# The impact of AI on the financial sector and supervision

DeNederlandscheBank
EUROSYSTEEM

AFM

# Key messages

- Artificial intelligence (AI) is advancing rapidly in the financial sector, in the Netherlands as elsewhere. Although AI provides substantial opportunities, it also presents risks. On the one hand, it can create better customer service and more personalised products, and offer institutions opportunities to increase their revenues and cut costs. On the other hand, there are risks, including in the areas of data quality, data protection, explainability, incorrect results, discrimination and exclusion, and greater dependence on third parties.
- Financial institutions are expected to use AI responsibly. The objectives of supervision and the standards institutions are required to comply with are technology-agnostic, and apply just as well when AI is used. The use of AI is part of existing processes and services supervised by the AFM (the Dutch Authority for the Financial Markets) and DNB (De Nederlandsche Bank) under the financial legislation and regulations in force.
- AI has consequences for the way in which the AFM and DNB supervise the financial sector. The AFM and DNB will need to expand their knowledge in this area to enable them to assess institutions' use of AI. Regulatory methods and procedures will need to be developed or changed in certain cases to take new technologies into account. Supervision will focus *inter alia* on managing the risks of using AI applications, the modalities of those applications, and the results.
- Additional requirements may need to be laid down, depending on how AI is being used, an existing example being the rules on algorithmic trading on the capital market. Specific legislation on the responsible use of AI is currently sparse. General standards will need to be clarified or specified as the use of AI and its importance to the financial sector increases. This should preferably happen in a harmonised way at the European level, *inter alia* for the sake of regulatory convergence and in order to create a level playing field.
- Regulation will need to achieve a good balance between the responsible use of AI on the one hand, and room for innovation on the other.
- The European AI regulation designates AI systems used to check the creditworthiness of natural persons and for risk assessment and pricing in the case of life and health insurance as high-risk. Additional requirements will therefore apply to the development and the controlled and responsible use of those AI applications, most of them in line with existing rules on risk management and governance. In the case of other AI applications, institutions will be encouraged to comply with requirements for high-risk AI systems voluntarily, *inter alia* by observing European guidelines on trustworthy AI. The AFM and DNB fully support this.
- The AI regulation will require financial institutions using AI to pay particular attention to the proper protection of fundamental rights. In the case of high-risk applications, institutions will be required to assess the impact on the fundamental rights of individuals or groups.
- Where AI is used for financial services, the principle underlying the AI regulation is that the existing financial regulators (i.e. the AFM and DNB in the Netherlands) will also take responsibility for monitoring compliance with the regulation. Broad coordination and collaboration between regulators of AI is desirable at both the European and national level, especially in view of the particular risks to fundamental rights posed by AI systems.

# Summary

**The rapid development of artificial intelligence (AI) poses challenges for the regulatory work of the AFM and DNB.** AI is advancing rapidly worldwide, including in the financial sector. Although it provides many opportunities, it also presents substantial risks to institutions and society. If the risks of AI materialise, that could have an impact on the financial soundness and integrity of institutions, and affect customers' interests and the relationships between financial market players. These developments confront the AFM (the Dutch Authority for the Financial Markets) and DNB (De Nederlandsche Bank) with the question of how the responsible use of AI in the financial sector can be guaranteed and AI supervision can be shaped. The aim of supervision is to ensure that institutions adequately manage the risks of AI applications, while leaving enough room for innovation within the limits of the law. AI is changing rapidly, and regulation is changing in tandem. Hence the findings of this report do not provide ready-made answers to the issues mentioned. The aim of the report is to formulate criteria and areas of attention when shaping AI supervision.

**Essentially, AI involves computer systems carrying out tasks normally requiring human intelligence, or surpassing it.** AI applications are advanced statistical approaches to large quantities of data that enable predictions, recommendations and new content to be generated. The uses of AI applications have really taken off, thanks to an exponential increase in computing power and data, and the progress of computer science. The rise of generative AI, which enables new content to be created, marks a new phase in AI evolution.

**Dutch financial institutions have been using AI for some time now and are experimenting with advanced AI models, which suggests that the use of AI will continue to grow in the years to come.** AI is used e.g. for fraud prevention and detection, and to combat money laundering, terrorism financing and cyber crime, to assess creditworthiness and verify identities. It also helps staff to work more efficiently, for instance by summarising telephone conversations with customers automatically. Dutch institutions say that they regard the ethical use of AI as very important, and they are not trying to push the boundaries of what is permitted and possible under the current and forthcoming regulation. Many institutions also state that they are cautious about using generative AI at the moment, albeit they recognise the potential of this technology. Some of them are gradually starting to use generative AI for support processes. New use cases are expected to be discovered in the future that financial institutions could include in their operations. Financial institutions might also be inclined to use more AI for reasons of competition. Too much caution could adversely affect their competitiveness vis-à-vis foreign financial institutions and non-financial players on the Dutch market.

