers on a stop sign, they can trick machine learning algorithms into misclassifying it 100% of the time, despite the fact that the stop sign would still be readily identifiable to any human.¹³⁷ As more of the financial world adopts artificial intelligence algorithms into their business strategies, the potential for adversarial strategies rises exponentially. For example, if a criminal learns that a bank's algorithm identifies transactions as fraudulent only when certain patterns of behavior occur, he can simply avoid the behaviors that trigger a fraud alert. Or if borrowers know the algorithm by which lenders rate their credit risk, they can simply manipulate the relevant variables to ensure they are viewed as creditworthy. Perhaps even more worrisome is the possibility that a competitor might manipulate stock markets based on knowledge of how other financial institutions' investment algorithms work. Cybersecurity is already a major worry for financial institutions. This worry only increases as algorithms perform more and more of the significant actions present in financial markets.

III. REGULATING ARTIFICIAL FINANCIAL INTELLIGENCE

Artificial financial intelligence raises a number of potential concerns, from its perpetuation of historical trends, to its effect on human decision-making, to its capacity to magnify and exacerbate volatility. But the mere *possibility* of harm, without more, does not suggest that we need new laws, or even new regulatory approaches. To justify new laws, we need to know that current laws are not working. To date, this threshold has not been met. The financial industry is one of the most heavily regulated industries in the

¹³⁴ See Tom C. W. Lin, *The New Financial Industry*, 65 ALA. L. REV. 567 (2014).

¹³⁵ See Calo, *supra* note 31, at 419–20; Tom C. W. Lin, *Financial Weapons of War*, 100 MINN. L. REV. 1377 (2016).

¹³⁶ See J. D. Tygar, *Adversarial Machine Learning*, 15 IEEE INTERNET COMPUTING 4, 4 (2011).

¹³⁷ See Kevin Eykholt et al., *Robust Physical-World Attacks on Deep Learning Visual Classification*, arXiv 1707.08945, Apr. 2018, <https://arxiv.org/abs/1707.08945>.

United States, and the kaleidoscope of regulators and regulations cover nearly every activity that a financial institution could conceivably think of doing. The Consumer Financial Protection Bureau has authority to protect consumers from unfair or abusive practices in the financial markets, and it has been proactive in monitoring new financial technologies.¹³⁸ The Securities and Exchange Commission (SEC) is tasked with protecting investors and promoting fair, efficient capital markets, and it has itself adopted machine learning methods to monitor markets.¹³⁹ The Financial Stability Oversight Council monitors the stability of the nation's financial system and responds to emerging risks, and it regularly reports on the effects of automated trading systems and financial technologies.¹⁴⁰ Artificial financial intelligence has yet to be shown to fall into gaps in this framework.

That said, this does not mean that current laws are perfectly attuned to the risks of artificial intelligence. It also does not mean that new uses of artificial intelligence will not render current financial regulation ineffective. If anything, it is hoped that this Article serves to highlight to bankers, policymakers, judges, and academics the importance of artificial financial intelligence as a field, and to highlight the need for thoughtful discussions both of how it is being used and how it should be used.

In particular, regulators should keep in mind three broad goals in adopting policies for addressing artificial financial intelligence: ensuring fairness, facilitating efficiency, and promoting stability. These goals are not particularly novel. They are the hallmarks of financial regulation, and they form the backbone of most statutory frameworks in the industry.¹⁴¹ They do, however, have different valences when applied to the world of artificial financial intelligence. This Part will discuss the ways in which financial regulators should use their existing authority to ensure that artificial financial intelligence does not undermine the free and fair functioning of financial markets.

A. Artificial Fairness

First, regulators should review artificial financial intelligence to ensure that it is being implemented in a way that treats consumers and investors fairly. *Fairness* here is meant to be an inclusive term. Fair markets are ones that are free of discrimination,¹⁴² abusive practices,¹⁴³ and fraud.¹⁴⁴ These are broad and open-ended concepts, but they are essential underpinnings of our

¹³⁸ See Van Loo, *supra* note 11, at 532–33.

¹³⁹ See Scott W. Bauguess, Acting Dir. and Acting Chief Economist, DERA, Champagne Keynote Address at OpRisk North America 2017: The Role of Big Data, Machine Learning, and AI in Assessing Risks: a Regulatory Perspective (June 21, 2017), <https://www.sec.gov/news/speech/bauguess-big-data-ai>.

¹⁴⁰ See Financial Stability Oversight Council, 2018 ANNUAL REPORT (2018), <https://home.treasury.gov/system/files/261/FSOC2018AnnualReport.pdf>.

¹⁴¹ See William Magnuson, *Financial Regulation in the Bitcoin Era*, 23 STAN. J. L. BUS. & FIN. 159 (2018).

¹⁴² See Citron & Pasquale, *supra* note 60, at 16–18.

financial system. More importantly, they have been given tangible and well-defined meaning by regulators tasked with overseeing the industry.

In terms of discrimination, regulating artificial financial intelligence will require measures to prevent algorithms from discriminating against particular disfavored or minority groups. There is already ample legislative authority to act on this front. The Equal Credit Opportunity Act prohibits lenders from discriminating against potential borrowers on the basis of race, color, religion, national origin, sex, marital status, or age.¹⁴⁵ The Fair Housing Act prohibits banks from considering similar characteristics when making mortgage decisions.¹⁴⁶ Regulation B prohibits discrimination in credit scoring systems.¹⁴⁷ The key here will be identifying when discrimination is occurring on prohibited grounds. Given the complex and, in many cases, inscrutable operations of artificial intelligence algorithms, regulators will need to develop methods for auditing financial algorithms and identifying root causes.¹⁴⁸ After all, there are many other variables that are correlated with race, sex, religion, and age, and artificial intelligence algorithms might well identify those variables and use them in impermissible ways, even if they are prevented from directly using the prohibited characteristics in decisions.

