The Regulation of AI Trading from an AI Life Cycle Perspective

ALESSIO AZZUTTI[*], WOLF-GEORG RINGE[†] & H. SIEGFRIED STIEHL[‡]

Abstract

Among innovative technologies, Artificial Intelligence (AI) is often avouched as the game changer in the provision of financial services. In this regard, the algorithmic trading domain is no exception. The impact of AI in the industry is a catalyst for transformation in the operations and the structure of capital markets. In effect, AI adds a further layer of system complexity, given its potential to alter the composition and behaviour of market actors, as well as the relationships among them.

Despite the many expected benefits, the wide use of AI could also impose new and unprecedented risks to market participants and financial stability. Specifically, owing to the potential of AI trading to disrupt markets and cause harm, global financial regulators are faced today with the daunting task of *how* best to approach its regulation in order to foster innovation and competition without sacrificing market stability and integrity. While there are common challenges, each market player faces problems unique to the context-specific use of AI. In other words, there are no one-size-fits-all solutions for regulating AI in automated trading. Rather, any effective and future-proof AI-targeting regulation should be proportionate to the particular and additional risks arising from specific applications (eg, due to the specific AI methods applied with their respective capability, validity and criticality). Therefore, financial regulators face a multi-faceted challenge. They must first define the additional risks posed by specific use cases that call for more in-depth scrutiny and, hence, identify the technical specificities that can facilitate the occurrence of those risks. Based on this assessment, they finally need to determine which AI characteristics require special regulatory treatment.

Inspired by the EU AI Act proposal, this paper examines the advantages of a 'rule-based' and 'risk-oriented' regulatory approach, combining both *ex-ante* and *ex-post* regulatory measures, that needs to be put in perspective with the 'AI life cycle'. By advocating for a multi-stakeholder engagement in AI regulatory governance, it proposes a way forward to assist financial regulators and industry players – but even actors in public education – in understanding, identifying and mitigating the risks associated with automated trading through an engineering approach for the purpose of complexity mastering.

Keywords: algorithmic trading; artificial intelligence; regulatory governance; AI life cycle
JEL Classification: G18; G28; G38; K22; K23; M15; M48; O33; O38

[*] Alessio Azzutti is a PhD Candidate in Law at Universität Hamburg and a Research Associate at the Centre for Banking & Finance Law, National University of Singapore.

[†] Wolf-Georg Ringe is a Professor of Law and Finance; Director of the Institute of Law & Economics, Universität Hamburg, and a Visiting Professor, University of Oxford.

[‡] H. Siegfried Stiehl is a Professor Emeritus at the Department of Computer Science, Universität Hamburg.

TABLE OF CONTENTS

1. Introduction

In a previous study,[1] we discussed how capital markets trading that is driven by Artificial Intelligence (AI) may result in a number of unintended consequences and even engage in market manipulation regardless of human intent. We also showed how existing legal frameworks generally fall short of ensuring capital markets' safety and integrity due to the risks created by the potential of autonomous and black box AI trading to result in market misbehaviour and harm. In consequence, we argued that global financial regulators should embrace a holistic approach and acquire interdisciplinary knowledge to advance their regulatory science towards mitigating the additional risks created by AI. This chapter attempts to go one step further by providing a novel conceptualisation of capital markets' enhanced system complexity introduced by the widespread use of AI trading techniques and strategies by financial market players. Understanding the main factors of complexity due to AI, particularly its subfield of Machine Learning (ML),[2] can help better identify risks regarding technologically-related aspects of AI-driven trading. Therefore, it should be seen as a prerequisite for global financial regulators to refine their regulatory approach and technique in response to the implications of technology sophistication within the ramification of algorithmic trading.

The aim of the present chapter is to disentangle the novel sources of capital markets' complexity due to AI that are relevant to financial regulation. It offers an initial attempt to explain how the sophistication in AI tools for financial trading may influence market functioning and calls for regulatory responses to risks associated with it. As will be discussed below, not only does autonomous and often black box ML-based trading lead to new agency problems within firms, but the widespread reliance on ML methods by market participants also contribute to the emergence of new algorithmic trading strategies and to shaping market interactions. In particular, the speed and pace of advancements in AI within the automated trading domain can expose markets to novel forms of market disruptions and unintended consequences, including unprecedented risks of market manipulation and even 'tacit' collusion by autonomous AI trading systems regardless of human intent. While being surely fascinating, technological developments in algorithmic trading due to AI also raise several challenging questions on how the law should react given the additional risks associated with them. In this context, we will investigate whether the special risks associated with ML-powered algorithmic trading are adequately captured by existing regulatory frameworks. The answer is, in part, no. And, indeed, the incremental risks associated with AI trading and their possible negative consequence to market stability and integrity urge us to consider and analyse whether those have outpaced existing risk-mitigating measures, such as established risk management practices and other control tools.[3] As our analysis will reveal, existing regulatory requirements and other legal safeguards have become obsolete, thus perilously inadequate to deal with the additional layer of complexity given the technical specificities of AI.

---

[1] Alessio Azzutti, Wolf-Georg Ringe and H. Siegfried Stiehl, 'Machine Learning, Market Manipulation, and Collusion on Capital Markets: Why the "Black Box" Matters' (2021) 73 University of Pennsylvania Journal of International Law 79.

[2] For ML, we refer to mathematical methods for computational learning from data, which does not require standard programming paradigms by human experts. ML are data-driven, empirical approaches that can assist, or even replace, humans in cognitive and financial decision-making tasks.

[3] *See* Senior Supervisors Group, 'Algorithmic Trading Briefing Note' (*Federal Reserve Bank of New York*, April 2015) <https://newyorkfed.org/medialibrary/media/newsevents/news/banking/2015/SSG-algorithmic-trading-2015.pdf> accessed 16 September 2022.

We proceed as follows. Section 2 gives a technical but high-level overview of different AI generations applied to financial trading and presents the sub-field of ML methods as the main cause of increased system complexity. It also suggests how continuous and rapid progress in computational finance research allows most sophisticated market actors to research on increasingly autonomous AI-based trading 'agents'[4]. As will be discussed, these developments raise several ethical and legal questions mainly due to the increasingly autonomous and often opaque nature of specific ML-based trading systems and strategies. With a focus on algorithmic market misbehaviour, Section 3 examines some of the emerging risks associated with AI trading, including forms of market manipulation and 'tacit' collusion. It also highlights how agency attributed to AI can hamper financial authorities' ability to regulate AI trading behaviour and, thus, ensure meaningful oversight of market conduct rules. Building on this, Section 4 pinpoints a number of shortcomings in existing legal systems, regulatory frameworks, and supervisory mechanisms in order to cope with the technical specificities of AI trading and additional risks to market stability and integrity. Therefore, adding it all up will be instrumental for Section 5 to provide guiding principles for regulating AI in algorithmic trading. Inspired by the recent EU AI Act proposal[5], this paper advances the idea of a rules-based and risk-oriented regulatory approach. It emphasises the relevance of the concept of 'AI life cycle' (derived from classical software engineering), which relates to the institutional aspects of AI governance. Eventually, Section 6 concludes.

2. ML and algorithmic trading: A primer

When analysing global capital markets under the lens of Complexity Theory, we can observe how they have been gradually evolving as an increasingly complex (eco-)system.[6] Along time and space, their system complexity has steadily expanded along different but interconnected dimensions. Whereas complexity in finance is often understood and debated in terms of growing levels of sophistication in financial modelling and products,[7] it can also refer to fast-changing and adaptive capital markets' socio-technical system, its functioning and behaviour, given the evolving and dynamic nature of economic relationships among market players, also due to technological innovation.[8]

[4] By 'trading agent' or 'AI-based agent', we generally refer to 'software agents'. Whereas there is no commonly accepted definition of 'software agents' in the literature, we here use the definition given by Franklin Stan and Art Graesser, 'It is an agent, or just a program?: A taxonomy for autonomous agents.' in Jörg P. Müller, Michael J. Wooldridge and Nicholas R. Jennings (eds), *Intelligent agents III agent theories, architectures, and languages* (Springer 1997) 21-35. The authors define an autonomous agent as '*a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future*'.

[5] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts' (21 April 2021), COM(2021) 206 final [hereinafter EU AI Act].

[6] In physics, a complex system is composed by a number of sub-units or parts that interact with each other through competitive, nonlinear collaboration leading to emergent, self-organised system behaviour, which in turn affects the behaviour of the individual parts. Applying complexity theory to observe and understand the global capital markets system, its functioning, and evolutionary dynamics is by no means a new idea. *See, eg*, Cheng-Yun Tsang, 'Rethinking Modern Financial Ecology and Its Regulatory Implications' 32(3) Banking & Finance Law Review 461.

[7] Cf Steven L. Schwarcz, 'Regulating Complexity in Financial Markets' (2009) 87 Washington University Law Review 211.

[8] See, eg, Tsang (n 6), 470-73.

Undoubtedly, the first emergence of electronic trading has contributed to increasing levels of overall system complexity.[9] It is often argued, indeed, that the ultra-fast and interconnected nature of algorithmic trading has contributed to the creation of new risks for financial stability and market integrity.[10] In response, most advanced jurisdictions have promulgated *ad hoc* regulation with the goal to mitigate these risks by counteracting the effects of flawed or malicious design, development and use of algorithmic trading systems and strategies.[11] Nevertheless, due to advancements in AI and particularly ML research, the sophistication of algorithmic trading practises, businesses, and related industries today adds an extra layer of complexity on top of a dynamic system – ie global capital markets – that naturally tend to be quite complex anyway.[12] In this trend towards enhanced system complexity, there are also some negative aspects, which mainly have a two-fold bearing. On one hand, complexity matters for those market players, such as financial institutions, in need to make a productive use of it to inform their business decisions under conditions of uncertainty while also comply with the law. On the other hand, financial regulators and supervisors must take utmost account of capital markets' complexity spurred by AI in order to safely navigate all technical and related regulatory aspects while pursuing their institutional mandates.

With these aspects in mind, the following provides a high-level overview of how a new generation of AI-driven algorithmic trading techniques and strategies, based on ML, have been shaping the relationship between complexity and financial markets. As will be argued below, developments in algorithmic trading urges global regulators to start assessing the need to improve their regulatory theory and practice in order to understand sources of system complexity and their properties (eg, stability, fragility, etc.) from an interdisciplinary perspective.

2.1. AI-induced complexity in global capital markets

The incremental and increasingly widespread use of AI techniques and tools in finance is among the most fascinating developments in financial technology (FinTech) in recent decades.[13] Today, the case of financial institutions using ML to automate tasks within their trading and investment businesses just represents one of the many ways AI has been fueling an escalating trend towards greater system complexity in our global capital markets.[14]

---

[9] See generally Andrei A. Kirilenko and Andrew W. Lo, 'Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents' (2013) 27(2) Journal of Economic Perspectives 51-72; see also Neil Johnson and others, 'Abrupt rise of new machine ecology beyond human response time' (2013) 3 Scientific Reports 2627.

[10] See, eg, Victor Galaz and Jon Pierre, 'Superconnected, Complex and Ultrafast: Governance of Hyperfunctionality in Financial Markets' (2017) 3(2) Complexity, Governance & Network 12.

[11] See Yesha Yadav, 'Algorithmic Trading and Market Regulation' in Walter Mattli (ed), *Global Algorithmic Capital Markets: High Frequency Trading, Dark Pools, and Regulatory Challenges* (Oxford University Press 2019) 232.

[12] Cf Hilbert Martin and David Darmon, 'How Complexity and Uncertainty Grew with Algorithmic Trading' (2020) 22 Entropy 499.

[13] For a historical account of different FinTech waves in the history of global finance, see Douglas W. Arner, Jànos Barberis and Ross P. Buckley, 'The Evolution of FinTech: A New Post-Crisis Paradigm?' (2015) 47 Georgetown Journal of International Law 1271.