**AI offers substantial opportunities for both customers and financial institutions.** AI provides opportunities to make smarter (more efficient) use of data, e.g. enabling products to be offered that are more tailored to individual customer needs. It can also speed up the provision of services and possibly reduce the barrier to accessing financial services. It offers financial institutions opportunities to grow their revenue

by tapping into new customer groups, increasing cross-selling opportunities and optimising prices. Behind the scenes, AI can make processes more efficient, thus reducing the costs for institutions. It can also help financial institutions to gauge risks better and more consistently, provided they have good-quality, complete data and models that are explainable and verifiable. Lastly, AI has the potential to improve defences against cyber attacks and combat financial crime more effectively and efficiently.

**The use of AI applications, however, also presents risks that could have an impact on financial institutions and their customers.** These risks therefore fall within the supervisory sphere of the AFM and DNB. AI could cause socially undesirable outcomes, for example, as a result of discrimination and exclusion, possibly resulting from using biased data or data of insufficient quality. The fact that large quantities of data need to be processed also leads to security and data protection risks. Advanced AI models, moreover, can take complex decisions that are difficult to explain, thus undermining the explainability and transparency of financial outcomes. The use of AI by financial institutions could also lead to dependency risks, because suppliers of AI technology currently only constitute a small (mainly non-European) group. Other risks include inadequate governance frameworks and the high energy consumption of AI.

**Regulation will need to achieve a good balance between the responsible use of AI and room for innovation.** What impact AI will have on the financial sector in the years to come is difficult to gauge. A scenario analysis of trends in the development of regulation and innovation in AI applications underscores the importance of balance. A scenario in which a balance is achieved between growing innovation and adaptive regulation seems to offer the best prospects. Conversely, scenarios in which there is an imbalance between regulation and innovation seem to produce less favourable results. This emphasises the importance of a pro-active, adaptive attitude on the part of policy-makers and regulators. Regulation needs to be able to create a framework for responsibility and transparency that at the same time provides enough room for innovation and rapid developments in AI products and services.

**The objectives of supervision and the standards that institutions must comply with are technology-agnostic, and apply just as well when AI is used.** Institutions are expected to use AI responsibly. AI systems must not jeopardise the financial soundness and integrity of financial institutions, nor may the use of AI harm customers' interests or the integrity of relationships between market players. The AFM and DNB apply the existing regulatory framework to AI systems as well. For instance, institutions must comply with the requirements of sound and ethical operational management. This means setting up proper risk management that includes the risks of using AI. Standards relating to customers' interests (e.g. the duty of care), standards relating to product development and distribution, and standards to prevent excessive borrowing also apply to the use of AI.

**AI will have consequences for the way in which the AFM and DNB supervise the financial sector.** The AFM and DNB will need to expand their knowledge in this area to enable them to assess the use of AI by institutions. Regulatory methods and procedures will need to be developed or changed in certain cases to take account of new technologies. Supervision will focus *inter*

*alia* on deciding about and managing the risks of using AI applications, the modalities of those applications, and the results. As the use of AI and its importance to the financial sector increases, the AFM and DNB will step up their supervision of the way in which institutions manage the risks of AI use: this year, for instance, DNB is carrying out a thematic examination of the use of AI by insurance companies.

**Further guidance, and in certain cases specific regulations, may be needed to provide institutions with clarity and enable effective supervision.** Specific references to AI systems in the current regulations are few and far between. AI is explicitly included in sector-based financial legislation in several areas: for example, the rules on algorithmic trading on capital markets, specific requirements for the use of AI in consumer credit, and rules on automated advice. With a view to aspects such as regulatory certainty, it may be desirable to clarify in other areas too what is expected of an institution in guidance documents, and where necessary in specific regulations. Regulations on banks' and insurance companies' internal models lay down many specific requirements with which AI models also need to comply, which can make the use of certain AI models difficult. If those models are found to have added value, it may be necessary to change the current framework of standards.

**The European AI regulation lays down more detailed requirements regarding the controlled, responsible use of AI used to check the creditworthiness of natural persons and for risk assessment and pricing in the case of life and health insurance,** designating these AI applications as high-risk. Consequently, there are requirements regarding *inter alia* risk management, data quality, technical documentation, human supervision, robustness and transparency for users, in line with existing rules on risk management and governance. Institutions also need to assess the impact of these AI systems on the fundamental rights of individuals or groups. The AI regulation recommends that applications that it does not designate as high-risk should also comply with the requirements for high-risk systems as far as possible. The AFM and DNB fully support this.

**Regulation on the use of AI in the financial sector will remain fragmented, hence standards need to be clarified, and there needs to be close collaboration between the regulators.** At the European level, the sector-based regulatory frameworks need to link up with the complementary AI regulation. The European supervisory authorities (ESAs) can provide further clarification where necessary. In order to avoid supervision itself becoming fragmented, and given the sector-based expertise and experience that has been acquired, the supervisory role under the AI regulation is assigned in principle to the existing financial regulators insofar as it relates to financial services. The current allocation of supervision between the AFM and DNB is expected to apply also to the supervision of AI systems.