Promoting fairness in artificial financial intelligence will also require greater efforts to ensure that financial institutions properly disclose the nature of artificial intelligence risks to investors. These disclosures must be both accurate and understandable. Again, there is ample legislative authority for regulators to act here. The Securities Exchange Act of 1934 prohibits manipulation, deception and fraud in connection with the sale of securities.¹⁴⁹ Rule 10b-5 prohibits individuals from making untrue statements of material fact or omitting material facts necessary to make statements not misleading in connection with the sale of securities.¹⁵⁰ The Investment Company Act requires financial institutions to disclose ample information about investment policies.¹⁵¹ The Investment Advisers Act requires financial advisors to give suitable investment advice to consumers, taking into account the client's financial situation, investment experience, and investment objectives.¹⁵² Regulators should use this broad legislative authority to set forth

¹⁴³ See Stephen Choi, *Regulating Investors Not Issuers: A Market-Based Proposal*, 88 CAL. L. REV. 279 (2000); Jerry W. Markham, *Protecting the Institutional Investor – Jungle Predator or Shorn Lamb?*, 12 YALE J. REG. 345 (1995).

¹⁴⁴ See John C. Coffee, Jr. & Hillary A. Sale, *Redesigning the SEC: Does the Treasury Have a Better Idea?*, 95 VA. L. REV. 707, 761-62 (2009).

¹⁴⁵ 15 U.S.C. § 1691 (2012).

¹⁴⁶ 42 U.S.C. § 3605 (2012).

¹⁴⁷ Equal Credit Opportunity Act (Regulation B), 12 C.F.R. § 202.5 (2011).

¹⁴⁸ See Citron & Pasquale, *supra* note 60, at 28.

¹⁴⁹ Securities Exchange Act of 1934 §10(b), 15 U.S.C. §78j(b).

¹⁵⁰ 17 C.F.R. § 240.10b-5(b) (2008).

¹⁵¹ See Securities and Exchange Commission, Registration Form Used by Open-End Management Investment Companies, 63 FED. REG. 13,916, 13,917 (Mar. 23, 1998).

¹⁵² See Rules and Regulations, Investment Advisers Act of 1940 17 C.F.R. § 275 (2012).

clear guidelines on what sorts of artificial financial intelligence products are appropriate for investors, how those products can be marketed, and what disclosures must be made.

Regulators will also need to conduct regular reviews of artificial financial intelligence to monitor for potentially harmful effects on consumers, such as excessive interest rates or improper financial advice.¹⁵³ Currently, too little information is available about how financial institutions are using artificial intelligence, and how artificial intelligence is affecting financial markets. In 2019, Senator Elizabeth Warren wrote an open letter to financial regulators asking them how they were responding to algorithmic finance, including what they were doing to combat discrimination, how they were overseeing fair lending laws, and whether they are conducting analyses of the impact of algorithms on borrowers.¹⁵⁴ These are basic questions which should be easily answerable. The fact that they are not suggests that regulators either are not acting to the full extent of their statutory authority, or that they are not being forthcoming about their priorities and procedures.

Ensuring the fairness of artificial financial intelligence will thus require a searching analysis of the machine learning process, from algorithm design, to data gathering, to implementation procedures. Because artificial intelligence is by its very nature difficult to understand and explain, the burden on regulators will be high.¹⁵⁵ It may require regulators to hire experts on machine learning, or deploy machine learning themselves.¹⁵⁶ But the benefits are also high. If private and public sector actors can do it right, they could greatly improve the fairness of financial markets.

B. Artificial Efficiency

Second, financial regulators must be mindful of artificial intelligence's effects on the efficiency of financial markets. If artificial financial intelligence leads to the mispricing of stocks, or asset bubbles, or a reduction in competition, or higher fees, then regulators must be ready to respond. Again, there is ample authority for them to do so. Indeed, promoting efficient markets is at the core of financial regulators' mission. The SEC's stated mission is to "maintain fair, orderly and efficient markets."¹⁵⁷ The Commodity Futures Trading Commission's ("CFTC") mission is to "foster open, transpar-

¹⁵³ See Robert Bartlett et al., *Consumer-Lending Discrimination in the Fintech Era* (NBER, Working Paper No. 25942, 2019); Rory Van Loo, *The Missing Regulatory State: Monitoring Businesses in an Age of Surveillance*, 72 VAND. L. REV. 1563, 1622 (2019).

¹⁵⁴ Letter from Senator Elizabeth Warren to Federal Reserve, Federal Deposit Insurance Corporation, Office of the Comptroller of the Currency, and Consumer Financial Protection Bureau (June 10, 2019), <https://www.warren.senate.gov/imo/media/doc/2019.6.10%20Letter%20to%20Regulators%20on%20Fintech%20FINAL.pdf>.

¹⁵⁵ But not insurmountably so. See Kroll et al., *supra* note 88.

¹⁵⁶ See Bauguess, *supra* note 139.

¹⁵⁷ See 15 U.S.C. § 78c(f).

ent, competitive and financially sound markets.”¹⁵⁸ The Financial Industry Regulatory Authority’s (“FINRA”) mission is to “promote market integrity.”¹⁵⁹

In many ways, artificial intelligence might be expected to *improve* efficiency in financial markets. After all, machine learning algorithms are being deployed for the very purpose of improving the speed and accuracy of fraud detection, risk management, and investment decisions.¹⁶⁰ The great advantage of machine learning strategies is that they can find unnoticed, but material, patterns within data and then make predictions based on those patterns.¹⁶¹ Financial institutions have strong incentives to make sure that their algorithms work accurately and reliably.

But the mere fact that artificial intelligence algorithms are intended to improve efficiency, or even that financial institutions have incentives to ensure that they do so, does not mean that these algorithms will achieve their intentions, or that in achieving them, they will not create unexpected consequences elsewhere. History is littered with examples of algorithms acting unexpectedly, or causing unexpected harm.¹⁶² Regulators must therefore be attuned to the mechanisms by which artificial intelligence algorithms might create inefficiency in financial markets.

Although a thorough assessment of the causes of inefficiency within artificial financial intelligence would require a close analysis of the source code at issue, a few plausible risks are readily apparent. The first has to do with macro trends in the economy. As discussed earlier, artificial financial intelligence relies for its efficacy on data sets that, necessarily, involve historical information.¹⁶³ But if the historical trend that the information reflects is itself undesirable or unsustainable, then artificial intelligence algorithms based on it might reinforce or strengthen the trend.¹⁶⁴ They might, for example, pile into an industry that had experienced rapid growth in recent months, or rapidly exit an industry that had experienced a slump. Indeed, so-called momentum trading is a popular strategy among quantitative funds.¹⁶⁵ Again, this might be a good thing if it leads to assets more quickly reflecting their objective value. But it might equally be the result of a poorly designed algorithm.