[14] Cf The Board of The International Organization of Securities Commissions (IOSCO), 'The use of artificial intelligence and machine learning by market intermediaries and asset managers', (IOSCO, September 2021) 1-3 <https://iosco.org/library/pubdocs/pdf/IOSCOPD684.pdf> accessed 16 September 2022; and Organisation for Economic

Two main contributing, concatenated, and path-dependent factors have allowed ML research to flourish in the financial industry. On one hand, the rise of AI-driven trading responds to market players' special business needs, also with a view at best serving their clients.[15] For instance, investment firms can today apply ML to a number of innovative tasks. These include sentiment analysis based on social media content, automated text analysis of financial reports and other documents, visual chart analysis, optimising tasks within the trading cycle (eg, pre-trade analysis, trading signal generation, trade execution, etc.), up to research even on artificial trading agents. On the other, these developments have been facilitated by parallel progress in ITC products, services and solutions related to AI.[16] First, both computational power and data storage have become relatively cheap and accessible, with technological market players offering outsourcing services (eg, cloud computing, software as a service and even AI as a service).[17] Second, thanks to both market and regulatory developments, there is an exploding availability of data that algorithms can process to inform market participants' trading and investment decisions.[18] Besides traditional market data and other financial data, so-called 'alternative data' are increasingly used by firms to extract valuable insights.[19] Finally, ML methods and libraries are more easily accessible (ie via open access tools and/or software as a service),[20] and their innovative applications have contributed to a paradigm shift from 'gold old-fashioned AI'[21] (GOFAI) to ML and deep learning approaches to solve challenging financial problems and tasks.[22]

In need for more speedy information and greater analytical power to deal with uncertainty driven by the complexity of financial markets (eg, to find profitable trading opportunities), most sophisticated global market players already make a substantial use of ML in the conduct of their trading activities. Looking at the future, one could safely assume that AI/ML adoption within the domain of financial trading will continue to be widespread and even increasing. Given this trend, the following provides a high-level overview of the main ML paradigms currently utilised and researched by financial market players.

Co-operation and Development (OECD), 'Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers' (OECD 2021) 21-29 <https://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf> accessed 16 September 2022.

[15] See Financial Stability Board (FSB), 'Artificial intelligence and machine learning in financial services: Market developments and financial stability implications' (FSB, 1 November 2017) 7-10 <https://fsb.org/wp-content/uploads/P011117.pdf> accessed 16 September 2022; see also IOSCO (n 14) 6-8; and OECD (n 14) 19-20.

[16] FSB (n 15) 7-8.

[17] See generally Dave Cliff, Dan Brown, and Philip Treleaven, 'Technology Trends in the Financial Market: A 2020 Vision' (UK Government Office for Science 2010) <https://webarchive.nationalarchives.gov.uk/ukgwa/20121212135622/http://www.bis.gov.uk/assets/bispartners/foresight/docs/computer-trading/11-1222-dr3-technology-trends-in-financial-markets.pdf> accessed 16 September 2022.

[18] FSB (n 15) 9.

[19] For a reference material on 'alternative data', see Denev Alexander and Saeed Amen, *The Book of Alternative Data: A Guide for Investors, Traders, and Risk Managers* (John Wiley & Sons 2020).

[20] For instance, the 'GitHub' community (https://github.com/community) allow users to discuss, interact with and collaborate on software projects.

[21] For an introduction, see Margaret A. Boden, 'GOFAI' Chapter in Keith Frankish, Milton Keynes and William M. Ramsey (eds) *The Cambridge Handbook of Artificial Intelligence* (Cambridge University Press 2014) 89.

[22] See FSB (n 15) 10.

2.2. An 'ABC' of ML in algorithmic trading

Depending on the actual level of human involvement in the learning process (eg, data curation, model choice, hyperparameters optimisation, human-in-the-loop-and-control, etc.), the broad range of ML methods can be categorised into three main paradigms, each of which is based on mathematical frameworks.

First, in '*supervised learning*' (SL), computer trading algorithms are trained with empirical data pre-labelled by human experts in order to learn a function that maps from input to output.[23] SL methods are computational tools that can be used for, eg, statistical regression and classification.[24] For instance, a human trader can feed an SL algorithm with historical market data (eg, assets' prices, returns, volatility) to train a ML model in order to predict a given financial instrument's price movement or to classify different assets according to specific distinguishing criteria (eg, in terms of risk).[25]

Second, in '*unsupervised learning*' (UL), algorithms infer patterns (eg, regularities) from input data by identifying similar but distinctive features with limited or even no human feedback.[26] UL methods can be generally used in cluster analysis in presence of high-dimensional data.[27] For example, a human trader can use UL methods to run a pre-trade cluster analysis among a given portfolio of financial instruments according, for instance, to their likelihood to have a positive daily return in light of past observations to inform subsequent trading decision-making.[28]

Third, '*reinforcement learning*' (RL) allows for the creation of software agents[29], the knowledge and behaviour of which develop through a series of reinforcements (ie rewards and punishments) in order to reach pre-set goals defined as objective function.[30] In other words, RL agents take action to achieve a pre-defined objective (eg, maximise profits) and/or optimise a given utility function (eg, limit risks). RL agents are, hence, oriented at discovering best policy action, via trial and error, within a partially unknown and unpredictable dynamic action environment,[31] such as the globalised financial system. Their ultimate goal is to constantly refine and improve their knowledge about a specific environment to achieve the best rewards.[32] RL defines a highly heterogeneous category of computational approaches inspired by how human knowledge develops through cognitive experience and

[23] Russell Stuart and Peter Norvig, *Artificial Intelligence. A Modern Approach* (4th edn, Pearson 2022) 671.
[24] FSB (n 15) 5.
[25] Azzutti, Ringe and Stiehl (n 1) 86.
[26] Russell and Norvig (n 23) 671.
[27] Cf Ira Assent, 'Clustering high dimensional data' (2012) 2 Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 340-350.
[28] See, eg, Mansoor Momeni, Maryam Mohseni, and Mansour Soofi, 'Clustering Stock Market Companies Via K-Means Algorithms' (2015) 4(5) Kuwait Chapter of Arabian Journal of Business and Management Review <https://platform.almanhal.com/Files/Articles/75353> accessed 16 September 2022.
[29] See n 4 above.
[30] Russell and Norvig (n 23) 840-42.
[31] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction* (2nd edn, A Bradford Book 2018) 37-57.
[32] In doing so, RL agents face a constant trade-off between exploration/exploitation of new/previously successful policy action within the specific application domain's space and time. ibid 19-35.

interactions with the living environment.[33] Relatively, RL-based trading agents can resemble how human traders generally acquire market knowledge through trading experience (eg, through successes and failures).[34] Unlike both SL and UL methods,[35] RL can allow for the creation of trading agents able to explore their environment to learn the best trading strategies for profit maximisation under some sort of risks control, by also accounting for real markets' constraints (eg, liquidity and transaction costs, market impact, etc.).[36] In the field of high-frequency trading (HFT), for instance, RL methods can be used to enhance trading performance by predicting directional price movements from order book signals or optimising trade execution or smart order routing tasks.[37] As will be discussed below, RL can be used as a foundational paradigm to establish end-to-end ML approaches such as autonomous trading agents.

2.3. From 'deep learning' to autonomous trading agents

As a more recent subfield in ML, '*deep learning*' (DL) methods have substantially contributed to give fresh impetus to ML research with tremendous promise behind a variety of applications including classification of audio, image and video data but also financial trading.[38] DL methods are computational approaches based on a model of the human cortical structure in order to best analyse input data by learning on multiple abstraction levels (ie through so-called 'hidden layers').[39] Apart from increasing computational cost, DL methods can offer a number of benefits over traditional linear statistics approaches to finance, as they can learn non-linear functions to discriminate latent patterns in data.[40] As DL is also used in combination with other ML methods, a recent wave of published works in computational finance research is exploring the potential offered by different ML-based approaches in financial trading.[41] As we will see, DL makes it possible to achieve increasingly capable and autonomous algorithmic trading systems up to the point of artificial trading agents.[42] The mentioned benefits, however, come with some fundamental drawbacks. First, these methods can be

[33] ibid 314-347.
[34] Azzutti, Ringe and Stiehl (n 1) 88.
[35] Vangelis Bacoyannis and others, 'Idiosyncrasies and challenges of data driven learning in electronic trading' (NIPS 2018 Workshop on Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy, Montréal, Canada, 2018) <https://arxiv.org/abs/1811.09549> accessed 16 September 2022.
[36] See generally Thomas G. Fischer, 'Reinforcement learning in financial markets – a survey' (2018) FAU Discussion Papers in Economics, No. 12/2018 <http://hdl.handle.net/10419/183139> accessed 16 September 2022, who provides for a literature review of various RL algorithms and their application to solve financial trading task.
[37] See Michael Kearns and Yuriy Nevmyvaka, 'Machine Learning for Market Microstructure and High Frequency Trading' in David Easley, Marcos Lopez de Prado and Maureen O'Hara (eds), *High-Frequency Trading. New Realities for Traders, Markets and Regulators* (Risk Books 2013) 91-124.
[38] See generally Ahmet Murat Ozbayoglu, Mehmet Ugur Gudelek and Omer Berat Sezer, 'Deep learning for financial application: A survey' (2021) 93 Applied Soft Computing 106384 <https://doi.org/10.1016/j.asoc.2020.106384> accessed 16 September 2022.
[39] Russell and Norvig (n 23) 801-839.
[40] See, eg, Darwin Choi, Jiang Wenxi, and Zhang Chao, 'Alpha Go Everywhere: Machine Learning and International Stock Returns' (2020) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3489679> accessed 16 September 2022, which provides evidence of DL superiority over linear models in predicting stock returns.
[41] For an overview, see Ozbayoglu, Gudelek, and Sezer (n 38).
[42] See Ngoc Duy Nguyen, Thanh Nguyen and Saeid Nahavandi, 'System Design Perspective for Human-Level Agents Using Deep Reinforcement Learning: A Survey' (2017) 5 IEEE 27091-27102 <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8119919> accessed 16 September 2022.

easily prone to over-fitting.[43] When a model overfits training data, it has learnt a function that may fail to perform well when confronted with new, unseen observations and their statistical properties.[44] As will be discussed in detail in Section 4.1, DL methods – like other ML paradigms – are also well-known for their strong dependency on data quantity and quality. In addition, as another more general issue in ML research, DL methods are not free from human-induced bias (eg, inductive bias from statistically not relevant training data).[45] But, more importantly, DL-powered systems can often operate like a black box whenever they fail to provide insight into how they process data in order to assess the validity of the output. In fact, their outcome may be unpredictable, uninterpretable, and unexplainable for human experts, even in the eyes of their developers, deployers, and users.[46] As will be explained, the often opaque nature of specific ML methods and systems contributes to the growing complexity of global capital markets which, if not adequately regulated, can result in risks for financial stability and market integrity.

Within increasingly fast, interconnected, and data-rich global capital markets, steady progress in ML research paves the way for establishing incrementally capable algorithmic trading systems until the point to witness the emergence of (truly) autonomous trading agents in the foreseeable future.[47] In our previous work, we have discussed how the combination of *deep* and *reinforcement* learning (DRL) methods can first exemplify this technological trend towards full autonomy in algorithmic trading.[48] DRL-based agents have already achieved superior-to-human performance in a number of application domains, including playing chess, Go, and Atari video games.[49] Solving such strategic problems comes along with large amount of training and ample need of computational power for a DRL agent to master its large action space[50]. Apart from that, DRL has to be carefully tailored to the domain and task at hand.[51] Applying DRL to concrete use cases in complex environments, such as trading on capital markets, is a rather challenging task. Given financial markets' non-deterministic properties, the large number of players involved and their different sorts of interaction can make

[43] See, eg, Shihao Gu, Bryan Kelly and Dacheng Xiu, 'Empirical Asset Pricing via Machine Learning' (2020) 33 The Review of Financial Studies 2223.
[44] For a technical account on how to deal with issues of overfitting in financial ML, see Marcos López De Prado, *Advances in Financial Machine Learning* (1st edn, Wiley 2018).
[45] Anirudh Goyal and Yoshua Bengio, 'Inductive Biases for Deep Learning of Higher-Level Cognition' (*arxiv*, 30 November 2020) <https://arxiv.org/abs/2011.15091> accessed 16 September 2022.
[46] See generally Carlos Zednik, 'Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence' (2021) 34 Philosophy & Technology 265.
[47] Adriano Koshiyama, Nick Firoozye and Philip Treleaven, 'Algorithms in Future Capital Markets' (Proceedings of the First ACM International Conference on AI in Finance, October 2020), which also discusses 'adversarial learning', 'transfer' and 'meta-learning' as other disrupting computational learning approaches <https://dl.acm.org/doi/abs/10.1145/3383455.3422539> accessed 16 September 2022.
[48] Azzutti, Ringe and Stiehl (n 1) 90-92.
[49] See n 42.
[50] In RL-based methods, the action space refers to the set of actions to be taken by the agent to achieve a goal starting out from a pre-set initial configuration with initialisation within a specific environment. For a discussion of different RL methods applied to financial trading from the perspective of action space, see Fischer (n 36).
[51] Cf Robert Kirk and others, 'A Survey of Generalisation in Deep Reinforcement Learning' (*arxiv*, 30 January 2022) <https://arxiv.org/pdf/2111.09794.pdf> accessed 16 September 2022. The authors argue that DRL-models are not silver bullets, but require to be carefully tailored to the problem at hand, apart from data availability and quality and computational burden issues.

artificial trading agents' action space prohibitively large, thus inefficient with regard to algorithms' time and memory complexity as their input size increases.[52]

Nevertheless, a DRL-based trading agent can benefit from the advantages offered by combining DL and RL in a unique system. For instance, a trading agent could first extract meaningful trading signals (eg, based on technical indicators) from a complex and dynamic market environment via DL and, on this basis, use an RL algorithm to find the best trading policy by observing and interacting with markets.[53] However, the degree of complexity for a given DRL agent to efficiently discover the action space and, thus, safely operate in a given market context strictly depends on specific use cases.[54] To illustrate, an algorithmic agent called to optimise a given task (eg, trade execution) requires a different degree of model complexity and computational power than a trading algorithm operating across multiple venues and assets based on a statistical arbitrage strategy. In the latter case, in fact, the agent is confronted with a higher number of variables to monitor simultaneously and its action space can be substantially larger.