**Broad coordination of AI supervision is desirable at the national level as well.** Use cases of AI can be found throughout the public and private sectors. Existing regulatory frameworks are already being applied outside the financial sector too, e.g. to AI applications in health care. The risks of AI systems are best assessed in their specific contexts: this requires sector-based expertise, which limits the extent to which the details of standards can be harmonised. Collaboration between the regulators is desirable for the sake of effective and efficient supervision, while respecting those boundaries.

# Contents

# Introduction

**Artificial intelligence (AI) has made substantial strides in recent years.** Many people have started to appreciate the direct added value of AI in their work or private lives, in particular because of the breakthrough of generative AI. Given the opportunities for increasing productivity and developing innovative products, organisations worldwide have started looking at what AI could offer them, and a host of new companies have entered the market offering new applications. Financial institutions have been using systems and models involving AI for some time now. Although it provides opportunities, the use of AI also presents risks to financial institutions and society as a whole. For instance, AI models that use inadequate data will assess risks incorrectly, and they can contain biases that disadvantage particular customer groups. AI can therefore have an impact on the financial soundness and integrity of institutions, and on customers' interests. This brings it within the supervisory sphere of the AFM and DNB in the financial sector. AFM and DNB supervision focuses on the responsible use of AI, taking advantage of the benefits that AI can bring while properly managing the risks.

**The AFM and DNB face similar issues regarding AI supervision.** The technological developments in the financial sector, and the introduction of the European AI regulation in the future, are driving the AFM and DNB as regulators to reflect on the requirements laid down for the use of AI in the financial sector and on what form supervision should take. AI is changing rapidly, and the regulations are changing in tandem, hence the findings of this report do not provide ready-made answers to the issues mentioned. The aim of the report is to formulate criteria and areas of concern when shaping AI supervision. The AFM and DNB are keen to enter into a dialogue with the sector, their customers, authorities and other stakeholders on the opportunities and risks of AI.

**The report is structured as follows.** Chapter 1 broadly describes how the report defines AI and what forms AI can take. It then provides an overview of current AI applications in the financial sector and the opportunities and risks they involve. Chapter 2 discusses the supervisory tasks associated with AI. It outlines the current regulatory frameworks and explains the future changes in AI supervision with the advent of the AI regulation. The report does not consider the use of AI applications by the AFM and DNB themselves (AI-assisted supervision) or the impact that wide-ranging AI-led changes in the economy could have on the financial sector.

# 1 AI in the financial sector

There are many different ways of explaining what AI involves, but the AFM and DNB use the definition in the AI regulation (§1.1). Financial institutions are already using AI models in various areas (§1.2), including to monitor transactions so as to prevent money laundering and terrorism financing, and to make processes more efficient. They are also carrying out a lot of experimentation with AI, which suggests that the use of AI will increase sharply in the years to come. The use of AI in the financial sector brings many opportunities for both customers and financial institutions (§1.3). At the same time, it presents new risks that could have an impact on the way in which financial institutions deal with their customers and on their financial soundness (§1.4). Future use of AI will depend ***inter alia*** on developments in innovations and regulation. Various future scenarios also have differing implications for financial sector supervision (§1.5).

## 1.1 What do we mean by AI?

**Essentially, AI involves computer systems carrying out tasks normally requiring human intelligence.** The precise definition of AI (artificial intelligence) is under discussion in the scientific and political communities and in practice. The AFM and DNB use the definition proposed by the OECD[1] (Organization for Economic Cooperation and Development) that has been adopted in the European Union's AI regulation, as it will have a major influence on the scope of policy and regulation.

The definition of an AI system in the AI regulation is *"a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments."*

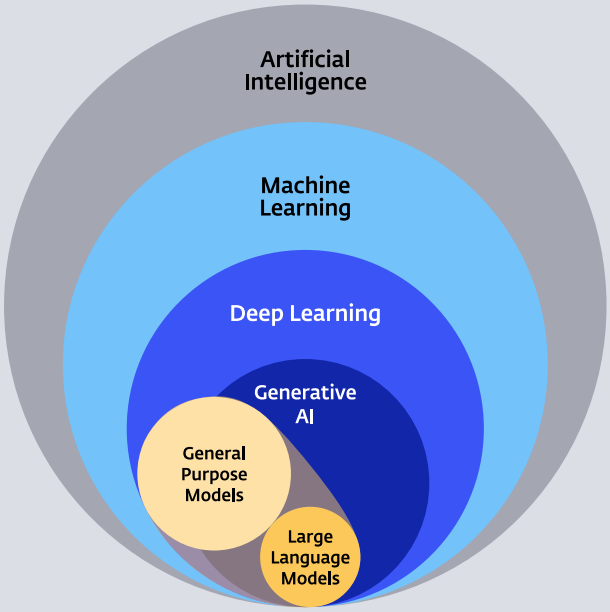
**There are various types of AI, which are explained in Box 1.** They differ in the degree of autonomy and the learning element that is contained within the AI system.