¹⁵⁸ See COMMODITY FUTURES TRADING COMMISSION, FY 2018 AGENCY FINANCIAL REPORT 5 (Nov. 2018), <https://www.cftc.gov/system/files/2019/04/24/2018afr.pdf>.

¹⁵⁹ *Our Mission*, FINRA, <https://www.finra.org/about/our-mission>. (last visited May 14, 2020).

¹⁶⁰ See *supra* Part I.B.

¹⁶¹ See Lehr & Ohm, *supra* note 33, at 684–88.

¹⁶² See VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR (2018).

¹⁶³ See *supra* Part I.B.

¹⁶⁴ See Robin Wigglesworth, *Volatility: How “Algos” Changed the Rhythm of the Market*, FINANCIAL TIMES, Jan. 8, 2019, <https://www.ft.com/content/fdc1c064-1142-11e9-a581-4ff78404524e>.

¹⁶⁵ See Matt Prewitt, *High-Frequency Trading: Should Regulators Do More*, 19 MICH. TELECOMM. & TECH. L. REV. 131, 135 (2012).

Another form of inefficiency might arise not so much from unintended actions as intentional cooperation. Artificial intelligence algorithms increase the risk that firms will collude to raise prices, even in the absence of a formal agreement to do so.¹⁶⁶ This presents a serious antitrust risk—indeed, the Federal Trade Commission is so worried about this problem that it has held hearings on the topic.¹⁶⁷ Imagine, for example, that an artificial intelligence algorithm deployed by banks to set mortgage rates learns that every time it lowers its rate, its competitor does as well, or, potentially worse, every time it raises its rate, its competitor does as well. The interaction of these algorithms might well lead to higher prices for all. This is not merely hypothetical. As early as 1994, the Department of Justice (DOJ) found that airlines used a joint computerized booking system to set collusive ticket prices.¹⁶⁸ More recently, the DOJ has charged Amazon sellers with using algorithm-based software to engage in unlawful price fixing.¹⁶⁹

Other forms of inefficiency might be even more pernicious. Opportunistic market actors could potentially exploit knowledge of the workings of other firms' artificial intelligence algorithms to trigger unexpected behavior.¹⁷⁰ As mentioned before, researchers have worried for some time about the potential of *adversarial AI* to insert poisoned data into artificial intelligence systems in order to defeat their intended functioning.¹⁷¹ This type of strategy has obvious applications in the world of finance. Hedge funds with short positions on a stock might try to cause other firms' algorithms to trigger sell orders.¹⁷² Fraudsters might intentionally change their patterns of behavior in order to avoid fraud detection algorithms.¹⁷³ Hackers might attempt to create financial panic by spreading false information on social media platforms or news sites.¹⁷⁴ All of these sorts of behavior could introduce new and dangerous forms of inefficiency into markets, and regulators must be ready

¹⁶⁶ See Kuhn & Tadelis, *Algorithmic Collusion* (2017).

¹⁶⁷ See FEDERAL TRADE COMMISSION, THE COMPETITION AND CONSUMER PROTECTION ISSUES OF ALGORITHMS, ARTIFICIAL INTELLIGENCE, AND PREDICTIVE ANALYTICS (Federal Trade Commission 2018), <https://www.ftc.gov/news-events/audio-video/audio/algorithmic-collusion>.

¹⁶⁸ See Antonio Capobianco, *Algorithms and Collusion*, ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT 5 (2017).

¹⁶⁹ See Azriel Ezrachi & Maurice E. Stucke, *Artificial Intelligence & Collusion: When Computers Inhibit Competition*, 2017 U. ILL. L. REV. 1775, 1777 (2017).

¹⁷⁰ See, e.g., Adam Janofsky, *AI Could Make Cyberattacks More Dangerous, Harder to Detect*, WALL ST. J., Nov. 13, 2019, <https://www.wsj.com/articles/ai-could-make-cyberattacks-more-dangerous-harder-to-detect-1542128667>.

¹⁷¹ See *supra* Part II.C.

¹⁷² For empirical evidence of hedge fund manipulation of stock prices, see Itzhak Ben-David et al., *Do Hedge Funds Manipulate Stock Prices?*, 68 J. FIN. 2383 (2013).

¹⁷³ See Mary Frances Zeager et al., *Adversarial Learning in Credit Card Fraud Detection*, 2017 SYS. AND INFO. ENGINEERING DESIGN SYMP. 112, <https://ieeexplore.ieee.org/abstract/document/7937699>.

¹⁷⁴ See Michelle Cantos, *Breaking the Bank: Weakness in Financial AI Applications*, FIRE EYE (Mar. 13, 2019), <https://www.fireeye.com/blog/threat-research/2019/03/breaking-the-bank-weakness-in-financial-ai-applications.html>.

to identify and respond to them before they mature into more general market failures.

C. Artificial Risk

Finally, regulators must monitor artificial intelligence's effects on systemic risk. Observers in a wide array of areas, from national security to labor to elections, have discussed the ways in which artificial intelligence might lead to major shifts in systemic stability. Financial regulators must be similarly wary of artificial financial intelligence's effects on the stability of markets. Doing so will turn primarily on identifying the pathways through which artificial financial intelligence could create contagion among financial actors and then acting to increase the resilience of participants and systems to market shocks.

Assessing the systemic risk of artificial financial intelligence requires a careful understanding of the uses of artificial intelligence and the ways in which it is changing the nature of industry risks.¹⁷⁵ We have already seen that artificial financial intelligence might lead to more volatile markets by, for example, leading to faster and larger transaction volumes.¹⁷⁶ It might also lead to more financial bubbles, or bigger ones, if it turns out that artificial intelligence algorithms tend to pile into "fashionable" market trends.¹⁷⁷ Other risks turn on the artificial intelligence industry itself. For example, if financial institutions rely heavily on third-party providers for artificial intelligence algorithms or the data used to drive those algorithms, then those third-party providers might well become systemically important themselves, despite their non-financial focus.¹⁷⁸ Additionally, if artificial financial intelligence leads financial institutions to focus on new variables or sectors (such as social media posts or browsing activity), then those new variables could suddenly become the source of interconnectedness, and thus risk.¹⁷⁹

It just so happens that debt, one of the centerpieces for artificial financial intelligence today, is also a crucial factor in assessing the degree of systemic risk within the market.¹⁸⁰ Debt, after all, was at the heart of the last financial crisis and is one of the primary foci of many systemic risk regulators today.¹⁸¹ If financial institutions turn increasingly to artificial intelli-

¹⁷⁵ See FINANCIAL STABILITY BOARD, ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN FINANCIAL SERVICES: MARKET DEVELOPMENTS AND FINANCIAL STABILITY IMPLICATIONS, at 29–31 (2017).

