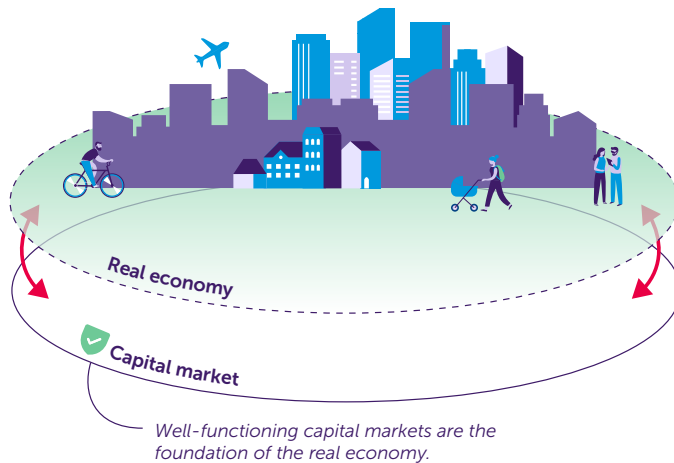


AI in Capital Markets: Balancing Innovation and Integrity

Introduction and purpose

Well-functioning capital markets are the foundation of the real economy. They support economic growth, help channel capital into productive investment, and contribute to households' financial wellbeing.



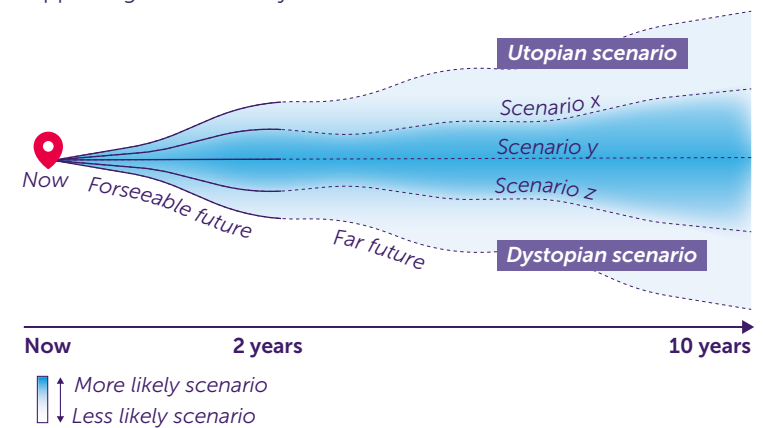
AI is an integral part of the way capital markets operate and is becoming more deeply embedded as its capabilities continue to evolve. As a result, the range of AI applications is expanding across the trade lifecycle: from the analysis of information before trades are made, to the execution of orders and the processing of transactions afterwards.

This exploratory AFM report looks at how the use of AI in capital markets could affect market integrity. It does not claim to be exhaustive or definitive. By considering both opportunities and risks of AI across the trading lifecycle, it is intended as a starting point for further debate, research, and coordinated action.

To illustrate what AI could mean for capital markets, **two extreme scenarios are laid out:** one optimistic and one pessimistic. Over the next 10 years, neither extreme is likely and the future will probably include elements of both. The point is that the choices being made today shape the direction markets will take.

In the utopian scenario, capital markets are fully trusted, transparent and essentially frictionless. AI ensures prices are accurate and information reaches everyone quickly and evenly. Supervision acts preventively and in real-time. Firms use self-auditing systems that continuously check trading activity in the background. Manipulative trading behaviour is difficult to execute by design, as trading agents have built-in constraints and actively steer trading decisions to support proper market functioning.

In the dystopian scenario, capital markets are opaque, unfair, and widely untrusted. With humans no longer meaningfully in the loop, AI allows engineered information to directly shape decisions. At the same time, autonomous trading amplifies harmful behaviour and self-reinforcing reactions between trading models distort prices. As trust breaks down, public markets completely lose their role in supporting the economy.



Concrete opportunities and risks today

This report examines current AI applications across the three phases of the trading lifecycle: pre-trade, execution, and post-trade.

Used responsibly, AI can improve the workings of capital markets.

Before a trade, it can widen access to financial insight and support more informed or personalised investment decisions. During trading, it helps deliver faster and better order execution and supports liquidity provision. It also enables earlier signal detection, helping markets reflect more accurate prices. In the post-trade phase, it can reduce the cost and burden of reporting.

The same capabilities that make AI powerful also expose markets to vulnerabilities. Existing risks return in new forms, while new ones emerge faster and with fewer warning signs. Traditional oversight is challenged most by self-learning models, particularly when they are given greater autonomy or are pursuing misaligned goals.