1 OECD (2023), Updates to the OECD's definition of an AI system explained

## Box 1 The various types of AI

AI is a broad concept that can be perceived in different ways. AI applications are in effect advanced statistical approaches to data that enable predictions, recommendations and new content to be generated. There is a distinction between e.g. machine learning and deep learning. Generative Artificial Intelligence (GenAI), which has recently grown sharply in popularity and is capable of generating new content (e.g. ChatGPT and Midjourney), is a deep learning technique. See Figure 1.

### Figure 1 The relationships between different types of AI




**Machine learning** is a type of AI in which the model learns from existing data and is capable of enhancing that data, enabling decisions or predictions to be made. The model 'recognises' patterns, as it were, and can develop them in more detail. The model can learn in various ways: for instance, it can use labelled data (supervised learning), non-labelled data (unsupervised learning) or feedback-based data to enhance the AI system (reinforcement learning).

**Deep learning** is a type of machine learning that employs various learning layers to simulate the learning process in the human brain, using what are known as 'neural networks' to analyse data and make decisions. Neural networks have an input layer (the input data to the network), a hidden layer or layers (intermediate layers where the combinations and transformations take place), and an output layer (the result of the computations based on the input provided). The complexity of deep learning models makes it difficult or impossible to relate the results to specific inputs, which is why they are referred to as 'black box' models.

**Generative AI** employs deep learning to create new content based on 'prompts' and existing data, e.g. new texts, images, audio and video.

**Large language models (LLMs)** are generative AI models that make use of language, taking huge quantities of data from human language use (e.g. books and online forums) to learn to generate texts. The AI regulation refers to general-purpose AI models, i.e. AI models with wide-ranging possibilities that can be used for multiple purposes rather than one specific purpose. These models can provide a basis for a wide range of specific uses: there are LLMs, for instance, that can do programming.

## 1.2 Current uses in the financial sector

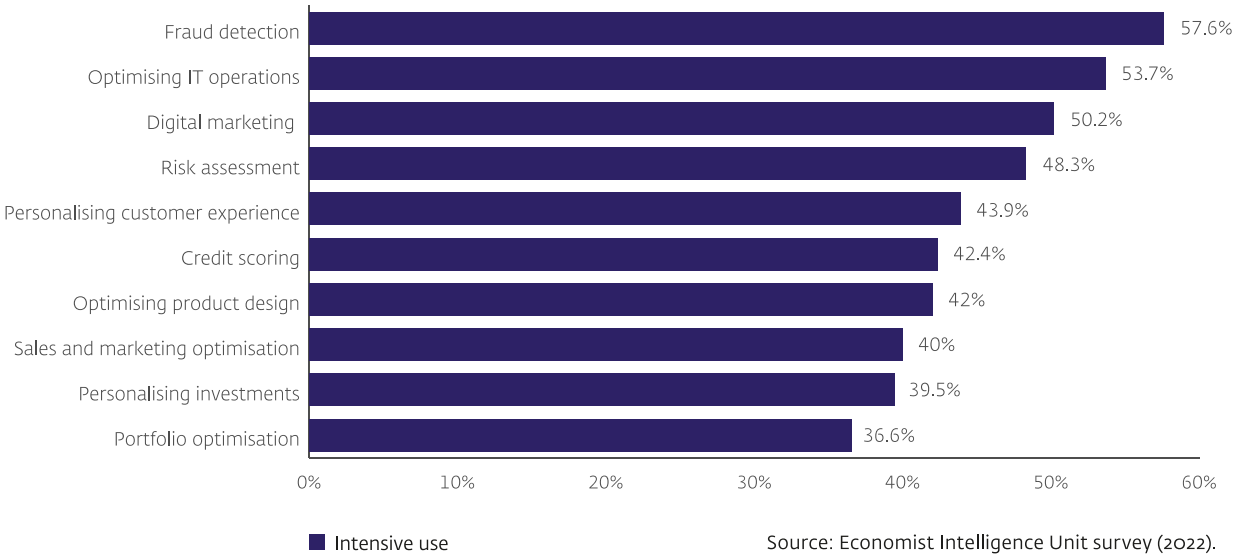
**AI is rapidly advancing in the financial sector worldwide.** A survey conducted by The Economist in 2022 found that banks worldwide mainly use AI to detect fraud (58% using it intensively and 32% to some extent.[2] Almost all the banks in that survey said they use AI in various areas or intend to do so during the next three years (see also Figure 2). The use of AI is also advancing rapidly in the insurance industry. An American survey in 2023, for example, found that 88% of car insurance companies and 58% of life insurance companies were using or intended to use AI.[3] On the capital markets, AI can be found in the machine learning models used for algorithmic trading.[4]

**Financial institutions in Europe are also making full use of AI.** An ECB survey in 2022 found that 60% of the 105 major European banks are using AI.[5] It found that AI was being used mainly in chatbots, credit rating and algorithmic trading.[6] European insurance companies are also using AI. A 2023 survey by EIOPA, the European Insurance and Occupational Pensions Authority, showed that AI was being used by half of the respondents in the non-life insurance sector and by 24% in the life insurance sector. European insurance companies are using AI mainly for customer service (41%), fraud detection (39%) and claims handling (37%). Chatbots are the most common use. AI is also being used to make cross-selling recommendations and improve product pricing.[7] The AI models currently being used are mainly simple, explainable ones, but insurance companies also expect to start using more complex models in the near future.