According to a growing number of published work in computational finance, there is at least first evidence of the benefits of DRL applications in algorithmic trading.[55] As will be discussed in Section 3, DRL-based systems can be used to establish autonomous trading agents that can discover how to manipulate markets as an optimal and rational behaviour when some conducive market conditions and factors are present. To clarify, DRL is only one possible approach towards achieving increasing levels of autonomy and sophistication in algorithmic trading. In contrast to the ML paradigms mentioned above as stand-alone approaches, one should be aware that the most innovative and promising algorithmic trading systems employed by financial institutions can combine several ML components. In effect, modern algorithmic trading systems should be better conceived as complex 'ecosystems' of algorithms, which potentially also require the involvement of the human factor (ie in order to achieve accurate, reproducible, reliable, predictable and explainable outcomes).[56] Different ML methods and techniques can be combined in complex architectures, integrated or hybrid systems, or ensemble strategies. Various agents may act in concert in multi-agent systems under different learning and collaborative/competitive approaches with the final goal, in any case, to generate *alpha* (ie extra-returns).[57] However, a key issue is to which degree such complex systems comply not only with explainability requirements, as discussed in Section 4.2, but also, eg, those of accountability and liability.

[52] Cf Jan-Alexander Posth and others, 'The Applicability of Self-Play Algorithms to Trading and Forecasting Financial Markets' (2021) 4 Frontiers in Artificial Intelligence 668465 <https://doi.org/10.3389/frai.2021.668465> accessed 16 September 2022.
[53] Cf Yang Li, Wanshang Zheng and Zibin Zheng, 'Deep Robust Reinforcement Learning for Practical Algorithmic Trading' (2019) 7 IEEE Access <https://ieeexplore.ieee.org/document/8786132> accessed 16 September 2022.
[54] See, eg, Adrian Millea, 'Deep Reinforcement Learning for Trading – Critical Survey' (2021) 6(11) Data 119-144 <https://www.mdpi.com/2306-5729/6/11/119> accessed 16 September 2022.
[55] But see Tidor-Vlad Pricope, 'Deep Reinforcement Learning in Quantitative Algorithmic Trading: A Review' (*arxiv*, 31 May 2021) <https://arxiv.org/abs/2106.00123> accessed 16 September 2022.
[56] See generally Koshiyama, Firoozye, and Treleaven (n 47).
[57] Longbing Cao, 'AI in Finance: Challenges, Techniques, and Opportunities' (2022) 55(3) ACM Computing Surveys (CSUR) 1-38 <https://arxiv.org/abs/2107.09051> accessed 16 September 2022.

Overall, in light of the state-of-the-art in ML research, emerging trends and future directions, it is reasonable to expect financial institutions to invest in and adopt increasingly capable and sophisticated AI trading tools. With the prospect of greater automation of ML-powered trading up to the point of autonomous trading agents, it will become necessary to ensure that the use of AI systems is reliable for the private interest of profit but also for the public interest of preserving the stability of financial markets.

3. AI trading 'mis-behaviour': mapping the risks

Algorithmic trading and, notably, HFT have effectively contributed to making financial trading faster, more interconnected, cheaper and, perhaps, markets overall more efficient.[58] By contrast, increasingly algorithmic-dominated markets have introduced new financial and technology-related risks.[59] With a focus on the latter, this section explores the many ways in which the increasing reliance on AI agency can exacerbate old regulatory struggles and even result in new risks to capital markets' safety and integrity through novel forms of algorithmic market manipulation and 'tacit' collusion.

Because of AI trading systems' increasing autonomy and growing complexity, it will be increasingly challenging for enforcement bodies to assess the real motives of a given AI-driven market misbehaviour or crime as those might be due to several different causes. The autonomous and often opaque nature of specific ML methods and strategies cast doubts on many aspects of financial regulation and supervision. On the one hand, financial institutions may need to compromise technological innovation with the legal obligation to comply with market conduct rules and other regulatory requirements. On the other, financial supervisors face the burdensome task of safeguarding financial stability and market integrity by detecting, investigating, and prosecuting algorithmic misbehaviour. Nevertheless, whenever AI-driven misbehaviour occurs, and manipulation passes undetected, malicious AI users expose markets to fragility, creating a number of negative externalities and harm to society. If we fail to adequately govern and regulate AI trading, global capital markets will likely be exposed to market failures and systemic instability. Some of these risks are discussed below.

3.1. AI trading 'mis-behaviour': the four basic scenarios

Unconscious, negligent, and malicious use of AI in financial trading can lead to a number of market inefficiencies and even be the cause of systemic risk, thus leading to more fragile global capital markets.[60] In short, there are four basic scenarios in which AI trading can cause market distortions, including disrupting market events (eg, flash crashes), market misbehaviours or crime.

First, an AI system can be involved in a market accident or crime as a victim itself. In this case, AI is induced to misbehave (ie 'manipulated' by a third party motivated by some personal interest). For

[58] Gaia Balp and Giovanni Strampelli, 'Preserving Capital Markets Efficiency in the High-Frequency Trading Era' (2018) 2 University of Illinois Journal of Law, Technology & Policy 349.
[59] See n 9.
[60] Jón Daníelsson, Robert Macrae and Andreas Uthemann, 'Artificial intelligence and systemic risk' (2022) 140 Journal of Banking and Finance 106290 <https://doi.org/10.1016/j.jbankfin.2021.106290> accessed 16 September 2022.

instance, a trading system can either be deceived or even hacked. Imagine a human trader attempting to lead a rival trading algorithm to make mistakes by misleading it.[61] An attacker can, in principle, influence the behaviour of an algorithm trading system by introducing manipulated information in the input data stream. As heavily reliant on data, ML algorithms are indeed susceptible to input manipulation (ie adversarial attacks).[62] Second, algorithmic trading can result in a number of unintended consequences (eg, market accidents) due to some operational failures emerging from poor human design, use, or control (eg, due to a system bug).[63] For instance, most popular events of exceptional market disturbance are often the result of some flawed design leading to software errors.[64] Other times, market accidents may be due to technical glitches in trading algorithms that lead to abnormal trading conditions, putting some market participants at an unfair disadvantage.[65] Even with more deterministic systems, the unintended consequences of algorithmic trading can often defy human comprehension to the point of complicating effective law enforcement. Third, more underhandedly, malicious actors can consciously develop and train manipulative AI systems for unfair market practices seeking to gain illicit profits that would not otherwise be available to them. It is a well-known problem that prosecuting humans for AI trading misconduct and harm can be somewhat burdensome for enforcement bodies. This is because the latter would need to ascertain the exact contribution in liability among a vast and often opaque number of possible individuals involved, to different extents, behind a given AI project.[66] Therefore, the use of increasingly powerful but sophisticated AI applications for financial trading can cause additional information asymmetry between financial firms and market conduct supervisors, thus exposing markets to risks of market failures. Finally, we have the trickiest scenario in which autonomous AI trading agents are so advanced that they can self-learn how to game market rules as an optimal and rational strategy regardless of any specific human intent.[67] As will be discussed, this is the most challenging scenario for public authorities to regulate algorithmic trading behaviour and enforce market conduct rules *vis-à-vis* increasingly autonomous and often black box AI trading systems. Using the RL research as an illustrative case, the remainder of the section discusses some fundamental practical and technical challenges that can ultimately preclude RL-based agents from discovering manipulation in an autonomous way as an optimal and rational strategy to achieve the best rewards.

3.2. ML and market manipulation

[61] Jakob Arnoldi, 'Computer Algorithms, Market Manipulation and the Institutionalization of High Frequency Trading' (2015) 47 Georgetown Journal of International Law 1271.
[62] See Elior Nehemya and others, 'Taking Over the Stock Market: Adversarial Perturbations Against Algorithmic Traders' in Yuxiao Dong, Nicolas Kourtellis, Barbara Hammer and Jose A. Lozano (eds), *Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track. ECML PKDD 2021. Lecture Notes in Computer Science Vol 12978* (Springer 2021) 221-236.
[63] See Dario Amodei and others, 'Concrete Problems in AI Safety' (*arvix*, 25 July 2016) <https://arxiv.org/pdf/1606.06565.pdf> accessed 16 September 2022.
[64] Knight Capital Americas LLC, File No. 3-15570 (*Securities and Exchange Commission*, 16 October 2013) <https://www.sec.gov/litigation/admin/2013/34-70694.pdf> accessed 16 September 2022.
[65] *Quoine Pte Ltd v B2C2 Ltd* [2020] SGCA(I) 02.
[66] Azzutti, Ringe and Stiehl (n 1) 119. See also Jonathan Tan Ming En, 'Non-Deterministic Artificial Intelligence Systems and the Future of the Law on Unilateral Mistakes in Singapore' (2022) 34 Singapore Academy of Law Journal 91, who discusses the issue from a common law perspective using Singapore's legal framework as a case study.
[67] Azzutti, Ringe and Stiehl (n 1) 118.

Thanks to progress in computational finance, algorithmic trading practices and strategies continue to evolve, and new forms of market manipulation can be expected to arise.[68] Under the current algorithmic revolution, AI offers innovative tools for malicious market participants to optimise traditional forms of manipulation and even discover new manipulative strategies. With these risks in mind, it is therefore time to reflect on how to deal with the incremental complexity introduced by AI-driven trading and associated risks to the safety and integrity of global capital markets.

It is a well-established fact that algorithmic trading can often distort markets and hamper their stability due to their innate characteristics of speed, automation, and interconnectedness.[69] There is also increasing evidence of human traders at financial institutions consciously developing and using trading algorithms to accomplish profitable but unlawful trading strategies. Indeed, the growing number of prosecuted cases against algorithmic traders manipulating markets clearly indicates financial authorities' increased focus on risks to market integrity associated with technologically-advanced forms of trading.[70] Under this perspective, ML methods and techniques can be expected to contribute further to discovering new and insidious forms of market manipulation. Though it also seems reasonable to assume that the potential of a given AI trading system to manipulate markets successfully can largely depend on a number of both micro- and macro-economic factors. Among these, for instance, one could think of a given AI user's degree of market access and market power as two main practical aspects to assess the ability to effectively manipulate markets while, at the same time, being exposed to some financial risks.[71]

In our previous work, we conceptualised some emerging risks, focusing on AI trading market manipulation and using the proprietary trading industry as a case study.[72] Among the many possible fruitful applications, we pointed out how some of the most aggressive, 'predatory' HFT strategies can benefit from the optimisation power offered by AI tools. Also some recent studies show how ML can be used to implement sophisticated manipulative HFT strategies by taking advantage of 'deep' knowledge about trading platforms' IT infrastructures and their mechanical functioning.[73] While HFT is undoubtedly an exciting field for ML research and practice, it should not be seen as the only suitable financial trading application domain. AI trading should not exclusively be lumped together with ultrafast strategies. Indeed, the potential offered by increasingly 'intelligent' trading algorithms can also be expected to develop in less speedy and less transparent markets.

---

[68] Tom W. Lin, 'The New Market Manipulation' (2017) 66 Emory Law Journal 1253; and Azzutti, Ringe and Stiehl (n 1) 97.

[69] Yadav Yesha 'How Algorithmic Trading Undermines Efficiency in Capital Markets' (2015) 68 Vanderbilt Law Review 1607.

[70] See, eg, Merritt B. Fox, Lawrence R. Glosten and Sue S. Guan, 'Spoofing and its Regulation' [2021] Columbia Business Law Review 1244.

[71] Cf Edward Leung and others, 'The Promises and Pitfalls of Machine Learning for Predicting Stock Returns' (2021) 3(2) The Journal of Financial Data Science 21-50. In comparing ML methods to more traditional approaches, the authors argue that the successful performance of any given ML trading system and strategy ultimately depends on the ability of that system/strategy '*to take risk and implement trades efficiently*'.

[72] See Azzutti, Ringe and Stiehl (n 1) 98-102.

[73] Vasilios Mavroudis, 'Market Manipulation as a Security Problem' (EuroSec '19: Proceedings of the 12th European Workshop on System Security, 29 March 2019) 1-6 <https://dl.acm.org/doi/10.1145/3301417.3312493> accessed 16 September 2022.