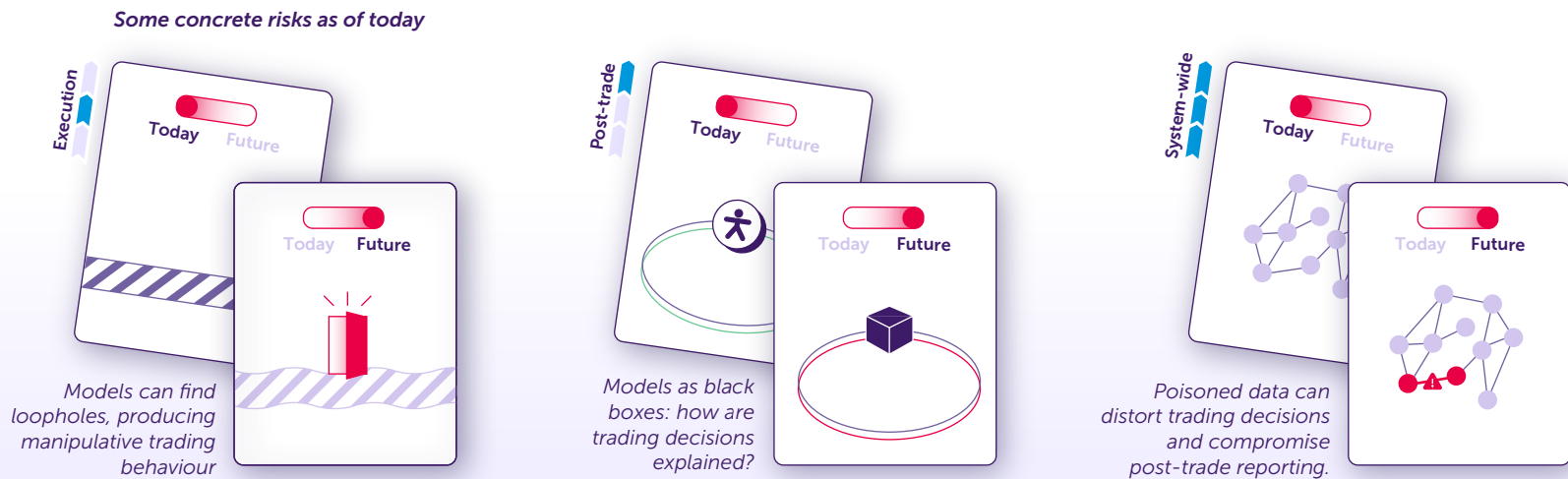
In the pre-trade phase, two risks appear. The first is AI-washing, which involves overstating or misrepresenting a firm's use of AI. The second is inappropriate investment advice: tools such as chatbots and AI investment assistants can give the impression of expertise without actual accuracy, or the regulatory protections investors would expect from formal advice.

In the execution stage, self-learning models can produce trading behaviour that is manipulative even if unintended. When targets are defined too narrowly or guardrails are weak or ignored, models can find loopholes. Another risk is that trading is drawn away from public exchanges because visible orders are more exposed to being exploited.

Post-trade, AI is changing how trading decisions are monitored, explained, and controlled after the fact. Firms may struggle to explain why a system executed a trade at a particular moment, especially when models operate as black boxes. Looking ahead, oversight may be further complicated by weak audit trails, flawed AI-generated compliance reports, and limited transparency from third-party vendors.

AI is also creating system-wide risks across the entire trading lifecycle. Data or other unregulated information streams that feed AI models can be poisoned, thereby distorting trading decisions and compromising post-trade reporting. This can be further reinforced when AI models operate in highly interconnected markets and respond to one another's trading behaviour, potentially destabilizing markets even without malicious intent. And as these systems become more autonomous, there is a risk that AI agents find ways to trade with other agents, without a human in the loop.

On the next page, a visual overview of the risks is presented.



AI in capital markets: a visual risk overview

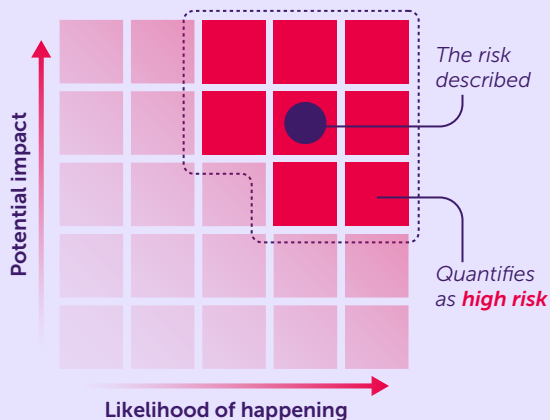
This page presents the main risks that we have identified. For a complete description of the risks, please refer to our full exploratory report [here](#).