2 The Economist Intelligence Unit (2022), Banking on a game-changer: AI in financial services
3 Insurance Newsnet (2023), 58% of life insurers use artificial intelligence or are interested: NAIC survey
4 AFM (2023), Proprietary trading firms use machine learning on a large scale in trading algorithms
5 ECB (2023), Supervising the future of banking: navigating the digital transformation
6 ECB (2023), Take-aways from the horizontal assessment of the survey on digital transformation and the use of fintech
7 EIOPA (2024), Report on the digitalization of the European insurance sector (to be published later this year)

## Figure 2 Areas in which AI is being used by banks worldwide

The percentage of banks making intensive use of AI in a particular area




Source: Economist Intelligence Unit survey (2022).

**Dutch financial institutions say that they have been using systems and models involving AI for some time now, and are experimenting with more advanced models.** In the ECB survey almost 60% of the significant Dutch banks stated that they are using one or more AI models. When interviewed, most Dutch banks said that they are mainly still exploring the possibilities, which suggests that AI use will increase in the years to come.[8] The Dutch banks that are currently using AI use it for creditworthiness assessments, fraud prevention and to combat cyber crime and financial crime. Natural language processing (NLP) is used to improve the quality of chatbots; it also helps staff to work more efficiently, for instance by summarising telephone conversations with customers automatically. The EIOPA survey showed that AI was being used by 56% of the Dutch respondents in the non-life insurance sector and by 50% in the life insurance sector. It found that AI was most commonly used in chatbots and for targeted online marketing and fraud detection. Most insurance companies expect that their use of AI will increase in the next three years.

**Dutch institutions recognise the potential of generative AI, but many of them are cautious about using it at the moment.** This emerges from a series of interviews with Dutch financial institutions. Many financial institutions ban the use of the public version of ChatGPT internally in order to ensure that confidential internal data does not end up on external servers, where it could be used e.g. to train the model in more detail. Some institutions do use a version that does not post the data input on a public server. Also, some institutions are gradually starting to use AI for support processes, as confirmed by OECD research. For instance, according to

8 For the purposes of this report we spoke to banks, payment institutions, pension providers and insurance companies.

the OECD, financial institutions are currently experimenting mainly with offline versions of LLMs that are used to support internal processes and activities, such as doing translations and searching for information.[9] Many banks and insurance companies say that recent developments in the field of generative AI have been more rapid than expected. At the same time, many institutions are currently cautious about using it themselves, as it causes a rapid increase in complexity. One bank, for instance, said that 'everything related to generative AI is banned', and a pension administration organisation said that generative AI did not yet fall within the scope of the AI development teams. New use cases are expected to be discovered in the future that financial institutions could include in their day-to-day operations.[10] In this case their attitude to generative AI could change.

**Increasing numbers of asset management companies are using AI for developing investment strategies, risk management and compliance.** This was found in a study by the European Securities and Markets Authority (ESMA).[11] Only a few asset management companies have developed a completely AI-based investment process so far. AI can help investment firms to optimise trading and post-trade processes, while reducing the market impact of large orders and the numbers of handling errors.[12] The AFM examined the use of machine learning in trading algorithms by the four major proprietary trading firms and found that they often use machine learning models, mainly supervised learning models at the moment. These models predict the price of a financial instrument that is being traded. The complexity of these models, however, presents risks: they must therefore be explainable and properly verified.

**Some institutions say that they are already using more AI in other countries, e.g. the United States.** Lack of scale in Europe, hence lack of adequate data to train AI models, would seem to be a factor here. Another explanation given is the more stringent privacy legislation. Uncertainty about regulation would also seem to be involved in institutions' caution about using AI in Europe, and it is thought that risk management could be more thorough in Europe than in the United States. Proposed AI use cases often need to be approved by the management board, even those intended solely to make non-essential repetitive tasks more efficient.

**Financial institutions say that they prefer to keep their AI models as simple as possible.** Institutions weigh up model performance against explainability. The more risky the activity, the more likely the outcome is to be in favour of a simple model, even if more advanced models produce better results, as explainability is particularly important in such cases. Some complex AI models are used to only a small extent or not at all because of uncertainty regarding ultimate supervision. Several institutions say that they are fundamentally against black box (e.g. deep learning) models and that they will not be using them in future. Whether they will still think the same way in a few years' time is debatable. Reasons of competition, for instance, might lead

9 OECD (2023), Generative Artificial Intelligence in Finance
10 Ibid.
11 ESMA (2023), Artificial intelligence in EU securities markets
12 Ibid.

them to start using more advanced AI in future, as too much caution might make them unable to compete with foreign financial institutions and non-financial players.