In Section 2, we have shown how, thanks to RL, financial institutions can nowadays research artificial trading agents that, once tasked with a trading goal, can successfully discover the best strategy with increasing autonomy by learning from experience. As losing control over trading algorithms can lead to several market failures, the risks introduced by increasingly autonomous trading agents should be thoroughly assessed by financial regulators, given their potential to hamper market stability and integrity. Still, the question of whether and to what extent human contribution is required to guide RL-based agents to discover manipulation as an optimal strategy remains unsettled. The human component may still be critical to induce AI trading agents to learn manipulative strategies. In principle, teaching AI-based agents to cheat market conduct rules can be done, eg, by training them with historical data or by providing feedback during their live operations on markets. In both cases, however, humans would provide clear guidance to AI to engage in market misconduct. Nevertheless, consciously implementing AI for manipulative or other disruptive market conduct can be challenging, given the many practical and technical limitations AI-based agents can generally face in real markets.[74] But, thanks to constant progress in ML research, some of these limitations might be overcome soon. According to a growing number of in-lab studies, in fact, there is at least first evidence of trading agents' ability to discover manipulation as an optimal strategy, regardless of any human intent.[75] However, it is not yet clear which the best learning frameworks (eg, via back-testing, simulated market environments, or real-time) are to enable ML-powered trading agents to autonomously discover and engage in market manipulation.

Overall, AI-driven trading's potential to hamper market integrity has become a 'moving' target for global financial regulators. Ascertaining whether AI trading misconduct is the result of specific human intent and strategic business planning or, instead, just a matter of negligent development and/or oversight has become increasingly challenging for enforcement authorities.[76] At the same time, global financial regulators remain cautious and vigilant regarding developments in AI within the ramification of algorithmic trading but have not taken any specific regulatory measures yet. The prospect, perhaps still most unlikely, of extremely *cynical* trading agents, which can discover manipulation while pursuing a pre-defined goal, seems more than mere imagination, thus leaving unanswered a number of troubling questions regarding the suitability of the current legal systems and regulatory framework in dealing with all such risks. In the following, we discuss how ML-based agents could even be involved in collusive scenarios.

3.3. Algorithmic trading agents and 'tacit' collusion

[74] See Azzutti, Ringe, and Stiehl (n 1) 97-98 and 101-02.
[75] Cf Enrique Martínez-Miranda, Peter McBurney and Matthew J. W. Howard, 'Learning Unfair Trading: A Market Manipulation Analysis from the Reinforcement Learning Perspective' (IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), 2016) 103-109 <https://doi.org/10.1109/EAIS.2016.7502499> accessed 16 September 2022; Takanobu Mizuta, 'Can an AI perform market manipulation at its own discretion? – A genetic algorithm learns in an artificial market simulation' (2020 IEEE Symposium Series on Computational Intelligence (SCCI), 21 May2020) 407-412 <https://doi.org/10.1109/SSCI47803.2020.9308349> accessed 16 September 2022; and Megan Shearer, Gabriel V. Rauterberg and Michael P. Wellman, 'Learning to Manipulate a Financial Benchmark' University of Michigan Law & Economic Research Paper No. 22-038 <http://dx.doi.org/10.2139/ssrn.4219227> accessed 16 September 2022.
[76] Azzutti, Ringe and Stiehl (n 1) 118-19.

Delegating cognitive agency and financial decision-making to ML-based trading agents can also expose markets to novel forms of algorithmic collusion. Notably, there are growing but still disputed concerns about competing software agents that autonomously learn to 'tacitly' collude.[77] First singled out by antitrust scholars,[78] global regulators have warned about these emerging threats more recently.[79] In light of the increasingly widespread adoption of AI trading tools by market players, it is safe to believe that unprecedented and harmful forms of algorithmic market coordination could soon materialise on global capital markets.[80] Whenever this was the case, we would urge global financial regulators to monitor and examine these technological developments with attention and scrutiny to ensure the fair and proper functioning of capital markets.

According to state-of-the-art computational economics research, two main technical barriers could preclude AI agents from autonomously forming cartels. First, algorithmic cartels rely on certain conducive market settings to establish and sustain supra-competitive equilibria. In a previous study, we discussed the relevance of a number of facilitating market factors for 'tacit' collusion between competing trading algorithms to occur (ie market transparency, trading frequency, traded assets homogeneity, market power, entry barriers, and technological sophistication).[81] To the extent that these facilitating factors are present in specific segments of capital markets, competing trading algorithms could more likely show the tendency to achieve cartel-like outcomes. However, some clarifications are warranted here. Firstly, it is reasonable to assume that a given software agent's capability to endeavour collective forms of market manipulation strictly depends on its ability to take risks and execute trading in an economically efficient way.[82] Secondly, the very level of sophistication level of AI trading systems can also play a central role in the emergence of collusion without any direct human involvement.[83] For instance, consider the case of RL-based agents. As a primary technical limitation, the capability of competing RL-based trading agents to self-learn how to coordinate behaviours autonomously can be largely constrained by the complexity and high-dimensionality of their action environment (ie the specific market application). In fact, given a potentially infinite number of variables to monitor simultaneously, RL-based agents might find it hard or even impossible to achieve dynamic strategic coordination with rivals' algorithms while also pursuing a pre-defined business objective. At least in principle, however, it seems plausible that algorithmic trading agents can be trained by humans to discover ways to coordinate with rivals (eg, by reverse engineering their algorithms) or even find autonomous ways to collude as an optimal behaviour.[84] Regarding the latter possibility, some technical and conceptual aspects remain underexplored. For instance, it is still unclear whether RL-based agents can coordinate behaviour

[77] For a literature review, see Florian E. Dorner, 'Algorithmic collusion: A critical review' (*arxiv*, 10 October 2021) <https://arxiv.org/abs/2110.04740> accessed 16 September 2022.
[78] See, eg, Ariel Ezrachi and Maurice E. Stucke, 'Artificial Intelligence & Collusion: When Computers Inhibit Competition' [2017] University of Illinois Law Review 1775.
[79] OECD, 'Algorithms and Collusion: Competition Policy in the Digital Age' (OECD 2017) <https://oecd.org/daf/competition/Algorithms-and-colllusion-competition-policy-in-the-digital-age.pdf> accessed 16 September 2022.
[80] See generally OECD (n 14) 31-32; see also Azzutti, Ringe and Stiehl (n 1) 103-04 and 112.
[81] See Azzutti, Ringe and Stiehl (n 1) 104-08 and references quoted therein.
[82] See n 71.
[83] See n 77.
[84] Azzutti, Ringe and Stiehl (n 1) 109.

without any direct communication or strategy signalling among them.[85] While it is hard to assess the exact need for communication required by RL algorithms to coordinate their trading strategies, the greater analytical capabilities offered by their combination with DL could offer a fruitful avenue for future research in this field.[86]

Notwithstanding all these technical difficulties, some recent computational research studies provide at least the first evidence, from in-lab experiments, about the ability of independent RL-based agents to coordinate behaviours in dynamic competitive games. [87] The spectacular conclusion of these studies is that 'tacit' collusion will almost inevitably arise as an optimal strategy in all those markets increasingly dominated by algorithms.[88] As these results derive from simulated and controlled environments, the reader should thus treat them with caution. No shadow of doubt, RL-based agents have shown a rather *natural* attitude to collude under very simplified constraints and stylised market settings. However, this conclusion can hardly be generalised in the context of more complex domains such as trading in capital markets. When exposed to real market configurations, RL agents may indeed find it extremely hard to scale well and master the action space's higher dimensional complexity before strategically coordinating behaviours with their rivals to achieve the best rewards.[89] Arguably, due to the increased analytical capability offered by DL methods that could help relax many of these mathematical constraints, some forms of algorithmic coordination between competing AI trading agents could become easier to accomplish soon.[90] Nevertheless, it is certainly fascinating to observe that the scientific community is employing and researching the same ML paradigms as discussed above (ie DRL methods) to assess novel and emerging risks of algorithmic collusion. Therefore, one can even assume that global capital markets will soon face new and unprecedented risks of algorithmic collusion, the likelihood and related social harm of which still remain to be concretely determined. In our view, however, the likelihood of these risks materialising will depend heavily on the exact technical capabilities and relative limitations of specific ML use cases. In particular, the materialisation of such risks will greatly depend on the ability of RL-based agents to explore market abuse in an increasingly autonomous manner through self-learning. In the following section, we discuss how the often black box nature of most sophisticated ML-based methods and strategies can result in a number of regulatory and supervisory failures.

[85] Cf Ulrich Schwalbe, 'Algorithms, Machine Learning, and Collusion' (2018) 14(4) Journal of Competition Law & Economics 568.
[86] See Ashwin Ittoo and Nicolas Petit, 'Algorithmic Pricing Agents and Tacit Collusion: A Technological Perspective' in Hervé Jacquemin and Alexandre De Streel (eds) *L'intelligence artificielle et le droit* (Larcier 2017) 241.
[87] Stephanie Assad and others, 'Autonomous algorithmic collusion: economic research and policy implications' (2021) 37 Oxford Review of Economic Policy 459.
[88] Calvano Emilio and others, 'Artificial Intelligence, Algorithmic Pricing, and Collusion' (2020) 110 American Economic Review 3267.
[89] See Ittoo and Petit (n 86); see also Malte Jeschonneck, 'Collusion among Autonomous Pricing Algorithms Utilizing Function Approximation Methods' (2021) DICE Discussion Paper No. 370 <http://hdl.handle.net/10419/240913> accessed 16 September 2022.
[90] See Matthias Hettich, 'Algorithmic Collusion: Insights from Deep Learning' (2021) CQE Working Papers 9421 <https://ideas.repec.org/p/zbw/dicedp/372.html> accessed 16 September 2022.

4. The *status quo*: uncertainties and failures

In the above, we have discussed the potential offered by AI tools for financial trading, both for good and evil. Section 2 has examined some of the ways financial institutions can benefit from adopting and using ML methods and techniques. Then, in Section 3, we have outlined how ML-powered algorithmic trading can expose markets to new risks, including novel forms of market manipulation and 'tacit' collusion by increasingly capable and autonomous trading agents (eg, thanks to RL). Therefore, in light of these emerging threats to markets' safety and integrity, this section addresses how existing financial law and regulation can fall short of dealing with the increased complexity invoked by AI trading. We first outline some technical and practical challenges RL-based agents can generally face to operate reliably in real market settings. In addition, since black-box AI trading can result in misconduct and harm, there is a need to explore the ability of the current legal systems and enforcement regimes to deal with such technology-specific risks and related issues of liability.

4.1. Challenges of the 'deep computational finance' research

Like any other ML research field, '*deep* computational finance' integrates interdisciplinary scientific knowledge and methodologies.[91] More than merely advancing financial practice, the development, training and testing of algorithmic trading systems is nowadays related to an engineering approach for the purpose of complexity mastering,[92] in the context of the 'AI life cycle'.[93] As based on a complex mix of emerging and fast-evolving technologies, the development and implementation of trustworthy AI infrastructures can indeed present significant challenges to market actors.[94] To disentangle AI-driven complexity, the following addresses some technical specificities and related practical challenges to achieve effective and reliable use of AI tools in algorithmic trading. It then discusses critical ethical and legal issues associated with the often opaque nature of specific ML methods.

