



Cercle IA et finance

Paris, 4 February 2025

The foundations of trustworthy AI in the financial sector

**Speech by Denis Beau
First Deputy Governor of the Banque de France**

Ladies and gentlemen,

First of all, I'd like to thank the organizers for their invitation to launch this event focusing on the Paris financial centre's AI strategy: just days before the **international AI Action Summit**, this gives me the opportunity to reiterate our determination at the Banque de France and the ACPR to **take action** on this major issue for the industry – and to do so in concert with all financial sector players. The summit will also be an opportunity for the Banque de France to reaffirm its commitment by organising a side event on 11 February, featuring a round table discussion on ethical and inclusive AI.

AI – as you are already well aware, is being increasingly used in the financial sector, whether to assess credit risk, set insurance rates or estimate asset volatility. For a supervisor, its impact is potentially double-edged: while AI is a source of **opportunities** for the sector – including for its supervisor – it is also a **new vector of risk**. This ambivalent impact partly explains the **regulatory framework** that has just been introduced in Europe.

The European Union has proven itself a pioneer in this area by adopting the AI Act in the summer of 2024. However, this legislation raises legitimate questions, especially for the financial sector: is there not a risk of hampering innovation in the name of controlling risk? I would like to reiterate, before you today, a **strongly held conviction** that may seem iconoclastic in the current environment: **in the long run, regulating AI-related risks is good for competitiveness in both Europe and France. Without regulation, there can be no trust – and therefore no sustainable innovation.**

Because my opening remarks this morning are from a supervisor's perspective, I will discuss the opportunities and risks (I), then the conditions necessary for effective regulation of AI in the financial sector (II).

*

I/ To get a bit of perspective on things, I would like to revisit an initial observation: **AI, combined with an abundance of available data, is a powerful vector of transformation for the financial sector.**

1/ Our observations show that **AI is increasingly being used by financial institutions** along all segments of the value chain: i) to improve the “user experience”, ii) to automate and streamline internal processes, and iii) to control risks, particularly in the battle against fraud and against money laundering and the financing of terrorism.

The emergence of generative AI two years ago has triggered **a revolution in the accessibility of AI technology**, thanks to the possibility of interacting with algorithms using

natural language – via Large Language Models (LLMs) – **which makes adoption considerably easier**. Generative AI is also boosting **innovation** within companies as computer code can now be written by a much broader group of people.

If harnessed properly, AI can therefore **boost the efficiency** of financial institutions, **increase their revenues** and provide them with **risk management** solutions.

2/ **However, there is a downside, and the power of the solutions developed is accompanied by significant risks**, both for each of the players in the financial system and for the stability of the system as a whole. I would like to mention three of these risks.

The first is that these technologies may be put to improper use. The **complexity** and **newness** of certain modelling techniques can result in more errors, either in systems design or use. This poses a risk not only for customers, but also for institutions' financial health, as a poorly calibrated model could generate systematic losses. **These risks are compounded by two factors.** First, **the adjustment of the parameters of certain models in real-time**, which is one of their strengths, can also result in **rapid drift**. Second, certain AI systems are **particularly opaque**, generating a “**black box**” phenomenon.

The second risk I would like to highlight is **cyber risk**, which has become the number one operational risk in the financial sector over the past few years. **AI amplifies this risk** – both in terms of the danger posed by attackers and because it represents a new area of vulnerability. Conversely, we should be aware that AI can also enhance IT security, for example, by helping to detect suspicious behaviour.

Lastly, I'd like to highlight a third risk, which could become increasingly significant in the future, namely **environmental risk**. In the absence of reliable data provided by businesses or a commonly accepted basis of calculation, quantification of this risk is still subject to considerable variability. Nevertheless, it is clear that **training** the most recent generative AI models is a **very energy-intensive process**... and that if current trends continue, their **regular use by billions of customers will be even more so**. These factors naturally suggest that AI should be **used rather frugally**. In other words, AI systems should only be used when necessary.

II/ I would now like to turn to **aspects of regulation, legislation and control**, and primarily to the **European AI Act**. This will mainly concern the financial sector for two use cases: creditworthiness assessment for granting loans to individuals, and risk assessment and pricing in health and life insurance. The main impacts of this legislation will be felt from August 2026,

and as **market surveillance authority**, the ACPR should be responsible for ensuring that it is properly applied.