**Institutions regard it as important to keep a 'human in the loop'.** AI models or systems are therefore not always completely automated: an employee of the financial institution assesses the AI-generated result before it is finally passed and used. The employee (and ultimately the management) remains responsible for the content. Many institutions, for example, have a Chief Technology Officer (or similar) on their board, who is responsible for matters relating to innovation and digitisation.

**In order to limit the risks associated with using AI, existing risk management frameworks are used, in some cases supplemented by specific AI frameworks.** Institutions are accustomed to using models, and they can also use their existing model risk management frameworks for AI models. Some institutions have drawn up specific requirements for AI models: one insurance company, for instance, uses model score cards that pose a number of questions for each model. These could relate e.g. to the purpose of the model, what data it uses, and how it can be prevented from discriminating.

**Financial institutions say that they regard the ethical use of AI as very important, and they are not trying to push the boundaries of what is permitted and possible under the current and forthcoming regulation.** The majority of the institutions have not set up specific internal AI ethics committees to deal with ethical dilemmas; instead they use existing ethics committees to deal with AI-related matters. Institutions in various sectors say that they need shared definitions and interpretations of ethical principles (fairness, explainability, etc.) in order to provide more guidance on grey areas. To this end, insurance companies are already using the Dutch Association of Insurers' Ethical Framework,[13] which asks them to carry out a number of additional checks when using AI.

## 1.3 Opportunities offered by AI in the financial sector

**AI offers opportunities for both customers and financial institutions themselves.** We will first consider the opportunities that AI offers for better, faster customer service and then the possibilities that it offers for the financial sector to generate higher revenue and cut costs.

### Better, faster service

**AI can improve customer service by enabling more personalised products and faster service.** The more efficient use of data enables products and services to be offered that are more tailored to individual customer needs. AI also has the potential to speed up certain steps in various customer interactions (e.g. loan and insurance applications), resulting in faster service.

### Higher revenue

**AI can help financial institutions to tap into new customer groups.** It can e.g. enable risk assessments to be carried out on customers who have a limited credit history or none.[14] If this results in products such as consumer credit or insurance being able to be offered responsibly

13 Dutch Association of Insurers (2020), Ethical Framework
14 OECD (2023), Generative Artificial Intelligence in Finance

to people who were not previously eligible, it improves financial inclusion. Whether new customer groups can be acquired, and if so how, depends on whether the financial institutions have sufficient suitable data and whether they are able to use models effectively. Privacy and ethical considerations are also involved.

**AI can help to increase cross-selling opportunities.** Existing customers can be reached better, and services can be sold in a more targeted manner. AI can help e.g. to predict what product a customer will buy next, or to gauge how likely a customer is to buy a particular product. McKinsey outlines how fintechs in Asia, including banks operating solely online, use information such as behaviour, bank history and customer satisfaction to produce a profile that predicts customers' future actions and goals.[15] AI can also be used in the insurance sector to optimise sales channels and create new marketing channels.

**AI can help institutions to optimise their prices.** Insurance companies, for example, can use AI for behavioural pricing, i.e. to set premiums based on customer behaviour, which can help them to gauge or reduce expected claims and maximise their profits. It can also produce a perception of a fairer market, as risk-averse consumers bear less of the expense for the high-risk behaviour of others.[16]

While pricing strategies of this kind are not new, insurance companies using AI can focus more on behavioural pricing by using new data – often a combination of 'new' internal data (e.g. by analysing clicking and closing information from the website) and external data.[17] Although it provides opportunities, price personalisation also presents risks (e.g. of undermining solidarity, discrimination and exclusion: see ∫1.4).

### Lower costs

**AI can make processes more efficient, enabling financial products and services to be offered at lower costs.** AI helps financial institutions to increase the automation of processes and speed them up. It also offers opportunities to enhance IT operations by integrating various data sources, and by making the building and testing of software systems more efficient.[18] Financial market traders can use AI to reduce friction and increase efficiency and trading speed. Trading algorithms can identify potential trading opportunities and act on them without human intervention. AI helps to manage risks and order books better so as to streamline execution and work more efficiently.[19] Credit rating agencies, proxy advisory firms[20] and other financial market players can also use AI to support and optimise certain activities: NLP, for instance, can help to provide easy access to publicly available documents systematically to support analysts.[21]