Our starting point is to acknowledge that today's AI trading systems look less like a traditional trading desk and more like experimental laboratories or industrial production lines.[95] Actually, the ML training process can be thought of as an experiment in itself, whereby the ability to assess its performance and ensure explainability is key to proving successful results.[96] Generally, the specific use case and application domain heavily influence any real-life ML solution. As data-driven approaches, ML performance is highly dependent on the abundant availability and high quality of

[91] We use 'deep computational finance' as an umbrella term to generally refer to DL applications to financial trading, including those that somehow leverage the potential of RL methods (ie DRL-based trading agents).
[92] Cf Imane Bakkar and others, 'Software Validation and Artificial Intelligence – A Primer' Bank of England Staff Working Paper No. 947 (October 2021) <https://bankofengland.co.uk/-/media/boe/files/working-paper/2021/software-validation-and-artificial-intelligence-in-finance-a-primer.pdf> accessed 16 September 2022.
[93] Daswin Da Silva and Damminda Alahakoon, 'An artificial intelligence life cycle: From conception to production' (2022) 3(6) Patterns 100489 <https://doi.org/10.1016/j.patter.2022.100489> accessed 16 September 2022.
[94] Cf Kelvin Lui and Jeff Karmiol, 'AI Infrastructure Reference Architecture' IBM Systems, 87016787USEN-00 (*IBM*, June 2018) <https://www.ibm.com/downloads/cas/W1JQBNJV> accessed 16 September 2022.
[95] ibid; see also n 92.
[96] Remy Kusters and others, 'Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities' (2020) 3 Frontiers in Big Data 577974 <https://www.frontiersin.org/articles/10.3389/fdata.2020.577974/full> accessed 16 September 2022.

training data. In fact, there can be a number of technical challenges for effective and reliable applications of ML methods to solving problems of a financial trading nature, which may ultimately hamper their safe use and adoption given the need to comply with legal and regulatory requirements.[97] One such technical issue concerns the key role of data in ML research. Data have always been critical for financial forecasting and decision-making under uncertainty. In a capital market context, data collection and analysis can assist humans to draw out meaningful insights to better understand the domain of financial markets and, thus, support financial decision-making such as finding profitable trading opportunities. As data-driven approaches to empirical discovery, thus, ML methods offer innovative tools to deal with financial uncertainty.[98] Nevertheless, ML methods are well-known for their strong dependence on data quality and quantity. Training a ML algorithm heavily relies on copious amount of data available, which may also be gathered and augmented from multiple sources.[99] In addition, training data must be of utmost quality, eg, by ensuring statistical representativeness including bias independence. Moreover, data scientists must take all the reasonable steps to deal with data inconsistency, such as irregular or invalid input data, which can negatively affect any learning method.[100] When data are not sufficient in volume, they can also be augmented through synthetised data.[101] More generally, ML methods applied to trading face the mathematical problem of dealing with capital markets' nondeterministic behavioural properties, which are notoriously hard to model. The noisy, non-stationary and sometimes simply unpredictable nature of financial markets complicate the task, also for most sophisticated ML approaches, to model market dynamics given all the complexity embedded in capital markets.[102] In order to achieve its goal through an optimal sequence of actions, a trading agent not only has to explore the full action space, which is partly defined by the stream and evolution of empirical data, but also has to take into account the effects of its own trading strategy as well as other market constraints.[103] Taken all together, these aspects can lead to a number of computational limitations, e.g., due to high-dimensional spatio-temporal data.[104]

To overcome some of these technical limitations, RL-based agents can leverage the potential offered by DL methods (ie DRL methods).[105] But recurring to DL methods can also render trading agents more prone to over-fitting and model selection issues, two common risks that underpin the crucial role of model validation tasks (eg, backtesting[106]) to make sure that a given model can actually achieve its intended purpose.[107] Finally, most sophisticated ML approaches to financial trading (eg,

[97] Eg, IOSCO (n 14) 9-13, OECD (n 14) 51, Bacoyannis and others (n 35) 6, FSB (n 15) 28.
[98] See generally Cris Doloc, *Computational Intelligence in Data-Driven Trading* (Wiley 2019) 15-35; but see Hansen Kristian Bondo and Christian Borch, 'The absorption and multiplication of uncertainty in machine-learning driven finance' (2021) 72 The British Journal of Sociology 1015, arguing that the use of ML methods introduces a new, deep and penetrating form of uncertainty (ie model uncertainty).
[99] Posth and others (n 52) 3-4; see also n 19 above.
[100] OECD (n 14) 37-38.
[101] Posth and others (n 52) 3.
[102] Bacoyannis and others (n 35) 2-3.
[103] ibid 4.
[104] ibid 5-6.
[105] But see n 51.
[106] In computational finance, backtesting refers to the methods used to test a given predictive model on historical data. It allows human experts to observe how a given trading strategy would have performed should it have been employed in the past. See De Prado (n 44) 151-156.
[107] Posth and others (n 101) 3-5.

based on DL) can often invoke black box problems, thus leading to various ethical and legal issues. Indeed, as history has shown, the use of data without a sound scientific methodology and in the absence of ethical considerations can lead to unintended consequences and even harm markets and society as a whole.[108]

4.2. Black-box trading and ethical-legal dilemmas

Whenever an algorithmic trading system or strategy entails black-box algorithms or processes, several ethical-legal questions arise.[109] Notably, financial institutions using AI trading tools must ensure their compliance with market conduct rules and other regulatory requirements, such as engaging in fair and permissible activities and taking due care of their algorithmic systems to avoid disrupting markets.[110] In a financial trading context, the black box problem arises whenever human experts cannot fully understand or explain *why* and *how* their trading algorithms have reached a particular solution/decision given specific data input, thus leading to accountability gap and associated liability issues.[111] Using profitable ML-powered trading systems without completely understanding and meaningfully controlling their behaviour can thus amount to a severe form of free-riding. Financial institutions utilising AI for unfair practices or illicit purposes externalise the costs of their risky activities to other market participants while also polluting the informativeness of markets.[112]

Now, very different can be the causes leading AI systems to behave as a black box.[113] First, the phenomenon can simply manifest as a deliberate design choice made by the same users in order to keep the details of their AI-driven trading secret to others with the aim to secure a competitive edge.[114] Alternatively, opacity can arise as an unintended consequence of using sophisticated trading systems due to a lack of humans' specialised expertise in their design, development, deployment, and use.[115] Moreover, opacity can become inevitable in those complex approaches allowing for RL-based trading agents.[116] Whereas the inner working of more traditional, hence non-ML, approaches to algorithmic trading is determined by human knowledge encapsulated in algorithms and data structures, ML methods render self-learning from empirical domain data up to dynamically adjust to changing situations due to new observations. Therefore, understanding and mitigating causes of opacity in ML-driven algorithmic trading systems are paramount to ensuring the acceptance of innovative AI trading

[108] See Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown 2016).
[109] See, eg, Azzutti, Ringe and Stiehl (n 1) 119-22.
[110] For an EU perspective, see Gerald Spindler, 'Control of Algorithms in Financial Markets: The Example of High-Frequency Trading' in Martin Ebers and Susana Navas (eds), *Algorithms and the Law* (Cambridge University Press 2020) 207.
[111] See n 109 above.
[112] Cf O'Neil (n 108); see also Ekaterina Svetlova, 'AI ethics and systemic risks in finance' (2022) AI and Ethics <https://link.springer.com/article/10.1007/s43681-021-00129-1> accessed 16 September 2022.
[113] See Jenna Burrell, 'How the 'Machine' Thinks: Understanding Opacity in Machine Learning Algorithms' (2016) 3(1) Big Data & Society 1-12 <https://journals.sagepub.com/doi/pdf/10.1177/2053951715622512> accessed 16 September 2022.
[114] ibid 3.
[115] ibid 4. But see De Prado (n 44) 15-16 and 113-14, who dismisses the black box problem in the professional algorithmic trading context as a misplaced argument.
[116] ibid 4-5.

solutions and strategies from a legal and regulatory perspective. In the literature, two seemingly competing research fields are emerging to solve the challenges posed by AI transparency, interpretability and explainability.[117] On the one hand, a regulatory solution could entail higher demand for transparency[118] in AI/ML systems.[119] To this end, several commentators argue in favour of model interpretability[120] as a remedy.[121] A more radical alternative, instead, would be 'opening the black box' by direct and transparent access to the AI decision making process.[122] But this solution raises several legal and practical challenges, as it can impair innovation and competition by forcing private organisations to disclose valuable intellectual property.[123] On the other hand, as less-intrusive solutions, *ex-post* approaches aim at ensuring ML explainability to interested stakeholders.[124] 'Explainable AI' (or XAI) is indeed consolidating as a fundamental field of interdisciplinary research in ML, with a growing number of applications also within the 'deep computational finance' community.[125]

Undoubtedly, the increasing sophistication in ML applied to algorithmic trading adds another layer of complexity to understanding capital markets' functioning and behaviour. When looking at state-of-the-art '*deep* computational finance' research, we encounter a number of problems. For instance, comparative analysis of ML research is always challenging (eg, through benchmarking). The computational finance academic community has not yet come out with globally recognised standard benchmark procedures and data to allow comparisons between competing ML methods and ensemble

[117] For a discussion of the relationship between transparency, interpretability, and explainability in human-agent systems, see Avi Rosenfeld and Ariella Richardson, 'Explainability in human-agent systems' (2019) 33 Autonomous Agents and Multi-Agent Systems 673.

[118] For an account on the interplay between different dimension of transparency and trust in ML-based systems, see John Zerilli, Umang Bhatt and Adrian Weller, 'How transparency modulates trust in artificial intelligence' (2022) 3(4) Patterns 100455 <https://doi.org/10.1016/j.patter.2022.100455> accessed 16 September 2022.

[119] Cf Inioluwa Deborah Raji and Yang Jingying, 'ABOUT ML: Annotation and Benchmarking on Understanding and Transparency of Machine Learning Lifecycles' (*arxiv*, 2019) <https://arxiv.org/abs/1912.06166> accessed 16 September 2022. However, the authors note '*transparency is ... the most prevalent principle in the ... literature ... [but] the intricacy and difficulty of translating the high-level ethical ideal of transparency into concrete engineering processes and requirements has been repeatedly referenced as a major challenge*'.

[120] See Rosenfeld and Richardson (n 117), which provide a review of relevant literature and discuss different types of approaches/tools for interpretability in ML; see also Cynthia Rudin and others, 'Interpretable Machine Learning: Fundamental Principles and 10 Grand Challenges' (2022) 16 Statistics Surveys 1, who define interpretable ML as a '*model [that] obeys a domain-specific set of constraints to allow it (or its predictions, or the data) to be more easily understood by humans.*'.

[121] Cynthia Rudin, 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead' (2019) 1 Nature Machine Intelligence 206; Rudin and others (n 120), discussing a number of technical challenges in interpretable ML, also in the context of RL methods.

[122] For instance, if passed, Regulation AT would have opened algorithmic traders' source code to inspection by US financial supervisors. See Megan Woodward, 'The Need for Speed: Regulatory Approaches to High Frequency Trading in the United States and the European Union' (2017) 50 Vanderbilt Journal of Transnational Law 1359, who discusses how transparency in algorithmic trading is pursued by financial regulation in the US and the EU.

[123] See, eg, Hilary J. Allen, 'Driverless Finance' (2020) 10 Harvard Business Law Review 157; and Azzutti, Ringe and Stiehl (n 1) 130.

[124] Alexandre Heuillet, Fabien Couthouis and Natalia Díaz-Rodríguez, 'Explainability in Deep Reinforcement Learning' (2021) 214 Knowledge-Based System 106685 <https://doi.org/10.1016/j.knosys.2020.106685> accessed 16 September 2022.

[125] Cf Alejandro Barredo Arrieta and others, 'Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI' (2020) 58 Information Fusion 82-115 <https://doi.org/10.1016/j.inffus.2019.12.012> accessed 16 September 2022.

strategies (ie in their theoretical limits, accuracy, and experimental success/failure results).[126] Insights from the finance industry are even rarer and/or less informative as proprietary trading's details regarding the nature and role of the utilised empirical data, as well as the learning process itself (or hyper parameters in general), are often kept secret for obvious commercial reasons. Indeed, the overall scarcity of information generally prohibits any comprehensive and effective comparison between different ML approaches,[127] while also scientific reproducibility is also impaired.[128] As a consequence, publicly negotiated and legally binding benchmarking concepts have to be designed and implemented according to the state-of-the-art in AI/ML performance and impact characterisation (ie including open data stewardship, standardisation of performance criteria and metrics, creation of independent bodies for testing and approval, etc.).

Nevertheless, it seems realistic to envisage the proliferation of increasingly autonomous AI trading agents anytime soon. But, since this development comes with unprecedented risks, it has to be principally questioned: (*i*) whether ML methods, lacking a sound theoretical basis and subject to risks of experimental falsification, should be allowed in areas critical to society in general, and (*ii*) whether democratic legislation, eg, for proper regulation, enforcement and liability, will be able to catch up with the pace of R&D. Indeed, in the sense of the late Stephen Hawking's warning,[129] also immanent in this domain is the threat of delegating agency to AI systems and their algorithms despite knowing that, if things go wrong, we might be unable to understand and control them, but also to keep the actors behind AI accountable and liable.

