

## Denis Beau: Artificial intelligence evolution and outcomes – digital trust, data and cloud

Introductory remarks by Mr Denis Beau, First Deputy Governor of the Bank of France, at the Point Zero Forum 2024, Zurich, 2 July 2024.

\* \* \*

Ladies and Gentlemen,

I am very pleased to open this panel on artificial intelligence (AI). AI is a **breakthrough innovation** that can lead to real economic disruption. This is especially true in the financial sector, where AI has been an important **main driver of transformation** in recent years. The advent of **generative AI** is expected to further accelerate this trend, not only *by increasing users' adoption* of AI tools, but also by **structurally accelerating the pace of innovation** (let's think, for example, of the new ability to generate computer code based on natural-language queries).

However, these significant developments raise a number of questions, including from the central banker and financial supervisor perspective with which I am looking at them. I would like to share some of these questions with you, before **expressing my views** about how we should tackle the issue.

1/ First, despite recent progress, **the underlying AI technology does not yet appear to be fully mature**, particularly as regards generative AI (GenAI). There are still a number of unanswered questions on this subject. **I will touch on two of them.**

**First, the question of general-purpose AI (GPAI) models:** how will they **perform** in a whole range of tasks relevant to the financial sector? This question really arises on two levels. Will **GPAI models become the standard for all uses**, to the detriment of specialized models? Will smaller, well-trained - that is, more specialized - models have the ability to live up to larger, more general-purpose models? These performance issues have many potential consequences, not least in terms of **competition**: if large GPAI models are introduced in all areas, we run a high risk of ending up in a **natural monopoly or oligopoly**, adding to the **already largely oligopolistic nature of the cloud market**.

**Second, the issue of the vulnerabilities of AI systems:** while we are starting to gain a clearer picture of the situation, research on this subject is far from complete. This is particularly true in the field of **cyber security for GenAI models**, with the recent discovery of the dangers of "indirect prompt injection". While this race between the 'sword' (development of new attack techniques) and the 'shield' (development of effective countermeasures) is traditional in the security field, our ability to **adequately secure AI systems** will have a major influence on the ability of different actors to make extensive use of this technology.

2/ Even though AI technologies are not yet fully mature, it seems to me that **central banks and financial supervisors should embrace them without delay**, for at least **three reasons**.

**First, to continue to carry out our missions effectively, by doing more and doing it better.** AI can of course help us become more efficient, by increasing the level of automation. But we also want **to offer new capabilities to agents.** For example, our LUCIA tool, an AI-based system with the capacity to analyze large volumes of banking transactions, allows us to assess the performance and relevance of banks' AML/CFT models during our on-site inspections

**Second, to develop critical expertise in AI.** Using AI for our own purposes allows us to gradually acquire a good command of the technology, and is a very effective way of **properly understanding its benefits and risks.** The virtues of **learning by doing** explain why internal uses of AI are very complementary to the supervision of AI systems deployed by the financial sector. For example, very recently, the ACPR, with the help of Banque de France's innovation center, Le Lab, organized a **"Suptech Tech Sprint"**, a hackathon intended to explore what generative AI can bring to the various supervisory functions. In three days, this event **revealed the potential of large language models (LLM) for supervision.**

**Finally, to drive the financial ecosystem, by sending a signal to the market that it too can – or must – take the plunge.** For example, the cutting-edge work being carried out at the Banque de France on post-quantum cryptography is raising awareness among private stakeholders about the need to address this threat. So, while it is clear to me that central banks and supervisors must seize the opportunities offered by AI, the question is: **How do we do that?**

3/ It seems to me that we must first **lay down a fundamental principle of governance: AI must be at the service of humanity and society, and not the other way round.** From this perspective, even if it does not solve all the problems, the ongoing adoption of the **European AI Act, the world's first binding text laying down the principles of "trustworthy AI", is a welcome step.** In particular, this text will increase consumer confidence, while providing legal certainty for economic operators. This governance principle can be supplemented by **three operational principles.**

**First, using AI proportionately and progressively. With a simple rule: the more critical the use case for our activity, the more we have to do it ourselves.** For institutions like ours, this goes to the fundamental issue of **data:** some of the data held by central banks and financial supervisors are **too confidential** to be stored on a third-party cloud infrastructure.

**Second, experimenting without delay, even with simple use cases, to find the right way of integrating AI into our activity, leading to an "augmented agent" rather than a "substituted agent".** Indeed, we can expect AI to significantly reshape the patterns of human-machine interactions. Finding the right combinations will encourage the adoption of the new tools, **by winning the buy-in of users,** which is a crucial issue.

**Third, collaborate with others, to share best operational practices and to build a coherent AI supervision framework.** Of course, I am thinking first of international cooperation, because AI-related issues are by their very nature global. In this area, while there may be nuances in terms of how to proceed, I note above all that **many**

**jurisdictions are expressing similar concerns, which should enable international cooperation to move forward.** But we also need to **cooperate with authorities in other sectors**, especially competition, cyber security, fundamental rights and even the green transition, as AI-related concerns are largely **interconnected**. In my view, these different forms of cooperation are an **essential condition** if we are to contribute to the emergence of the most relevant and resilient AI models, in other words, if we are to influence the development of the technology **in the direction of the general interest**.

Thank you for your attention.