15 McKinsey (2022), Building a winning AI neobank
16 AFM (2021), Personalising prices and terms and conditions in the insurance sector
17 AFM-DNB (2019), Artificial intelligence in the insurance sector, a fact-finding study
18 TU Delft (2023), The link between artificial intelligence (AI) and software engineering
19 OECD (2021), Artificial Intelligence, Machine Learning and Big Data in Finance Opportunities, Challenges and Implications for Policy Makers
20 Proxy advisory firms are organisations that advise institutional investors (investment funds, pension funds and insurance companies) how to vote at shareholders' meetings.
21 ESMA (2023), Artificial Intelligence in EU securities markets

**AI can help financial institutions to gauge risks better and more consistently, provided they have good-quality and complete data, as well as models that are explainable and verifiable.** Although structured data has always been used, data used as input to AI models helps to carry out what is known as 'sentiment analysis' and to create additional insight based on pattern recognition.[22] Whereas people can be inconsistent in their risk assessments, AI models are not. It should be noted, however, that AI models can contain systematic errors that produce incorrect results (for more information, see § 1.4).

**AI can also be used to support environmental, social and governance (ESG) investment.** Investors can use natural language processing to carry out targeted ESG assessments based on companies' corporate social responsibility reports.[23] DNB has also used NLP to study pension funds' awareness and implementation of sustainable investment.[24]

**AI can make combating financial and economic crime more effective and efficient by improving assessments and reducing the administrative burden on institutions and customers.**[25] Customer surveys are time-consuming and labour-intensive for financial institutions and their customers. AI can provide a solution by reducing the administrative burden on customers while at the same time improving risk analysis even more. According to the European Banking Authority (EBA), AI-based models have the potential to identify suspicious actors and activities better than non-AI models.[26] Transaction monitoring can also be improved using AI, and collaboration such as that currently taking place between five Dutch banks under the name Transaction Monitoring Netherlands (TMNL) makes this process even more effective. Combining transaction data from various banks' business customers enables correlations to be made that individual banks cannot make. It is not possible yet to analyse private transactions, as proper statutory privacy safeguards need to be introduced first.

**AI can improve defences against cyber attacks.** AI can help to detect weak links in a complex system, so that changes can be made to increase cyber security. AI techniques such as machine learning and deep learning will play an increasing role in AI-driven cyber security. These techniques enable certain cyber security processes to be automated and be 'smarter' than the current systems. Although AI can improve defences against cyber attacks, at the same time it can result in attacks becoming more sophisticated (see next section).

## 1.4 Risks of AI in the financial sector

**AI can yield significant benefits in the financial sector, but these go hand in hand with a variety of risks that can have an impact on financial institutions and their customers.** These risks thus fall within the scope of AFM and DNB supervision of financial institutions and financial stability, focusing on managing risks within institutions.

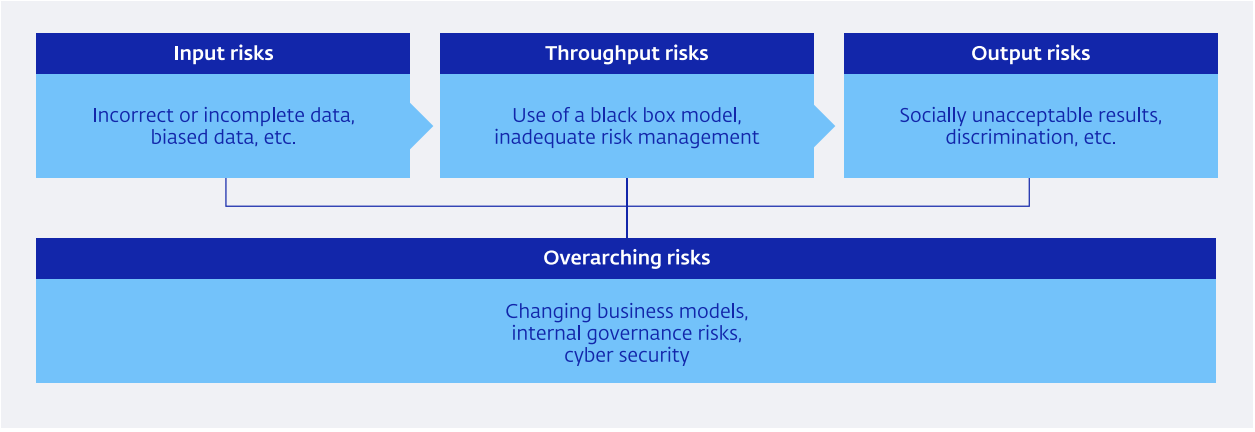
22 OECD (2021), Artificial Intelligence, Machine Learning and Big Data in Finance
23 ESMA (2023), Artificial Intelligence in EU securities markets
24 Rob Bauer, Dirk Broeders, Annick van Ool, DNB Working paper 770, Walk the green talk? A textual analysis of pension funds' disclosures of sustainable investing
25 DNB (2022), From recovery to balance
26 EBA (2020), Final Report on Big Data and Advanced Analytics

**Different risks occur in the various phases of an AI model.** To clarify these risks, it is useful to divide the AI process into three phases. An AI model needs data, which is selected and fed into the model: this is the input phase. The selected data is then subjected to particular operations and processing in the AI model: this is the throughput phase, also referred to as 'the system'. Following this processing in the system, the model produces results: this is the output phase. Particular risks occur in each of these three phases; there are also overall risks that occur in all three phases (see Figure 3).