4.3. Deficient regulatory regimes and uncertain law enforcement

As discussed above, ML-powered trading systems' technical specificities leave several unanswered questions regarding the integration of ML with financial theory and practice. In principle, financial institutions are required *de jure* to produce predictable, controllable and explainable trading outcomes and strategies. As such, human experts using ML methods should always be able to understand and provide effective reasoning behind their trading system's market activity to be compliant with the law. In addition, they should also be held responsible and accountable to various stakeholders (eg, their boss, clients, financial supervisors, etc.) for any negligent or improper use of algorithmic trading. Nevertheless, whenever a given AI system exhibits the black box issue, it can make any attribution of responsibility for misbehaviour and harm *de facto* impracticable. This limitation is particularly troublesome and worrisome whenever AI trading is involved in serious abuses, such as market manipulation, as enforcement bodies need to prove some relevant mental state with compelling evidence to succeed in enforcement.[130] Specifically, increasing levels of AI systems' autonomy and

[126] Cf Mark Haakman and others, 'AI lifecycle models need to be revised' (2021) 26 Empirical Software Engineering 95. The authors argue that, within the FinTech industry, the ML research has until now failed to address the challenges inherent to the ML lifecycle.
[127] Azzutti, Ringe and Stiehl (n 1) 92-93.
[128] Spyros Makridakis, Evangelos Spiliotis and Vassilios Assimakopoulos, 'Statistical and Machine Learning forecasting methods: Concerns and ways forward' (2018) 13(3) PlosONE, e0194889 <https://doi.org/10.1371/journal.pone.0194889> accessed 16 September 2022.
[129] See Rory Cellan-Jones, 'Stephen Hawking warns artificial intelligence could end mankind' (BBC, London, 2 December 2014) <https://www.bbc.com/news/technology-30290540> accessed 16 September 2022.
[130] Azzutti, Ringe and Stiehl (n 1) 118-19.

complexity can further exacerbate existing legal problems inherent to delegating financial decision-making to algorithms. With this in mind, the following gives a concise overview of the uncertain application of traditional liability rules. It also discusses some of the shortcomings of current regulatory and supervisory frameworks in dealing with the risks associated with AI trading.

4.3.1. Liability rules

Attributing liability for AI-driven accidents, market misconduct, and related harm can be highly problematic.[131] Existing legal systems are often criticised for their too vague definitions of market manipulation that, while maybe able to cope with traditional human misbehaviour, seem far from safely regulating AI trading potential for market manipulation and harm.[132] True, AI trading can represent a very different, unusual *animal* for financial law and regulation to deal with. Over time, complex, hybrid human-machine trading systems manifest and develop their 'behaviour' very differently from other law subjects (eg, human traders).[133]

Increased complexity introduced by AI can lead to liability issues for market misbehaviour and harm whenever something can go wrong or control is lost in AI trading. Established liability concepts underpinning existing market conduct rules can hardly find a safe application for cases of AI-driven forms of market manipulation. Given the automated, fast, interconnected, and increasingly sophisticated nature of algorithmic trading systems due to AI, traditional liability rules and tests – eg, intent, causation, foreseeability, negligence – will find an increasingly troublesome scope of application.[134] In most-advanced jurisdictions, enforcers need to prove a manipulator' *scienter* to count misconduct as a crime. Also, administrative sanctions are usually available to enforcement bodies, but the latter can find it hard to discern legitimate trading from unfair or abusive practices effectively.[135] A given AI misbehaviour can be due to several contingencies regardless of any specific human intent. It can merely be the outcome of counter-intuitive computational reasoning, an extrapolation of very latent patterns thanks to DL methods, or even the exploitation of strategies that human traders would not normally conceive. Moreover, the speed at which these systems operate and interact can preclude human experts from accurately foreseeing *a priori* algorithmic trading conduct and failures.[136]

In global algorithmic capital markets, the autonomous and black box nature of specific AI-driven trading strategies is inherently prone to undermining the safe application of established prohibitions of market misconduct as those rely on traditional liability rules. Even assuming the ability to detect AI trading misbehaviour, enforcement authorities would still need to find ways to ascertain the exact contribution of all possible individuals involved in designing, developing, using, and maintaining a

[131] See Section 3.1 above.
[132] See, eg, Azzutti, Ringe and Stiehl (n 1) 119-22; and Fletcher (n 70) 300-01.
[133] cf Iyad Rahwan and others, 'Machine Behaviour' (2019) 568 Nature 477-486 <https://doi.org/10.1038/s41586-019-1138-y> accessed 16 September 2022.
[134] See n 11 above.
[135] From a perspective of the US derivative markets, see Gregory Scopino, *Algo Bots and the Law: Technology, Automation, and the Regulation of Futures and Other Derivatives* (Cambridge University Press 2020) 265.
[136] Azzutti, Ringe and Stiehl (n 1) 121-22.

given AI system to impute liability. As discussed below, the current regulatory framework on the governance of algorithmic trading does not suffice to provide legal certainty and mitigation of AI-related risks.

4.3.2. Regulatory frameworks

Most advanced jurisdictions provide sector-specific regulation for the governance of algorithmic trading.[137] Whilst these sets of rules generally aim at mitigating risks to the orderly functioning of markets and their integrity, they are inherently entangled with somewhat outdated assumptions that hardly fit all the complexity of modern and future algorithmic trading.[138]

As a first concern, legal systems use their own definition of algorithmic trading. As such, this constitutes a possible source of regulatory arbitrage that obfuscates the exact scope of applicable regulation for AI-based forms of trading. As a common denominator, however, they are usually subject HFT practices to more stringent requirements due to their higher potential to distort markets and lead to systemic risk.[139] Apart from that, in all jurisdictions, algorithmic trading is not differentiated according to the specific uses of various AI technologies and applications, notwithstanding that the criticality of AI-related risks largely depends on the actual degree of complexity and capability of a given AI system.[140]

Secondly, the tasks of mitigating risks inherent to algorithmic trading are mostly left to their developers and users.[141] Financial institutions using algorithmic trading are subject to specific legal requirements. In most jurisdictions, for instance, investment firms using algorithmic trading must notify trading venues and competent authorities. They must also provide some details regarding their algorithmic trading systems and strategies (eg, strategies, risk controls, parameters relevant to execution, etc.).[142] However, existing regulatory frameworks remain *agnostic* about algorithmic trading technology itself, such as the exact degree of systems' sophistication and complexity. Regulatory focus is mostly put on algorithmic trading market behaviour and measurable outcomes (ie trading patterns).[143] Algorithmic traders' activity needs to be flagged, and firms must keep records of their trading history to support supervisory action.[144] Moreover, financial institutions must comply with a set of organisational requirements aimed at ensuring adequate investment in precautionary

[137] Kee H. Chung and Albert J. Lee, 'High-frequency Trading: Review of the Literature and Regulatory Initiatives around the World' (2016) 45(1) Asia-Pacific Journal of Financial Studies 7-23.
[138] Azzutti, Ringe and Stiehl (n 1) 122-26.
[139] Johannes Karremans and Magnus G. Scholler, 'MiFID II between European rule-making and national market surveillance: the case of high-frequency trading' in Adrienne Héritier and Magnus G. Scholler (eds), *Governing Finance in Europe* (Edward Elgar Publishing 2020) 32-51.
[140] See the EU AI Act (n 5). The EU proposal represents the first attempt to establish a risk-based regulatory framework on AI at the global level. See Lewin Schmitt, 'Mapping global AI governance: a nascent regime in a fragmented landscape' (2021) AI and Ethics 303.
[141] Joseph Lee and Lukas Schu, 'Regulation of Algorithmic Trading: Frameworks for Human Supervision and Direct Market Interventions' (2021) 33 European Business Law Review 193.
[142] ibid.
[143] Robert Seyfert, 'Algorithms as regulatory objects' (2021) Information, Communication & Society <https://doi.org/10.1080/1369118X.2021.1874035> accessed 16 September 2022.
[144] See Nathan Coombs, 'What is an algorithm? Financial regulation in the era of high-frequency trading' (2016) 45 Economy and Society 278.

measures. High *de-jure* requirements are put on the testing, validation, and deployment phases of algorithmic trading strategies to ensure reliable applications.[145] However, as recently highlighted by some regulatory reports, the conformity assessment of a given AI trading system largely remains an internal task for investment firms.[146] As a consequence, regulators and supervisors have no actual knowledge regarding how firms structure their algorithmic workflow from high-to-low levels of financial decision-making. Yet different AI applications can present very different risks. The architecture of a given algorithmic trading system usually comprises several distinct components, which can be combined in different ways. Taken together, these components form complex ecosystems of algorithms, including both software and hardware parts, some of which must be acquired by third parties. But, most importantly, the increasingly autonomous and often black box nature of specific ML approaches to financial trading can hamper financial institutions' ability to fulfil their compliance requirements.

Similar issues can arise for those regulatory responsibilities delegated to market operators. On the one hand, they serve as the first checker of algorithmic trading misbehaviour. For instance, stock exchanges usually provide simulation environments to trading firms to test algorithmic strategies' operational functioning and compatibility with their systems' capabilities and communication protocols.[147] Furthermore, trading venues have other established controls in place, such as arrangements for trading systems' operational resilience, circuit-breakers, and market surveillance facilities.[148] However, as watchdogs, market operators are well-known to face an incentive dilemma. They need to compromise rigorous screening on algorithms and market surveillance with the objectives of a profit-seeking private business under harsh competitive pressure from alternative venues.[149] Also, an algorithmic trading firm can simultaneously be the operator of a trading platform (eg, a dark pool), thus leading to a possible conflict of interests.[150] As another cause of ineffective oversight, algorithmic traders that use direct market access, an arrangement that allows firms to directly interact with a trading venue's order book by using the IT facilities of other financial institutions, can find ways to bypass some market access rules.[151] Finally, being exclusively in charge of the oversight of their own platforms, trading venues cannot provide for cross-market and cross-border market surveillance, which is the principal limit of current supervisory architecture against the ubiquity of certain AI trading strategies.[152]

As AI technology risks evolving within quite 'shadow' algorithmic trading markets, financial regulators must find innovative ways to promote a responsible culture in the design and use of AI. In

[145] Lee and Schu (n 141) 202-209; and Azzutti, Ringe and Stiehl (n 1) 123.
[146] ESMA, *MiFID II Review Report*, 28 September 2021 (ESMA 70-156-4572), 47-50.
[147] Patrick Raschner, 'Algorithms put to test: Control of algorithms in securities trading through mandatory market simulations', EBI Working Paper no. 87/2021, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3807935>.
[148] Lee and Schu (n 141) 217-221.
[149] See, eg, Yadav Yesha, 'Oversight Failure in Securities Markets' (2019) 104 Cornell Law Review 101.
[150] Danny Busch, 'MiFID II: regulating high frequency trading, other forms of algorithmic trading and direct electronic market access' (2016) 10 Law and Financial Markets Review 75.
[151] ESMA (n 146) 13-15 and 29-32; see also Alexander C. Culley, 'Does the deployment of algorithms combined with direct electronic access increase conduct risk? Evidence from the LME' (2022) Journal of Financial Regulation and Compliance (ahead-of-print), <https://doi.org/10.1108/JFRC-04-2022-0046> accessed 16 September 2022.
[152] Cf IOSCO, 'Technological Challenges to Effective Market Surveillance: Issues and Regulatory Tools – Final Report', (IOSCO April 2013) <https://www.iosco.org/library/pubdocs/pdf/IOSCOPD412.pdf> accessed 16 September 2022.

light of the above, it seems doubtful whether and effectively how self-regulation and delegation of fundamental supervisory tasks can together promote a 'good market conduct' culture within the AI trading industry.

4.3.3. Market Conduct Supervision

Financial regulation needs to be supported by solid enforcement to be effective. Enforcement, by its part, is intimately intertwined with financial supervisors' ability to detect, investigate, and prosecute cases of market misbehaviour in the best cost-efficient way, within a reasonable time, and in full respect of private interests in taking legal action.[153] In the evolving techno-sociological landscape of algorithmic-dominated global capital markets, however, growing complexity and interconnectedness introduced by AI trading can greatly contribute to complicating the oversight of markets by competent authorities.[154]

Like financial regulators, supervisors still show a rather limited understanding of AI and data technologies. Both face a substantial asymmetric information disadvantage against financial institutions developing and using AI trading tools (eg, due to the black-box). Furthermore, they face the daunting task of mitigating risks of market misbehaviour and crime through increasingly sophisticated but optimised AI-based trading strategies while being in a position of technological disadvantage.[155] Supervisory tasks and, particularly, market surveillance act mostly *ex-post*. Notwithstanding continuous efforts to achieve (almost) real-time, cross-market and cross-border surveillance thanks to enhanced cooperation among national financial supervisors, market conduct supervision is still mainly single-market oriented with delegated responsibilities upon market players themselves. Some *ex-ante* forms of supervision also exist. However, this other set of precautions mainly requires substantial investments in IT systems, testing facilities and other risk management tools and expertise by the same market players that need to be tailored to their specific use of trading technology. But being almost entirely delegated to private organisations (ie trading venues and investment firms), the effectiveness of organisational control systems is highly subjected to their willingness to contribute to socially positive outcomes. In prospecting a future dominated by artificial agents autonomously operating on markets, the direct auditing of AI-based trading systems could perhaps become more realistic thanks to the RegTech and SupTech research and practice (eg, via machine-readable regulation; XAI; etc.).[156] Finally, some progress towards AI-driven market conduct supervision can also be fairly expected to be made soon, especially when looking at the current trends and efforts being made by global financial supervisors to reduce their technological gap against market players.[157]

---

[153] See generally John Armour and others, *Principles of Financial Regulation* (OUP 2016) 577.
[154] Allen (n 123); Azzutti, Ringe and Stiehl (n 1) 125-26.
[155] FSB (n 15) 36.
[156] See generally FSB, 'The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions' (9 October 2020) 31-34 <https://fsb.org/wp-content/uploads/P091020.pdf> accessed 16 September 2022.
[157] See, eg, World Bank Group & Ministry of Foreign Affairs of the Netherlands, 'The Next Wave of Suptech Innovation: Suptech Solutions for Market Conduct Supervision' (World Bank, March 2021) <https://documents1.worldbank.org/curated/en/735871616428497205/pdf/The-Next-Wave-of-Suptech-Innovation-Suptech-Solutions-for-Market-Conduct-Supervision.pdf> accessed 16 September 2022.