¹⁷⁶ See *supra* Part II.C.

¹⁷⁷ *Id.*

¹⁷⁸ See FINANCIAL STABILITY BOARD, *supra* note 175, at 29.

¹⁷⁹ Cf. *id.* at 31 (discussing how widespread adoption of algorithms may uncover interconnections among variables that were previously thought to be unrelated).

¹⁸⁰ Hal S. Scott, *The Reduction of Systemic Risk in the United States Financial System*, 33 HARV. J. L. & PUB. POL'Y 671, 677 (2010). See generally Steven L. Schwarcz, *Systemic Risk*, 97 GEO. L. J. 193 (2008) (discussing the definition of systemic risk).

¹⁸¹ See Prasad Krishnamurthy, *Regulating Capital*, 4 HARV. BUS. L. REV. 1, 1 (2014) ("Most observers agree that the excessive debt or leverage of systemically important financial

gence to make decisions related to creditworthiness and loan activity, then the risk profile of this sector will turn heavily on the inner workings of the algorithms. As a result, systemic stability regulators will need to better understand the ways in which artificial financial intelligence affects debt markets and risk assessments.

Again, regulators have ample authority to act here. The Dodd-Frank Act created an entire department, the Financial Stability Oversight Council, to monitor systemic risk.¹⁸² Its mission is three-fold: first, to identify risks to the financial stability of the United States; second, to promote market discipline by eliminating expectations that the government will shield financial institutions from losses in the event of failure; and third, to respond to emerging threats to the stability of the financial system.¹⁸³ Artificial financial intelligence seems to fit squarely within the third category—it is an emerging risk, not a current one, to systemic stability. But in the future, if it becomes more widely adopted, it might well shift to the first category.

Given the regulatory ambit of the Financial Stability Oversight Council, and its broad-ranging mission to police systemic stability, statutory authority is unlikely to be an issue in regulating artificial financial intelligence's systemic risk. Competence and capacity, however, will be. Even within the private sector, there is a scarcity of expertise in machine learning fields.¹⁸⁴ The public sector suffers even more from this lack of experts and resources.¹⁸⁵ And without experts capable of understanding the workings of artificial intelligence algorithms, government regulators will struggle to monitor systemic risk within the sector.¹⁸⁶ One important step here would be to give regulators the ability to stress-test artificial financial intelligence systems, allowing them to provide synthetic inputs to a system and analyze the properties of the outputs. But doing so will first require regulators to gain a deep knowledge of machine learning and artificial intelligence systems, something they do not currently possess.¹⁸⁷

institutions (SIFIs) was a central reason why the housing crash of 2007–2009 led to a recession.”).

¹⁸² See Hilary J. Allen, *Putting the “Financial Stability” in Financial Stability Oversight Council*, 76 OH. ST. L. J. 1087, 1113–15 (2015).

¹⁸³ Dodd-Frank Wall Street Reform and Consumer Protection Act, 12 U.S.C. § 5322 (2012).

¹⁸⁴ See Bernard Marr, *The AI Skills Crisis and How to Close the Gap*, FORBES, June 25, 2018, <https://www.forbes.com/sites/bernardmarr/2018/06/25/the-ai-skills-crisis-and-how-to-close-the-gap/#57fa30be31f3>.

¹⁸⁵ Cf. Michael Kratsios, *Why the US Needs a Strategy for AI*, WIRED (Feb. 11, 2019) <https://www.wired.com/story/a-national-strategy-for-ai> (discussing the need for increased government focus on artificial intelligence to maintain a national competitive edge).

¹⁸⁶ See FINANCIAL STABILITY BOARD, *supra* note 175, at 34.

¹⁸⁷ Indeed, the need for greater machine learning expertise *within* government has been a mainstay of many recent proposals to improve governance of the AI industry. See Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L. J. 1147, 1216 (2017); Cary Coglianese & David Lehr, *Transparency and Algorithmic Governance*, 71 ADMIN. L. REV. 1, 30 (2019); Andrew Tutt, *An FDA for Algorithms*, 69 ADMIN. L. REV. 83, 114 (2017); Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085, 1093 (2018).

D. The Role of Self-Regulation

This Part has focused primarily on what *regulators* can do to minimize the risks and harms of artificial financial intelligence, but there is also much that financial institutions themselves can do. In fact, it is likely that self-regulation will be significantly more effective at cabining artificial intelligence's risks than regulatory enforcement actions could ever be.¹⁸⁸ Regulators are, by their very nature, outsiders. They do not know the inner workings of financial institutions nearly as well as insiders do, and they do not have the levels of expertise in machine learning that are available to the private sector. This means that, for artificial financial intelligence to fulfill its goal of making *intelligent* decisions, we will need bankers, engineers, and computer scientists to be cognizant of the broader social goals of finance. We cannot simply rely on regulators to step in to prevent financial institutions from taking risky actions—financial institutions themselves will need to be the primary gatekeepers of artificial intelligence.

What might self-regulation look like in artificial financial intelligence? Some general principles are clear. Bankers will need to consider the fairness of the artificial financial intelligence systems they develop, whether they may harm the efficiency of markets, and whether they might jeopardize systemic stability. They will need to be mindful of the data they use, the outputs they receive, the actions they take, and the monitoring systems they have in place. They may well need to go further and cooperate with other actors in the industry to ensure that their systems do not interact in dangerous ways.

But self-regulation must be more than just platitudes about corporate social responsibility.¹⁸⁹ It should also include efforts that look more like *hard law*. An illustrative example is FINRA.¹⁹⁰ FINRA is a self-regulatory body for brokerage firms formed in 2007.¹⁹¹ But despite being an industry-run

¹⁸⁸ Google itself has recognized the need for guiding principles for artificial intelligence in the private sector. In April 2018, Sergey Brin, Google's founder, wrote in a letter to shareholders that "[t]he new spring in artificial intelligence is the most significant development in computing in my lifetime" but that "such powerful tools also bring with them new questions and responsibilities," such as how they will affect employment, fairness, and safety. SERGEY BRIN, 2017 FOUNDERS' LETTER (2018), <https://abc.xyz/investor/founders-letters/2017/index.html>. In June 2018, Google CEO Sundar Pichai published an essay on Google's principles for AI, setting forth the objectives that guide Google's artificial intelligence strategies. See Sundar Pichai, *AI at Google: Our Principles* (June 7, 2018), <https://blog.google/technology/ai/ai-principles/>. In February 2019, the company went further, issuing a white paper arguing that governments need to take action to govern artificial intelligence. GOOGLE, PERSPECTIVES ON ISSUES IN AI GOVERNANCE (2018), <https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf>. But see Benjamin P. Edwards, *The Dark Side of Self-Regulation*, 85 U. CIN. L. REV. 573 (2018).