With this in mind, I would like to share **two simple messages** with you this morning: i) the risks linked to AI can essentially be handled within the existing risk management frameworks; ii) however, we should not underestimate certain new AI-related technical challenges.

1/ The AI Act will not lead to any major upheaval in the way risks are managed in the financial sector.

Financial institutions have a sound risk management culture, as well as robust **governance** and **internal control** systems. The Digital Operational Resilience Act (**DORA**), which has just come into force, rounds out the traditional regulatory framework with specific rules on operational resilience and IT risk management. **The financial sector is therefore well equipped** to meet the challenge of complying with the new regulations.

Admittedly, **the objectives** of the AI Act – first and foremost the protection of fundamental rights – and those of sectoral regulation – financial stability and the ability to meet commitments to customer– **differ. But operationally**, when the AI Act requires "high-risk systems" to have data governance, traceability and auditability, or guarantees of robustness, accuracy and cyber-security throughout the lifecycle, clearly, we are not in uncharted waters.

Rather, I would like to reiterate that **the usual principles of sound risk management and governance continue to apply under the AI Act**. Naturally these will guide the ACPR in assessing systems compliance when it is called upon to exercise its role of market surveillance authority. More specifically, our vision for deploying this new mission will be underpinned by three simple principles: (i) implementing "**market surveillance**" in accordance with the AI Act, i.e. primarily aimed at identifying systems likely to pose compliance problems; (ii) defining supervision priorities using a risk-based approach to ensure that the resources deployed are **proportionate** to the expected outcomes; and (iii) unlocking **all possible synergies with prudential supervision**. I believe that this was the intention of the European legislator when it entrusted national financial supervisors with the role of "market surveillance authority". It is also the best way of ensuring that we don't make the regulations any more complex at a time when our common objective should be to **simplify** them.

Naturally, **the principles** of good governance and internal control **also apply to algorithms not considered high-risk** by the AI Act, if they pose risks to the organisations concerned – think of the use of AI systems in market activities, for example. Here, lessons learned from implementing the AI Act and the resulting best practices will be invaluable for both supervisors and supervised entities.

2/ Nevertheless, the challenges posed by the use of AI should not be underestimated

Some of the issues raised by this technology are **definitely new**. Let me give you two examples. Firstly, **explainability**: with each advance in this field, artificial intelligence algorithms have become increasingly opaque and in a regulated sector like the financial sector, this is a problem. More specifically, **day-to-day users** of AI tools need to have a sufficient understanding of how they work and of their limitations if they are to make appropriate use of them and avoid the twin pitfalls of either blindly trusting the machine or systematically mistrusting it.

The second example is **fairness**. AI can accentuate biases present in data. Indeed, one of the aims of the AI Act is to detect and prevent such biases before they cause harm to citizens. This is a **technically complex** issue, as banning the use of certain protected variables is not enough to guarantee safe algorithms. This is particularly true for activities such as granting loans or pricing insurance, where **customer segmentation is part of normal business and risk management practices** in a competitive environment.

To address these new challenges and comply with the various regulatory requirements, **financial institutions will need to** acquire new human and technical resources and **upskill**. As market surveillance authority and prudential regulator, the ACPR will ensure that **risks are effectively managed**. Compliance with the AI Act will have to be more than just an internal administrative labelling exercise, and financial institutions will have to ensure that the algorithms are managed and monitored by competent people who understand their inner workings.

This means that the **financial supervisor** itself has to upskill and **adapt its tools and methods**. The ACPR has already published certain proposals in the past concerning the issue of explainability. It will eventually have to establish a doctrine on this topic as well as on algorithm fairness. We will also need to develop a specific methodology for auditing AI systems.

We cannot and must not take this methodological step forward alone. In addition to **unlocking synergies with other AI supervisors** in France and Europe, we need to **cooperate with the financial sector**. **Supervisors and supervised entities share many challenges** and they will overcome them more effectively if they are able to move forward together.

Events like today provide an opportunity to channel our collective efforts into a widely shared project. **It is by working together that we will be able to lay the foundations for trustworthy AI in the financial sector.**

I wish you fruitful discussions throughout this morning.