5. Disentangling complexity in AI trading

In the above, we have shown that due to constant progress being made in ML research, humans can today craft trading systems that enjoy greater autonomy and are capable to self-learn from empirical data. We have also discussed how ML increasingly presents universal challenges for financial regulators to understand the implications of the interplay between AI-driven trading, financial stability and market integrity. Even if existing regulatory and supervisory strategies might have done a good job with respect to first-generation AI algorithmic trading, it has to be prevented that they will fail to account for growing complexity introduced by most-advanced AI methods and techniques (ie based on ML and particularly DL) to be embedded in a globalised financial market with unprecedented complexity in terms of both structure, processes, and technology. By remaining almost entirely agnostic regarding the precise levels of AI trading systems' technological sophistication and complexity, the current regulatory approach does not allow regulators to develop a solid understanding of AI trading technology and how different ML methods can impact investment firms' operations as well as market functioning. The resulting picture is rather problematic given the many possible causes for ineffective regulation and supervision, mostly due to issues of transparency, accountability and liability in ML-based trading. Particularly, there may be a risk that industry players will not face adequate incentives to develop and use trading machines for fair practices and that fully comply with financial law and regulation.[158] Malicious market actors may even face great incentives to use black box ML for improper purposes. In this context, financial regulators will face a dangerous trade-off: finding an optimal balance between technological neutrality, economic efficiency and market integrity. As a necessary starting point to alleviate this growing tension, global financial regulators are being called upon to gain knowledge on AI trading at a deeper level – ie ML methods and their criticality – rather than just focusing on the tip of the ice-berg when it comes to regulating the risks associated with it.

5.1. Current regulatory policy trends

There is little doubt that technological innovation within the ramification of algorithmic trading gives rise to an increasing degree of complexity characterising the structure and functioning of global capital markets. Firstly, consider model complexity of specific ML approaches. While ML methods are generally intended to assist humans to handle market uncertainty, their often black box nature can constitute a source of uncertainty multiplication in itself.[159] This is particularly the case of DL-based systems as those can alter agency problems inherent to algorithmic trading in substantial ways.[160] Moreover, the widespread adoption and ever-greater reliance on ML trading tools can have several implications for capital markets, such as shaping economic relationships and degree of interconnectedness among market players.[161] In this evolving landscape, while financial regulators appear at least increasingly vigilant on the risks and related legal issues introduced by novel relationships between AI and algorithmic trading, no major regulatory policy reform has been taken across jurisdictions yet to tackle the additional uncertainty and incremental system complexity

[158] But see, eg, EU Commission (n 140).
[159] Hilbert and Darmon (n 12); see also Hansen and Borch (n 98).
[160] Azzutti, Ringe and Stiehl (n 1) 120.
[161] See n 159 above.

induced by advances in algorithmic trading due to AI/ML.[162] In a growing number of jurisdictions, however, policy makers have adopted soft-law instruments targeting AI applications in finance. With a main focus on retail financial services, however, financial regulators aim at guiding and promoting innovation in AI towards socially beneficial and responsible outcomes while avoiding risks posed by specific AI applications in risky domains for people fundamental rights and safety.[163] Differently, algorithmic trading, and particularly the proprietary trading industry remain a somewhat less explored AI application domain by global regulators.[164] In reality, current regulatory approaches appear rather unsuccessful to provide a sound normative framework for AI-related technology to develop and flourish within the financial trading industry in the interest of society.

5.2. The rationale to regulate AI trading

While policymakers worldwide seem to be taking somewhat different approaches with respect to AI governance,[165] no significant regulatory reform has addressed the financial trading industry to date. The only major exception pertains to AI-related regulatory guidance for consumer-facing financial services.[166] However, in light of the risks discussed above, the fundamental issue facing global financial regulators is not so much whether or not to regulate AI in financial trading, but rather *how* best to approach this challenging task in order to foster innovation and competition without sacrificing market integrity and safety. Following the mantra of technological neutrality, the current regulatory paradigm is largely focused on an *ex ante* control of AI trading at market access, rather than ensuring effective due diligence and final quality check of specific AI-powered trading tools. But shaping algorithms in order to avoid market misbehaviour is a goal that is deeply tied to supervisors' ability to check for compliance, detect market misconduct and ensure smooth enforcement action. As discussed in Section 4, however, both regulators and supervisors face a yet more daunting task when confronted with the speed and pace of AI research and practice, which are indeed often left without public scrutiny.[167] In our view, solutions to this problem can either be addressed with enhanced supervisory effectiveness or by novel regulatory approaches that are better able to disentangle complexity inherent to the growing sophistication in algorithmic trading. While greater supervisory capabilities can only be welcomed given, for instance, supervisors' traditional position of

[162] For example, the EU's MiFID II has only just come into force in 2018 but, as argued by some commentators, may already show some regulatory inefficiency due to the level of sophistication of algorithmic trading. See, eg, Karremans and Scholler (n 139).

[163] For instance, Singapore has been the first world jurisdiction to adopt a soft-law instrument to promote guiding principles to develop and use AI tools in the financial services industry. For a discussion on this and other recent policy trends, see ASIFMA, 'Enabling an Efficient Regulatory Environment for AI' (Asian Securities Industry & Financial Markets, June 2021) <https://asifma.org/wp-content/uploads/2021/06/enabling-an-efficient-regulatory-environment-for-ai-report_june-2021.pdf> accessed 16 September 2022.

[164] But see, eg, IOSCO (n 14); and OECD (n 14).

[165] For a general account of democratic AI governance issues from a comparative perspective, see Ingrid Schneider, 'Democratic Governance of Digital Platforms and Artificial Intelligence? Exploring Governance Models of China, the US, the EU and Mexico' (2020) 12(1) JeDEM – eJournal of eDemocracy and Open Government 1-24 <https://jedem.org/index.php/jedem/article/view/604/487> accessed 16 September 2022.

[166] See, eg, Jemio Prenio and Jeffrey Yong, 'Humans keeping AI in check – emerging regulatory expectations in the financial sector' FSI Insights on policy implementation No 35 (Bank of International Settlements, August 2021).

[167] Cf Carla Zoe Cremer and Jess Whittlestone, 'Artificial Canaries: Early Warnings Signs for Anticipatory and Democratic Governance of AI' (2021) 6(5) International Journal of Interactive Multimedia and Artificial Intelligence 100-109.

technological disadvantage *vis-à-vis* supervised entities, we here discuss merits to improve the regulation of algorithmic trading to account for AI specificities and additional risks−moving on from the assumption that both policymakers and regulators' lack proper awareness and a clear understanding about AI trading technology.[168] In principle, any AI-targeting regulation should build upon current legal and regulatory regimes on algorithmic trading and work incrementally.[169] By doing so, market actors can be incentivised to achieve more robust governance based on existing risk management requirements and practices, thus allowing regulation to promote ethical conduct in developing and using AI without necessarily burdening it with disproportional requirements.

Regulating the use of AI in the financial services industry should be based on an engineering approach *a la* De Silva and Alahakoon, targeting software components and architectures, their design goals and criticality according to specific use cases. This is intended to empower financial regulators to conduct, either themselves or by delegation to an independent body, a pivotal role in co-producing industry developments and markets along with market players.[170] To be operational, however, an engineering approach requires some technical understanding of the different AI methods, their capabilities, and associated risks according to the specific application domain, all elements underpinning financial regulators' necessity to develop interdisciplinary knowledge at the intersection of finance, law, and computer science.[171] To start, regulators should avoid any temptation to regulate AI through one-size-fits-all solutions.[172] This option could unjustifiably punish certain lower-level risk AI applications over more-risky ones, thus also impairing technological innovation.[173] In addition, as AI is a dynamic and evolutionary concept, also partly shaped by the particular efforts and interests of private organisations developing and using AI methods and techniques, any attempt to legally define it may undermine regulation by limiting its scope of application.[174] Now, in Section 2, we have seen that AI-driven algorithmic trading can entail very different ML methods, components, and scope of application within a given algorithmic trading strategy. Therefore, regulators need to carefully account for this complexity in order to determine the exact scope of application of new AI-targeting regulation. Their goal, however, should not be the one to regulate AI *per se*, but rather to prescribe rules for best mitigating risks posed by specific ML applications to trading. This last observation suggests that regulators should follow a proportionality principle. As pointed out in Section 3, the exact nature and magnitude of the additional risks posed by AI/ML depend on the specific methods

[168] For a recent initiative aimed at empower US policymakers to make better-informed decisions on complex issues as the regulation of AI, see the 'Stanford HAI Congressional Boot Camp on Artificial Intelligence' programme, held August 8-11, 2022. The initiative was organised by the Stanford University, California, in collaboration with leaders from Silicon Valley and pioneers from civil society organisations. The full programme is available at <https://hai.stanford.edu/sites/default/files/2022-08/Boot%20Camp%20on%20AI%20Full%20Booklet%20.pdf> accessed 16 September 2022. For a European initiative targeted at financial regulators, see the 'EU-Supervisory Digital Finance Academy', officially launched on 24 October 2022 and organised by the European University Institute. The full programme of the launch event is available at <https://eusdfa.eui.eu/wp-content/uploads/2022/10/EU-SDFA-Launch-event-final-programme-1.pdf> accessed 27 October 2022. Details of the curriculum are available at <https://eusdfa.eui.eu/wp-content/uploads/2022/10/EUSDFA-Curriculum.pdf> accessed 27 October 2022.
[169] Eg, OECD (n 14) at 51; and ASIFMA (n 163) 12.
[170] Cf Da Silva and Alahakoon (n 93).
[171] Azzutti, Ringe and Stiehl (n 1) 135.
[172] See Schuett, 'Defining the Scope of AI Regulations', 2021, 4, <https://doi.org/10.48550/arXiv.1909.01095>.
[173] Cf EU Commission (n 140); ESMA (n 146) 41-42 and 46.
[174] See Schuett (n 172) 7-11.

employed, their particular combination and application scope within a given trading system, as well as the specific techno-economic environment in which AI trading is called to operate.

To disentangle this complexity, any AI-targeting regulation should be proportionate to the particular and additional risks arising from using specific ML-powered algorithmic trading's methods, architectures or strategies. A similar approach, indeed, has been taken by the EU AI Act proposal, which provides a first useful conceptual and legal framework to classify different AI applications by delineating their dimensions of risk (and complexity). To the best of the authors' knowledge, this paper is the very first treatise to evaluate the merits of a risk-based regulatory approach, such as the one proposed by the EU AI Act, using algorithmic trading, in breadth and depth, as an explanatory use case. Accordingly, regulators should first define the ML additional risks that call for a more in-depth scrutiny, identify the precise ML technical properties that can facilitate those risks to occur, and finally determine which of these properties require a special regulatory treatment.[175] In other words, regulatory response should be proportional to the risks posed by specific ML approaches and applications. A risk-based approach can assist regulators in disentangling complexity by capturing the incremental risks introduced by increasingly diverse, complex, and sophisticated ML tools for financial trading. Whereas principle-based regulatory approaches can be preferable during the initial phases of a specific technology development, they may become unable to deliver market players the right incentives in developing safe and reliable innovation at later stages.[176] Once that technology matures and evolves in complexity, while also becoming a fundamental component of an established industry, then, regulators should assess the need for a more prescriptive approach whenever new risks manifest. In our context, all this suggest that regulators should best invest in developing a clearer understanding and stronger expertise to deal with both enhanced system complexity and the additional risks posed by ML and 'deep computational finance'.

5.3. Grounding the case for a rules-based and risk-oriented regulatory approach

By acknowledging the process of developing AI-driven trading systems as an industry line of production, we are prompted to reflect on the necessity of making the best use of on an engineering approach to regulating the 'AI life cycle'.[177] Indeed, whenever regulators lack adequate technical knowledge, the fulfilment of their institutional mandate will be likely undermined. On one extreme, they may risk over-regulating. On the other, regulators' lack of understanding – or even inaction – can lead to the establishment of a business culture of algorithmic 'far west'. Both scenarios should preferably be avoided. As an ideal, instead, regulators should maintain a role of co-producers of markets, along with economic actors, in order to install a culture of a fair regulated industry, which

[175] ibid 12-17.

