

Box 2.3

ARTIFICIAL INTELLIGENCE: AN AREA OF GROWING OPPORTUNITIES AND NEW SUPERVISORY CHALLENGES

In recent years, artificial intelligence (AI) has experienced a dramatic rise in popularity. This is due both to continuous advances in computational capabilities and to the emergence of large, state-of-the-art foundation models¹ capable of processing vast quantities of information. New AI models make it possible to tackle tasks normally associated with human intelligence, such as translating languages, answering questions or even reasoning.

These technologies are used by financial institutions, for example, to automate and optimise administrative tasks, analyse large volumes of data or in customer relationship applications. Thus, we are seeing AI-powered conversational assistants (chatbots) that can handle customer queries 24 hours a day, the use of advanced classification algorithms to detect financial fraud, and the application of AI to assess customer credit risk, among other applications.

From a general point of view, these general-purpose technologies have the potential to boost productivity across the entire economy. However, their use also entails risks, some of which are particularly relevant in the area of conduct supervision, where models can cause discrimination in the granting of credit, on the basis of biases in the data used to develop them, that violates the fundamental rights of customers. It is worth recalling the difficulty in explaining these models, as they use billions of parameters, thereby making it complicated to understand the factors that determine the results obtained, which is why they are known colloquially as “black boxes”.

To address this situation, Regulation 2024/1689 of 13 June 2024 on AI regulates the use of AI systems in the European Union (EU) with an approach based on risk levels and the protection of fundamental rights. It therefore identifies AI systems used in the granting of credit to natural persons as high-risk AI systems. In the case of these high-risk systems, developers and users must carry out comprehensive assessments of the risks associated with their models before they are put into production. In particular, the European regulation states they must comply with risk management and data governance requirements, have adequate technical documentation, analyse their results as well as have sufficient levels of accuracy, robustness and cybersecurity. It specifically stresses that AI systems must ensure that they are transparent and explainable. The latter entails providing

clear information on how the models work and how decisions are taken on the basis of said models. Similarly, they must apply supervisory mechanisms to ensure that automated decisions are reviewed (and, if necessary, corrected by persons).

Once these high-risk AI systems go to production, the regulation identifies that they must be supervised by market surveillance authorities. These authorities are designated at the national level and, in the case of credit models, the banking supervisory authority must also report incidents to the European Central Bank (ECB).

In view of the foreseeable designation of the Banco de España as the market surveillance authority for the use of AI in lending to physical persons, the Bank is already working towards assuming these powers. This surveillance work will include, inter alia: supervision and control of supervised institutions’ compliance with the regulation; assessment of the compliance of these models with the regulations; international collaboration with other market surveillance authorities for the exchange of information and the consistent application of the regulation throughout the EU; monitoring of incidents; and promotion of innovation.

In the exercise of these functions, the Banco de España will have to coordinate with other bodies, including the Spanish Agency for the Supervision of Artificial Intelligence. Thus, for example, the European Banking Authority and the ECB, together with the national competent authorities of the Single Supervisory Mechanism, have been working on the use of AI in regulatory capital calculation models, setting regulatory expectations and supervisory techniques. These expectations will be reflected in the update of the ECB’s Guide to Internal Models, which is due to be published in 2025.

In short, the rise of AI and the opportunities and challenges it brings have led the Banco de España to set up a cross-institutional working group which, together with the deployment of the market surveillance authority function, includes various lines of work to ensure the necessary human and computational resources, the deployment of use cases in the various functions exercised by the Bank and the development of an internal governance model that ensures the ethical and responsible use of AI.

¹ Foundation models are models trained on large quantities of data that can subsequently be specialised to perform specific tasks through retraining with more specific data.