¹⁸⁹ See generally Jonathan R. Macey, *Corporate Social Responsibility: A Law & Economics Perspective*, 17 CHAPMAN L. REV. 331 (2014); John M. Conley & Cynthia A. Williams, *Engage, Embed, and Embellish: Theory Versus Practice in the Corporate Social Responsibility Movement*, 31 J. CORP. L. 1 (2005).

¹⁹⁰ See generally Andrew F. Tuch, *The Self-Regulation of Investment Bankers*, 83 GEO. WASH. L. REV. 101 (2014).

¹⁹¹ *Id.* at 104.

organization, it does much that has the look and feel of law. Broker-dealers are obligated to join the body.¹⁹² It writes standards of conduct that bind its members.¹⁹³ It has the power to discipline members for violating its rules, including the ability to exclude brokers from practicing in the industry.¹⁹⁴ This sort of organization—a self-regulatory organization with teeth—would be helpful in establishing the rules of the road for artificial financial intelligence. Industry groups might, for example, develop a set of *best practices* on cybersecurity, machine learning algorithms, and data set development. In the end, these sorts of private sector endeavors are essential to ensure that artificial financial intelligence leads to fair, efficient, and stable outcomes.

Some of this work is already taking place within the technology industry.¹⁹⁵ Computer science engineers are increasingly incorporating social goals into their machine learning guides, and they are actively looking for ways to improve artificial intelligence algorithms' ability to take into account other external values.¹⁹⁶ One prominent example is TensorFlow.¹⁹⁷ TensorFlow is a machine learning system originally developed by Google for internal use but which was eventually released to the public on an open-source basis.¹⁹⁸ Its training modules now include the following note:

As a modeler and developer, think about how this data is used and the potential benefits and harm a model's predictions can cause. A model like this could reinforce societal biases and disparities. Is a feature relevant to the problem you want to solve or will it introduce bias?¹⁹⁹

It then includes a link to a longer discussion of “machine learning fairness,” with a set of recommended readings. While the efficacy of these warnings in a technical training setting might be questionable, it provides a good example of the ways in which computer scientists are trying to incorporate broader goals into their algorithms. Financial institutions might similarly work on introducing broader situational awareness into the engineers and bankers that create artificial financial intelligence systems.

¹⁹² *Id.* at 104–05.

¹⁹³ *Id.* at 105.

¹⁹⁴ See Barbara Black, *Punishing Bad Brokers: Self-Regulation and FINRA Sanctions*, 8 BROOK. J. CORP. FIN. & COM. L. 23 (2013).

¹⁹⁵ See Gabriel Goh et al., *Satisfying Real-World Goals With Dataset Constraints*, 29 ADV. NEURAL INF. PROC. SYS. 1 (2016); Lucas Dixon et al., *Measuring and Mitigating Unintended Bias in Text Classification*, 2018 AI, ETHICS, AND SOC'Y CONF. 67 (2018).

¹⁹⁶ See, e.g., Sanders Kleinfeld, *A New Course to Teach People About Fairness in Machine Learning*, GOOGLE AI (Oct. 18, 2018), <https://www.blog.google/technology/ai/new-course-teach-people-about-fairness-machine-learning/>.

¹⁹⁷ For a discussion of the history of TensorFlow's creation and its eventual release to the public, see Daniela Hernandez, *Google's Artificial-Intelligence Tool is Offered for Free*, WALL ST. J., May 12, 2016, <https://www.wsj.com/articles/googles-open-source-parsey-mcparseface-helps-machines-understand-english-1463088180>.

¹⁹⁸ *Id.*

¹⁹⁹ *TensorFlow Data Validation Tutorial*, TENSORFLOW, https://www.tensorflow.org/tfx/tutorials/data_validation/chicago_taxi. (last visited May 14, 2020).

IV. OBJECTIONS

This Article has argued that artificial financial intelligence poses a threat to free and fair financial markets, not because of its capacity to exceed human intelligence, but because of its capacity to exacerbate human error. Because of its reliance on potentially inaccurate datasets, its biasing effect on human decisionmakers, and its unpredictable interactions with other algorithms, artificial financial intelligence could lead to harmful changes in how the financial industry works. Regulators, thus, must be vigilant in ensuring that financial institutions maintain efficient, fair, and stable markets.

These arguments, however, are not uncontroversial. In fact, they run counter to three powerful theoretical strands in artificial intelligence scholarship. I will refer to these strands as the mirror critique, the balance critique, and the precautionary critique, but I might just as accurately have referred to them as the cynical view, the optimist's view, and the pessimist's view. This Part will briefly explain what these critiques are, identify how they differ from the views asserted in this Article, and offer a few thoughts on why I find them unpersuasive.

A. *The Mirror Critique*

One standard response to artificial intelligence today is to claim that there is nothing particularly new about it.²⁰⁰ Artificial intelligence may use different algorithms from other programs, or have greater processing capacities, but at heart, it is simply another form of computer code. And like all computer code, it has flaws. More specifically, the flaws of computer code mirror those of its creators.²⁰¹ If the programmer who writes the code is risk-prone, unthoughtful, or irrational, then the algorithms will reflect that. We should not be surprised, then, that machine learning algorithms are biased, overweight certain data, or respond unpredictably to changes. So, too, do humans.²⁰²

Proponents of the mirror critique might well agree with the premise of this Article that artificial financial intelligence has flaws in relation to its problematic reliance on faulty data. But they would disagree that these flaws are unique or new, or that they require special, tailored responses from regu-

²⁰⁰ See Dan Patterson, *Why AI is Nothing New*, TECHREPUBLIC (Jan. 22, 2019), <https://www.techrepublic.com/article/why-ai-is-nothing-new/>; Richard Socher, *AI Isn't Dangerous, But Human Bias Is*, WORLD ECON. F. (Jan. 17, 2019), <https://www.weforum.org/agenda/2019/01/ai-isn-t-dangerous-but-human-bias-is/>.