[176] Although being a more flexible approach, principles-based regulation can lead to uncertainty as regards to what is actually expected from firms to comply. See Cemal Karakas and Carla Stamegna, 'Financial technology (FinTech): Prospects and challenges for the EU' (European Parliament, March 2017) <https://europarl.europa.eu/RegData/etudes/BRIE/2017/599348/EPRS_BRI(2017)599348_EN.pdf> accessed 16 September 2022. For a discussion on the role of principles-based regulation in financial regulation, see Julia Black, 'The Rise, Fall and Fate of Principles Based Regulation' in: Kern Alexander and Niamh Moloney (eds), *Law Reform and Financial Markets* (Edward Elgar 2011) 3.

[177] Da Silva and Alahakoon (n 93).

is ultimately able to deliver financial stability and market integrity. To achieve this goal, not only they must properly understand technology-specific aspects of AI tools for financial trading that can be the cause of additional market uncertainty and risks, but also consider how those aspects are shaping agency problems within firms and the same business relationships among market players. As not all AI trading applications show the same criticality, financial regulators should therefore adopt a rules-based and risk-oriented regulatory approach. This could allow them to rely on a framework to classify different AI trading tools and applications depending on the perceived risks.[178]

Risk-based regulation can also allow public authorities to dispose of objective means for the use of enforcement resources and focusing their attention on relevant issues according to a well-defined risk framework.[179] The ability to categorise AI-based trading systems and components according to some risk-based metric can concretely contribute to update existing regulatory frameworks on the governance of algorithm trading. However, a risk-based categorisation of AI systems can also pose a number of challenges for its implementation.

Exactly aiming at providing further clarity and effectiveness to the proposed EU regulatory framework on AI, a recent interesting proposal suggests categorising AI systems according to a three-dimensional classification scheme based on: (*i*) the specific algorithmic methods employed in the system (ie 'Methods'); (*ii*) the capabilities to be achieved by that system (ie 'Capability'); and (*iii*) the level of criticality that can be attributed to it (ie 'Criticality').[180] In our view, this valuable approach can also provide fundamental insights in the context of our analysis. Indeed, a similar classification framework could be applied to the algorithmic trading domain. First, in terms of 'Methods', we have provided a high-level overview of how different ML paradigms can be used and combined within a given algorithmic trading system. We have also outlined some of the technical specificities and related risks of specific ML methods (eg, based on DL). Second, as regards to 'Capability', we have argued that any given AI application to financial trading can pose its own risks according to the specific use case. Thus, depending on their capabilities within the whole trading cycle and its socio-economic context, AI trading tools can present very different risk levels. Finally, in matters of 'Criticality', we have pointed out that the potential of a given AI trading system or agent to result in harm to others ultimately depends on a number of techno-economic factors (eg, the specific use case, ML model selection, complexity mastering with regard to action/state space and hyperparameters, trading efficiency, market power, cybersecurity, etc.), as well as regulatory and supervisory safeguards. As only one possible way to start disentangling AI-driven complexity, the above-mentioned framework offers the advantage of an interdisciplinary approach at the intersection of the Finance, Law, and Computer Science scientific communities.

[178] There is indeed an emerging consensus among different stakeholders in favour of risk-based approached to regulating AI applications. See, eg, Trilateral Research and EY, 'A survey of artificial intelligence risk assessment methodologies The global state of play and leading practices identified' (2021) <https://www.trilateralresearch.com/wp-content/uploads/2022/01/A-survey-of-AI-Risk-Assessment-Methodologies-full-report.pdf> accessed 16 September 2022.

[179] Julia Black and Robert Baldwin, 'Really Responsive Risk-Based Regulation' (2010) 32 Law & Policy 181.

[180] Thomas Schmid and others, 'The AI Methods, Capabilities and Criticality Grid' (2021) 35 KI – Künstliche Intelligenz 425.

Generally, a risk-based definition of algorithmic trading could provide greater legal certainty to market players, without necessarily raising compliance costs.[181] An appropriately calibrated risk-based regulation will give rise to incremental legal requirements based on the three pillars of accountability (ie human responsibility and liability), transparency (eg, regarding the ML model, computational process, and system architecture and strategy), and auditability (eg, compliance, fairness, security, safety).[182] In other words, to higher AI-driven risk levels, incremental legal and regulatory requirements should apply to financial institutions and their staff. Both *ex-ante* and *ex-post* regulatory tools can contribute to regulating AI trading and its behaviour.[183] *Ex-post* measures generally aim at ensuring human control under extreme circumstances and best guaranteeing transparency and explainability of AI outcomes and processes.[184] One often debated strategy is 'keep-the-user-in-the-loop-and-control' for safe and successful human-machine collaboration in automated AI systems.[185] In addition, thanks to the XAI research we could be soon able to develop increasingly autonomous ML-based trading that also allow humans to understand and interpret their inner functioning. Differently, *ex-ante* measures aim at regulating and shaping a given AI trading system behaviour before its deployment on markets.[186] While *ex-ante* regulation offers fascinating alternatives, their effectiveness can be impaired by human experts' knowledge and assumption about how AI can and should behave on markets and their ability to stress-test AI with proper quality data. Indeed, as data-driven approaches, ML-based trading requires to consider specific market settings, quality of input data, strategic goals, and stress scenarios in conducting testing, approval, and release. As both historical and synthetic data can fail to represent real markets behaviours, regulators face important challenges relating to the establishment of testing environments.[187] In addition, financial regulators are in a position of information asymmetry with respect to financial institutions as not being always able to determine – or just lack – access to all data relevant to testing AI. As *ex-ante* solution, however, mandating testing for AI trading systems – in order to check for their reliable and legally compliant development before their actual application on real markets – seems an inevitable regulatory policy innovation to be explored.

Ex-ante and *ex-post* regulatory solutions can contribute to disentangle the sources of enhanced system complexity introduced by AI. But as most innovative algorithmic trading can involve very different ML methods, techniques and strategies, financial regulation should provide, in a well-balanced way, for varying degrees of duties for both AI developers and users. In the following, we provide some preliminary thoughts on how regulators could *co-produce* greater standardisation of ML methods, components, and processes to render the potentially shadow business of AI algorithmic trading as a truly regulated industry in the interest of society. Greater standardisation and proactive regulation are intended together to strengthen AI governance and, at the same time, enable algorithmic trading to continue to thrive as a safe but nevertheless innovative regulated industry.

[181] But see ESMA (n 146) 45, 48, and 53-54.
[182] OECD (n 14) 56-58.
[183] See generally Allen (n 123).
[184] Azzutti, Ringe and Stiehl (n 1) 132.
[185] See Ross P. Buckley and others, 'Regulating Artificial Intelligence in Finance: Putting the Human in the Loop' (2021) 43 Sydney Law Review 43.
[186] See, eg, Allen (n 123) 196-201; Azzutti, Ringe and Stiehl (n 1) 130; see also Raschner (n 147).
[187] ibid.

5.4. Delving into the 'AI life cycle'

Delving into the 'AI life cycle' would allow financial regulators to both strengthen their knowledge and to develop further understanding about the risks associated with AI-induced system complexity. Based on a rule-based and risk-oriented regulatory approach, ML applications to financial trading could be subject to different regulatory requirements, such as testing and certification regimes. According to a recently published study by the Bank of England, software validation[188] for ML-powered systems presents today completely new challenges than more deterministic AI approaches to algorithmic trading.[189] Generally, software-related risks in ML-based systems can be grouped into three main categories: (*i*) model risk, ie the risks that a given ML algorithm cannot work as intended (eg, because of false model assumption); (*ii*) technology risk, ie operational failures (eg, due to incompatibility among system components or a system error); and (*iii*) data risk, ie all possible circumstances under which input data lack quality or availability.[190] Software validation, thus, ensures that complex software systems can work properly and meet their goals by mitigating software-related risks. However, as another recent study points out, while being perhaps able to address the peculiarities of more deterministic systems, established software validation frameworks within the financial sector perilously neglect key governance aspects of the AI lifecycle, including feasibility assessment, documentation, model monitoring and evaluation, and model risk assessment.[191] Under an engineering approach to regulating the 'AI life cycle', regulators could therefore start targeting risks arising from ML-powered algorithmic trading at the level of software validation.

More generally, based on a risk assessment, AI systems or components could be subject to more or less stricter pre-approval requirements (ie testing and certification) and other regulatory obligations (eg, on human control, re-validation, etc.). To illustrate, extremely 'high-risk' AI applications (or components) could even be prohibited when society cannot afford the related risks. For 'no-risk' or 'low-risk' AI trading tools, instead, an exemption regime could be provided.[192] As a general rule, however, regulators will apply stricter regulatory requirements with the increasing risk criticality of AI trading tools. Being able to rely on a framework as above to categorise AI trading systems according to their specific technical risks and exact use case is a more reliable tool for global financial regulators and supervisors to deal with additional complexity due to ML. While the above discussed framework is only one possible technical solutions to address the additional risks due to ML applications in such critical domains to society, the real challenge legislators/regulators face is to develop a solid and sufficient theoretical and technical knowledge for being aware of and deal with increasing levels of complexity introduced by AI and the resulting risks for the stability and integrity of global capital markets.

---

[188] In computer science, the term software validation refers to the part of the software development lifecycle aimed at verifying that a software system meets certain specifications and technical requirements in order to achieve its intended purpose.
[189] Bakkar and others (n 92).
[190] ibid.
[191] Mark Haakman and others, 'AI lifecycles models need to be revised: An exploratory study in Fintech' (2021) 26(95) Empirical Software Engineering, <https://doi.org/10.1007/s10664-021-09993-1> accessed 16 September 2022.
[192] EU Commission (n 140).

6. Conclusion

This paper has discussed increasing levels of system complexity of global capital markets and related additional risks due to constant progress in AI, particularly ML methods applied to financial trading from an interdisciplinary perspective. As industry and regulatory reports show, market players within the algorithmic trading industry are deeply interested in discovering how AI/ML methods and techniques can empower them to achieve better business performance.

Our analysis reveals that existing regulatory frameworks targeting the governance of algorithmic trading do not provide a stable ground for a secure implementation of AI trading tools. Although AI/ML methods and techniques have been taken up by the financial industry, we cannot safely conclude that the current financial regulation has been consistently updated to reflect the incremental risks in algorithmic trading due to rapid technological innovation. In effect, existing regulatory regimes do not help shed much light on the actual level of AI systems' sophistication and complexity achieved by financial institutions, as well as their criticality. By following a technology-agnostic regulatory approach, financial authorities have not been much engaged in developing necessary skills to grasp the inner functioning of algorithmic trading from an engineering-based AI/ML life cycle perspective. Being their focus mostly dedicated to the oversight of market behaviours, financial institutions have been enjoying a quite ample freedom in the use of trading technology. To counter this ever-widening lack of knowledge, financial regulators and supervisors need today to acquire more granular information to identify, assess, and analyse the risks associated with a given algorithmic trading system or strategy. In this chapter we argued that, in order to overcome current deficiencies and limitations of regulatory regimes applicable to advanced AI/ML in financial trading, the risk-based approach of the EU AI Act has to be put in perspective with regards to the novel 'AI life cycle' concept (derived from classical software engineering) which in turn is related to institutional AI governance issues. As a consequence, we advocate careful consideration of and in-depth discussion in the political, legal, institutional, industrial and academic arena on how to bring bear an effective synergy of the three currently rather disconnected regulatory layers of legislation/regulation, AI governance, and 'AI life cycle' for the sake of a future-proof financial trading sector.

Leaving financial institutions' research and implementation of AI trading tools as 'black boxes' put financial regulators and authorities in a position of asymmetric information disadvantage that can ultimately hamper their ability to conduct their institutional mandates effectively. As of today, there are no clear-cut solutions to the problem emerging on the horizon. At least, however, it has become clear that embracing an interdisciplinary approach is necessary to achieve effective AI regulation, thus, governance. To this end, there is a need to empower global financial authorities to fully understand the complexity inherent to AI/ML in finance. One way is to make policymakers and regulators better understand what is actually behind the 'AI life cycle' of financial institutions. Whenever financial authorities lack the resources and expertise to take upon this task alone, they may seek assistance or even delegate it to a specialised, newly established public authority or some other independent body (eg, independent trustee, sounding board). In other words, there are two main priorities to consider within the international fora of financial regulators in the near future. On the one hand, global financial regulators need to recruit well-qualified and independent experts and form

AI-dedicated units or divisions to work hand-in-hand with academic/industrial AI/ML experts of the finance sector. The latter aspect, nevertheless, also requires the resources necessary to attract and retain domain experts, as well as the equipment and infrastructure to motivate them. On the other hand, they must also think of effective ways to implement any future AI-targeting regulatory framework such as the one discussed in the present chapter. All in all, global financial regulators face today the universal challenge to render AI/ML and finance at the service of our global society.