The definition of a risk

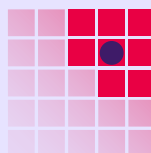
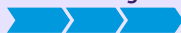
The AFM frames risk in accordance with this format:

“Certain developments, conditions, and behaviours of actor(s) that can lead to undesirable outcomes in the markets”.

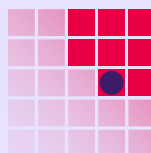
How to read this risk analysis?



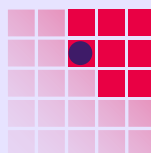
Crosscutting and system-wide



Poisoned data as a systemic risk for capital markets

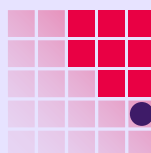
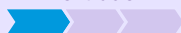


Agentic AI expands the risk surface in high-speed markets

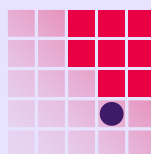


Concentration, correlated models, and common data sources as structural market vulnerabilities

Pre-trade

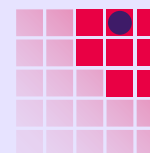


Treating GenAI as a substitute for investment advice: how inaccuracy and personalisation can harm retail investors

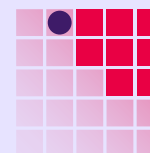


AI-washing: misleading investors through false or exaggerated AI claims

Execution

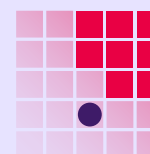


Harmful outcomes from the use of self-learning algorithms in trading



Order-flow exploitation and a decline in market participation

Post-trade



Opaque models: black-box behaviour and polluted regulatory data

Three conclusions

- 1. AI outcomes reflect human choices.** AI does not act in a vacuum. How a model behaves depends on the goals it is given, the limits built into it, the data it is trained on, and the environment in which it operates. A model can, for example, be designed to trade more cautiously in unusual market conditions. As trading systems become more autonomous, model design choices and ethics matter even more. Ongoing human oversight remains essential to check that trading models behave as intended, are ethical by design, and are not misdirected by poor or manipulated data.
- 2. When models move together, risks spread faster.** Markets become more vulnerable when AI trading models are similar or rely on the same inputs such as news sentiment. This can lead to correlated trading behaviour or feedback loops that push prices away from fundamentals and turn small shocks into larger market movements. Supervision needs to be adapted, and regulation reassessed to spot these system-wide risks and ensure they are mitigated timely and adequately.
- 3. Accountability remains with market participants.** AI does not dilute accountability. However complex or black box a model becomes, a market participant using it remains responsible for its trading behaviour. That means that AI trading models must remain controllable and compliant. Safeguards are needed to prevent model from exploiting loopholes, hiding problematic behaviour, or drifting into strategies that are profitable but harmful or illegal.

Three key areas for attention

- 1. Trustworthy AI models are essential to market integrity.** The AFM aims for markets in which AI trading systems are safe and reliable by design. This comes down to robust model validation, built-in safeguards, strong data governance, and continued human oversight. Model ethics, reliability, and transparency are not just compliance issues: they are key to market functioning and investor confidence, which in turns supports the real economy.
- 2. Supervising a mixed AI market.** Capital markets are likely to become a mix of governed, reliable AI models and less transparent or stable ones. Participants using well governed, transparent models can

expect predictable and proportionate supervision. High risk or opaque systems require closer scrutiny. Regulators should also consider whether today's market infrastructure remains suitable as trading becomes more autonomous. Before AI trading agents are used more widely, clear accountability rules must be in place, unauthorised trading environments must be identified and tackled.

- 3. Better understanding of market-wide AI effects.** Capital markets depend increasingly on the way AI trading models behave, both individually and in response to one another. Regulators need a clearer view of how these systems interact, and when those interactions could generate feedback loops, intensify market stress or increase trading costs through for example tacit collusion. This also requires a broader discussion about whether additional intervention tools are needed if model interactions disrupt fair and orderly markets.

Way forward

The AFM will take concrete steps to ensure that human oversight, transparency and accountability evolve alongside the growing use of AI in trading. Maintaining an agile and innovation-aware approach, the AFM aims to strengthen its supervisory capabilities and deepen internal expertise by translating the identified risks into actionable supervisory steps and defining the legislative mitigation measures they require. Externally, the AFM will maintain close and regular dialogue with the sector to exchange insights on a quickly innovating market and jointly anticipate potential risks and effective actions. This work will take place in close collaboration with national authorities such as ACM and the Dutch Ministry of Finance, as well as European and international bodies including the European Commission, ESMA, IOSCO, and other national competent authorities.