²⁰¹ See Jayne Williamson-Lee, *How Machines Inherit Their Creators' Biases*, MEDIUM (July 9, 2018), <https://medium.com/coinmonks/ai-doesnt-have-to-be-conscious-to-be-harmful-385d143bd311>; Jeff Cockrell, *A.I. is Only Human*, CHI. BOOTH REV., Summer 2019, <https://review.chicagobooth.edu/economics/2019/article/ai-only-human>; Kate Crawford, *Artificial Intelligence's White Guy Problem*, N.Y. TIMES (Jun. 25, 2016).

²⁰² For a recent summary of the many ways in which humans exhibit behavioral biases as a matter of psychology, see DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2013).

lators. If the flaws are precisely the ones that we worry about with *human* decisionmakers, then there is nothing special about artificial intelligence, and regulators should simply maintain their current strategies of monitoring and sanctioning bad actions when they take place.

Before responding to this critique, it is perhaps worthwhile to note some areas of agreement. This Article agrees that the “revolutionary” nature of artificial intelligence is often exaggerated.²⁰³ This Article agrees that artificial intelligence’s flaws are related to the flaws in human psychology.²⁰⁴ This Article agrees that we should not expect artificial intelligence to eliminate human error.²⁰⁵ All of these are important observations about the nature of artificial intelligence today, and they form part of this Article’s core argument. At the same time, the mirror critique is misguided in several ways.

First, while it may be true on a trivial level that the failures of artificial intelligence can be categorized in similar ways as the failures of human intelligence, there are also important ways in which the failures are different. Most legal problems today, if described at a sufficiently high level of generality, can be shown to be the product of similar problems. The concept of negligence in tort law bears striking similarities to the concept of fiduciary duties in corporate law, but this does not mean that they are the same thing or, more importantly, that we should treat tort law and corporate law the same. Similarly, while the general categories of problem are the same between human decisionmakers and artificial intelligence algorithms, the nature and effects of these problems differ. Some of the differences are quite simple: artificial intelligence algorithms can process information much more quickly than humans can, and thus we might expect that artificial financial intelligence might lead to errors propagating and spreading at higher rates. Other differences relate to the nature of decisionmaking in these systems: as described before, *adversarial artificial intelligence* demonstrates just how susceptible these algorithms are to intentional manipulation, and in ways that that most humans would spot.²⁰⁶ Certainly, there are also ways to fool humans. But the methods are different. Thus, even if we do believe that artificial intelligence’s flaws are also human flaws, we must also recognize that these flaws look very different when an algorithm is behind them.

A second way in which the mirror critique is unpersuasive is that there are some risks from artificial financial intelligence that seem to exist in an entirely new category. One of the risks that this Article has identified is the interaction *between* human decisionmakers and artificial financial intelligence. In other words, one of the problems with financial institutions implementing machine learning strategies is that human decisionmakers may rely too heavily on the recommendations these algorithms give.²⁰⁷ Studies have

²⁰³ See *supra* Part II.B.

²⁰⁴ *Id.*

²⁰⁵ *Id.*

²⁰⁶ See *supra* Part II.C.

²⁰⁷ See *supra* Part II.B.

shown, for example, that people tend to believe that algorithms are superior to humans in making financial decisions.²⁰⁸ Thus, if artificial financial intelligence proliferates, it may create problems related to overweighting by human bankers. This suggests that artificial intelligence will introduce a new risk that simply does not exist in traditional finance.

Third, setting aside objections relating to the extent of the harm and the nature of the harm, even if we concede that the categories of problem within artificial financial intelligence and traditional finance are the same, this does not necessarily lead to the conclusion that they should both be regulated in similar ways. The mere fact that two phenomena share similar risks does not suggest that they will respond to similar regulatory interventions. Even within human communities, we know that groups can respond quite differently to government action.²⁰⁹ The imposition of fines for undesirable behavior might cause some groups to reduce the incidence of that behavior but cause other groups to increase the incidence of that behavior.²¹⁰ Similarly, crafting tailored regulation to reduce the risks of artificial financial intelligence likely requires different approaches than those traditionally used. As an initial matter, simply understanding the risks of particular artificial intelligence algorithms requires different kinds of knowledge—regulators need expertise in computer science, neural networks, and data science, things that they have not traditionally focused on.²¹¹ Similarly, the interventions needed to reduce the risks of artificial intelligence are likely quite different than traditional interventions in finance. Simple disclosure obligations may not be sufficient for machine learning algorithms—source code is notoriously complex and inscrutable.²¹² Backtesting, dynamic testing, and zero knowledge proofs (some of the proposals for analyzing machine learning algorithms) are more promising, and they require different sorts of action from regulators.²¹³ All of this is to suggest that even if artificial financial intelligence presents similar categories of risk, the regulatory actions necessary to reduce these risks are new and require specialized knowledge.

²⁰⁸ See Packin, *supra* note 118.

²⁰⁹ See, e.g., Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338 (1997); Janice Nadler, *Expressive Law, Social Norms, and Social Groups*, 42 L. & Soc. Inquiry 60 (2017).

²¹⁰ See, e.g., Uri Gneezy & Aldo Rustichini, *A Fine is a Price*, 29 J. LEGAL STUD. 1 (2000).

²¹¹ See *supra* Part III.C.

²¹² See Kroll et al., *supra* note 88, at 638 (“Perhaps the most obvious approach is to disclose a system’s source code, but this is at best a partial solution to the problem of accountability for automated decisions. The source code of computer systems is illegible to nonexperts. In fact, even experts often struggle to understand what software code will do, as inspecting source code is a very limited way of predicting how a computer program will behave.”).

²¹³ *Id.* at 668.

