Date September 2026
Author Jimmy Kvarnström, Executive Director of Markets

**Finansinspektionen**
Box 7821
103 97 Stockholm
Tel +46 8 408 980 00
finansinspektionen@fi.se
www.fi.se

# Autonomy can be delegated to AI agents — accountability cannot

When algorithmic trading transformed European equities, it changed speed and scale, but not responsibility. The firms using those algorithms remained responsible for their use and effects. Agentic AI raises this accountability question in a broader setting. The answer remains the same.

Most AI in finance today informs human decisions. It may score credit, flag transactions or draft analysis. AI agents do something different. They act. They pursue objectives across chains of tasks, interact with other systems and agents, and adapt as they go. They can support stronger risk monitoring, automated controls and better services. But as agents become embedded in trading, risk management, compliance and customer processes, more decisions will happen at machine speed, without a human in each loop.

This raises specific challenges for both firms and supervisors. Accountability can blur when outcomes emerge from chains of automated steps rather than identifiable decisions. Auditability becomes harder when systems may not produce the same output every time and the basis for an action is difficult to reconstruct. When many agents interact, collective behaviour can emerge that no single firm designed. This may include correlated reactions, herding and, in markets, conduct that resembles coordination without agreement.

There is also a concentration and resilience risk. If many firms build agents on a small number of underlying models and providers, behaviour may correlate and operational dependencies may deepen — a systemic dimension that individual firms should assess and manage in their own model and provider choices. The same capabilities can also make it faster to find vulnerabilities and automate attacks, raising operational resilience risks.

When agents are supplied by third parties, it is necessary to know before deployment who performs the regulated activity and who answers for it.

This is not without precedent. When algorithmic trading grew, MiFID II did not prohibit automation. It required firms to demonstrate control. Algorithms had to be tested before deployment, monitored in operation, capable of being stopped immediately, and documented so that behaviour could be reconstructed afterwards. The principle was simple. Autonomy is acceptable where control is demonstrable. This principle also applies to agentic AI. It enables rather than restrains. Firms that can demonstrate control can deploy with confidence.

For firms, these principles translate into practical safeguards, extending the logic of MiFID II's algorithmic-trading rules to a new setting. Every agent should have a clearly identified owner, meaning a function with senior accountability for its objectives, boundaries and conduct. Mandates should be explicit, with limits on what an agent may do and the resources it may commit. Testing should precede deployment, monitoring should be continuous, and intervention, including immediate shutdown, should always be possible. Logging should allow firms to reconstruct what an agent did on what inputs and on what basis. Safeguards should be proportionate. An agent drafting analysis should not be governed in the same way as an agent committing capital.

Human oversight must be meaningful rather than nominal. With autonomous agents, oversight of every action is neither possible nor the point. What matters is oversight of the system — its design, objectives, boundaries and performance — and clarity about who exercises it. A human approving machine output faster than it can be read is not oversight. Governance should treat agents as the firm would treat anyone acting in its name, with defined authority and active supervision.

Supervision will also need to evolve. When decisions move into processes, supervisors must assess model governance, testing regimes, monitoring and controls, not only outcomes. This requires new capabilities and access to data that makes firm behaviour observable. It also requires supervisors to understand these systems from the inside, including by using them in their own work. The same technology that challenges supervision can strengthen it through sharper monitoring, faster analysis and earlier detection. This makes early dialogue between firms and supervisors important.

EU coordination would help before national practices diverge. The risks of agentic AI do not follow borders, and neither do the firms using it. Common expectations on accountability, auditability, testing, resilience and human oversight would give firms predictability and supervisors a shared baseline. This should build on technology-neutral rules and the AI Act, not duplicate them. Guidance can be common, while supervision should remain close to the firms and markets where agents operate.

The principle to uphold through all of this is well established. Firms may organise how they act through employees, systems or algorithms, but they cannot outsource responsibility for what is done in their name. Autonomy can be delegated to machines. Accountability cannot.

Autonomy is acceptable where control is demonstrable.

*Originally published in Views, The Eurofi Magazine, September2026*