B. The Precautionary Critique

Proponents of the precautionary critique take a different approach.²¹⁴ They assert that the risks of artificial intelligence are not just new, they are existential. They note the dramatic improvements in artificial intelligence's capacities in just the last few years and argue that we should assume that these improvements will continue into the future.²¹⁵ As a result, artificial intelligence could potentially, in the very near future, replace many, if not most, human tasks and lead to dramatic increases in inequality.²¹⁶ Thus, these AI pessimists assert, we need wide-ranging and intrusive new regulations to assure that artificial intelligence does not cause systemic harm to our economy and society.²¹⁷ Another related strand of this critique claims that, even if artificial intelligence does not present a risk of surpassing the broad cognitive capacities of humans *currently*, it may do so in the future. And if that is the case, then regulators need to address these future harms now, not wait for them to materialize. In other words, they should adopt something akin to a precautionary principle: regulators should take anticipatory action to protect people from plausible harm even before firm empirical evidence of that harm materializes.²¹⁸

²¹⁴ Notable proponents of the precautionary critique include Elon Musk (who has claimed that artificial intelligence is "potentially more dangerous than nukes"), Stephen Hawking (who has said that "the development of full artificial intelligence could spell the end of the human race."), and, sometimes, Bill Gates (who has said that "a few decades after [machines are doing most jobs] the intelligence is strong enough to be a concern . . . and [I] don't understand why some people are not concerned."). See Cade Metz, *Mark Zuckerberg, Elon Musk and the Feud Over Killer Robots*, N.Y. TIMES, June 9, 2018, <https://www.nytimes.com/2018/06/09/technology/elon-musk-mark-zuckerberg-artificial-intelligence.html>; Rory Cellan-Jones, *Stephen Hawking Warns Artificial Intelligence Could End Mankind*, BBC, Dec. 2, 2014, <https://www.bbc.com/news/technology-30290540>; Pete Holley, *Bill Gates on Dangers of Artificial Intelligence: "I Don't Understand Why Some People are Not Concerned"*, WASH. POST, Jan. 29, 2015, https://www.washingtonpost.com/news/the-switch/wp/2015/01/28/bill-gates-on-dangers-of-artificial-intelligence-dont-understand-why-some-people-are-not-concerned/?noredirect=on&utm_term=.f4a50453a03e.

²¹⁵ See Estlund, *supra* note 87, at 258 ("[T]he prospects for job destruction are eye-opening. Robotic and digital production of goods and services, coupled with advances in AI and machine learning, is poised to take over both routine or repetitive tasks and some more advanced tasks.").

²¹⁶ See ERIK BRYNJOLFSSON & ANDREW MCAFEE, *RACE AGAINST THE MACHINE: HOW THE DIGITAL REVOLUTION IS ACCELERATING INNOVATION, DRIVING PRODUCTIVITY, AND IRREVERSIBLY TRANSFORMING EMPLOYMENT AND THE ECONOMY* (2012); Laura Tyson & Michael Spence, Exploring the Effects of Technology on Income and Wealth Inequality, in *AFTER PIKETTY: THE AGENDA FOR ECONOMICS AND INEQUALITY* 170 (Heather Boushey, J. Bradford DeLong & Marshall Steinbaum ed. 2017); FORD, *supra* note 84.

²¹⁷ See Estlund, *supra* note 87, at 254 ("This Article charts a path for reforming that body of law [that is, employment law] in the face of justified anxiety and uncertainty about the future impact of automation on jobs.").

²¹⁸ See Stephen G. Wood et al., *Wither the Precautionary Principle? An American Assessment from an Administrative Law Perspective*, 54 AM. J. COMP. L. 581, 581 (2006); Lesley Wexler, *Limiting the Precautionary Principle: Weapons Regulation in the Face of Scientific Uncertainty*, 39 U.C. DAVIS L. REV. 459, 463 (2006).

The precautionary critique would take issue with the claim in this Article that artificial intelligence's limited capacity today suggests that its primary harms will stem from human error and its capacity to exacerbate this error. It would assert that this is a short-sighted view that fails to take into account foreseeable and near-term improvements in the technology. As a result, artificial financial intelligence presents a much greater risk to our financial system, and one that requires significant new regulatory initiatives and structures.

This Article concedes that if we believe that artificial intelligence will soon, or even eventually, lead to superhuman computers that will eliminate vast swathes of the work force or, as Elon Musk fears, swarms of killer robots, then we would need much more far-reaching reforms, both in law and society. Thus, to a certain extent, the arguments in this Article presuppose that the most extreme predictions about artificial intelligence are implausible. At the same time, if artificial intelligence does evolve to this super-state, our worries will be about much more than financial regulation.

But more importantly, the precautionary critique is itself susceptible to important criticism. If we adopted a strong precautionary approach, we might well end up stifling innovation by ratcheting up the costs of technological adoption.²¹⁹ It might also be a waste of regulatory resources—if we end up regulating something that never comes about, or that evolves in such a way that the regulation is ill-adapted to it, then we will have displaced useful, near-term regulatory action with useless, long-term action.²²⁰ A wait-and-see approach can be a prudent regulatory strategy if it allows industry participants, government bodies, and the public to gather more information about the costs and benefits of the technology, or let the technology mature into a more stable equilibrium.²²¹ In the case of artificial financial intelligence, where financial institutions are still experimenting and learning, regu-

²¹⁹ See, e.g., Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCI. 698, 698 (1998); Mark A. Lemley & A. Douglas Melamed, *Missing the Forest for the Trolls*, 113 COLUM. L. REV. 2117, 2125 (2013); Anthony Falzone, *Regulation and Technology*, 36 HARV. J. L. & PUB. POL'Y 105 (2013).

²²⁰ See Sharon B. Jacobs, *The Administrative State's Passive Virtues*, 66 ADMIN. L. REV. 565 (2014); John P. Dwyer, *Overregulation*, 15 ECOLOGY L. Q. 719 (1988).

²²¹ See Stephen M. Bainbridge, *Dodd-Frank: Quack Federal Corporate Governance Round II*, 95 MINN. L. REV. 1779 (2011); Paul G. Mahoney, *The Pernicious Art of Securities Regulation*, 66 U. CHI. L. REV. 1373 (1999); Larry Ribstein, *Bubble Laws*, 40 HOUSTON L. REV. 77 (2003); Roberta Romano, *Regulating in the Dark*, in REGULATORY BREAKDOWN: THE CRISIS OF CONFIDENCE IN U.S. REGULATION 86 (Cary Coglianese ed., 2012). But see Jonathan Masur & Eric A. Posner, *Unquantified Benefits and the Problem of Regulation Under Uncertainty*, 102 CORNELL L. REV. 87 (2016); Lisa Schultz Bressman, *Judicial Review of Agency Inaction: An Arbitrariness Approach*, 79 N.Y.U. L. REV. 1657 (2004); Michael D. Sant'Ambrogio, *Agency Delays: How a Principal-Agent Approach Can Inform Judicial and Executive Branch Review of Agency Foot-Dragging*, 79 GEO. WASH. L. REV. 1381 (2011); Glen Staszewski, *The Federal Inaction Commission*, 59 EMORY L. J. 359 (2009).

lators would be ill-advised to launch a wholesale restructuring of their regulation around a still-inchoate algorithm.²²²

C. The Balance Critique

If AI cynics say that artificial intelligence is nothing new, and AI pessimists say that artificial intelligence is a dire threat, AI optimists have a different view.²²³ While acknowledging that artificial intelligence has risks, they assert that these risks are more than outweighed by the benefits.²²⁴ Moreover, there are also risks to *not* adopting artificial intelligence. On balance, the benefits of incorporating artificial intelligence into business greatly exceed the costs. Thus, to proponents of this balance critique, rather than simply defaulting to an anti-technology position (like the AI pessimists do), we need to carefully balance the pros and cons, and we will find that the pros of encouraging the growth of artificial intelligence exceed the cons of doing so.

Again, this Article largely agrees with the substance of the balance critique. When regulating a technology, we need to take into account both the costs of the technology and the benefits of it.²²⁵ As these critics rightfully point out, there are harms from *not* adopting artificial intelligence strategies if they can play a useful role in reducing fraud or increasing the efficiency of capital markets.²²⁶ Companies are likely better judges of this than govern-

²²² See Steven L. Schwarcz, *Regulating Financial Change: A Functional Approach*, 100 MINN. L. REV. 1441, 1444 (2016) ("In thinking about regulating a dynamically changing financial system, it may be more effective—or at least instructive—to focus on the system's underlying, and thus less time-dependent, economic functions than to tie regulation to any specific financial architecture.").

²²³ See, e.g., Coglianese & Lehr, *supra* note 187, at 1223 ("For administrative agencies, what will distinguish the machine-learning era is not a substitution of human judgment with some foreign and unfamiliar methodology, but rather an evolution of human judgment to incorporate fundamentally similar—albeit more accurate and often practically more useful—processes of decision making made possible by advances in statistical knowledge, data storage, and digital computing. The analysis offered here provides an antidote to visceral reactions against the use of artificial intelligence in the public sector, reactions we hope will give way to a measured optimism capable of guiding and improving the future of the administrative state."); Steve Lohr *Another Use for A.I.: Finding Millions of Unregistered Voters*, N.Y. TIMES, Nov. 5, 2018, <https://www.nytimes.com/2018/11/05/technology/unregistered-voter-rolls.html>; Angus Loten, *AI Tool Helps Companies Detect Expense Account Fraud*, WALL ST. J., Feb. 26, 2019, <https://www.wsj.com/articles/ai-tool-helps-companies-detect-expense-account-fraud-11551175200>.

²²⁴ See Vikram Mahidhar & Thomas H. Davenport, *Why Companies That Wait to Adopt AI May Never Catch Up*, HARV. BUS. REV., Dec. 6, 2018, <https://hbr.org/2018/12/why-companies-that-wait-to-adopt-ai-may-never-catch-up>.

²²⁵ See COMM. ON CAPITAL MKTS. REGULATION, A BALANCED APPROACH TO COST-BENEFIT ANALYSIS REFORM (2013); Hester Peirce, *Economic Analysis by Federal Financial Regulators*, 9 J. L. ECON. & POL'Y 569 (2013). For a discussion of the difficulties of conducting such cost-benefit analysis in financial regulation, see John C. Coates IV, *Cost-Benefit Analysis of Financial Regulation: Case Studies and Implications*, 124 YALE L. J. 882 (2015).

²²⁶ See, e.g., Shani R. Else & Francis G.X. Pileggi, *Corporate Directors Must Consider Impact of Artificial Intelligence for Effective Corporate Governance*, BUS. L. TODAY, Feb. 12,

ment officials, given their direct involvement and their indirect financial incentives. Too often, the default view of regulators is that companies should bear the burden of proof to show that the benefits of their innovations exceed the costs.²²⁷ This may be good politics, but it is bad policy.

At the same time, it is the duty of financial regulators to uphold free, fair, and stable markets. This Article has identified several ways in which artificial financial intelligence might undermine these pillars of financial regulation, and thus regulators should be vigilant about the risks. More importantly, by identifying the risks and discussing ways to reduce them, we can *shift* the balance of benefits and harms, making artificial financial intelligence more socially useful and less socially damaging. Presumably this is a goal that all groups can support.

* * * * *

So we have rejected the mirror critique, the precautionary critique, and the balance critique, and, in the process, rejected the cynics, the pessimists, and the optimists. Where does that leave us? I would say with moderation. The most radical opinions often receive the most attention today, but this does not mean that they are the most accurate. This Article, at its core, asserts that artificial financial intelligence is important, it is new, and it is changing finance, but also that we are well-prepared to deal with it. We have a comprehensive body of laws and regulations attuned to dealing with new risks to the financial system, and our regulators are properly equipped to respond to the challenges of artificial intelligence. To be sure, they may need to adopt new regulatory tools or gain expertise in new areas of finance. But this is always the case with new financial technologies.²²⁸

CONCLUSION

In the final chapter of *Tell the Machine Goodnight*, in a chance encounter on a train, Pearl runs into the man she had met earlier in the book, whom the Apricity machine had told to amputate a finger. Pearl sees that the man is indeed missing his index finger. She approaches and asks him how he is doing. He replies, “Oh, good. You know. Getting along. As one does.” The reader is left in a state of uncertainty. Was the computer right? Was he happier as a result of taking its recommendation? Would he have been better off keeping his finger intact? The reader simply does not know. But the story gets at a deeper truth about artificial intelligence: thinking about artificial intelligence is an exercise in imagination. It requires understanding how the technology works and how it is currently being used, but it also requires us

2019, <https://businesslawtoday.org/2019/02/corporate-directors-must-consider-impact-artificial-intelligence-effective-corporate-governance/>.

²²⁷ See COMM. ON CAPITAL MKTS. REGULATION, *supra* note 225, at 9.

²²⁸ See William Magnuson, *Regulating Fintech*, 71 VAND. L. REV. 1167 (2018).

to imagine how it might work in the future, and to what new uses it might be put. While artificial intelligence is coming for finance, and, to a certain extent, has already arrived, its applications, for now, are limited and experimental. Its risks derive primarily from the failure of human, not machine, intelligence. We would do best to keep that in mind.