## Input risks

**AI models are trained on large quantities of data: where data is used in risk management or in dealing with customers, the results of the AI models are only as good as the data that they use.** The principle here is 'garbage in, garbage out': if an AI model is trained on incorrect or incomplete data, the results will also be incorrect or incomplete. An important prerequisite for using AI systems efficiently and effectively is ensuring good data quality and modernising institutions' older legacy systems. The financial regulators have been drawing institutions' attention to shortcomings in their data quality for some time now.[27] A particular problem with AI applications is that data can be biased: using historical data can result in the model copying existing inequalities and historical biases. For example, when using historical job application data to select future applicants, if that data is biased, certain sections of the population may be discriminated against. For instance, an AI model used by the tech company Amazon to select applicants based on historical data was found to have taught itself to practically exclude women.[28] AI applications therefore need to be validated regularly, so that changes can be made promptly when risks of this kind occur.

## Figure 3 The risks of the AI phases




27 DNB (2020), Transforming for trust; ECB (2020), ECB report on banks' ICAAP practices; DNB (2022), Good Practice: Robuuste pensioenadministratie (Good practice: robust pension administration) (in Dutch)
28 BBC News (2018), Amazon scrapped 'sexist AI' tool. Applications of this kind, however, are not covered by financial supervision.

**Using data also involves concerns regarding privacy and data protection. Data use must be duly substantiated, and data holders are responsible for protecting data properly.** Financial institutions potentially have a good deal of information on their customers, from the purchase of a new car or home to paying for the daily shopping at the supermarket. Under certain conditions this information can be used to develop an algorithm, or customer data can serve as input to an existing algorithm. Also, privacy safeguards are important if large quantities of personal data are being processed. An institution's data governance must therefore comply with the requirements of the General Data Protection Regulation (GDPR). This can create tension between the predictive power of an algorithm and the importance of a responsible, controlled, sound process in which customer privacy is properly safeguarded.

## Throughput risks

**Systematic errors in the AI model, known as 'model bias', can produce undesirable or incorrect results.** Model bias can occur if the right training data was not selected or the model parameters were incorrect, with the result that the algorithm is trained wrongly. Even if new, better data is fed into the same model, the same systematic errors may be repeated over and over again because of the way in which the model was trained, possibly producing undesirable or incorrect results. In order to obviate this risk, AI models should be enhanced using feedback loops. For example, the output can be monitored against standards and the results tested for identified risks. The required changes must then be made and the model refined. If this is not done, models are at risk of repeating the same mistakes ad infinitum, or errors may be increasingly self-reinforcing.

**Implementing AI in financial services can cause an explainability problem.** An AI model will often comprise a network of multiple interconnected layers and may have hundreds of parameters that contribute to producing a result: it will therefore not always be possible to trace results of AI models back to a single parameter, so that it can be explained to a customer why a particular decision was made. For example, if a credit or insurance application has been rejected and the customer is entitled to be given a reason for this rejection, it may be difficult to explain the AI model's decision to the customer because there is no clear data point that explains it. Using black box models of this kind can be at the expense of transparency and explainability. Scientists are working on technical solutions to improve the explainability of AI models.

## Output risks

**Institutions can use AI applications that produce socially undesirable outcomes.** For instance, a biased model can cause the undesirable financial exclusion of particular market segments, and the use of biased data can cause discrimination against particular sections of the population. This can be the case with AI models that assess creditworthiness. If they have been trained on historical data in which particular groups had less access to credit, they could perpetuate that bias by giving lower credit ratings to customers in those groups, possibly resulting in financial exclusion and unequal access to financial services.

**Using AI to personalise pricing can affect market transparency and solidarity.** AI-based pricing often requires large quantities of customer data, which can result in customers unconsciously 'paying' for personalised prices with their data, thus raising issues of privacy and ethics. It can

also reduce market transparency, making it more difficult for customers to compare products. In the insurance sector, personalising premiums can prejudice the principle of solidarity. This principle (that all insured persons should contribute to the shared costs) can come under pressure if premiums are calculated entirely on an individual basis. This can result in situations in which some customers pay far more than others, thus undermining the sense of solidarity and making particular groups potentially uninsurable.[29]

## Overall risks

**An important overall risk of using AI is a possible lack of fairness.** In the public debate on the risks of AI systems, requirements such as 'fairness' and 'responsible use' of AI are often mentioned. Using an AI model, for example, should not result in the unintentional unequal treatment of particular individuals or groups. A problem here is that there is as of yet no consensus in society as to what terms such as 'fair AI' or 'responsible AI' precisely mean and how institutions should modify their AI applications to comply with them. Institutions usually develop their own policies in this connection, based on their varying interpretations, hence there is a need to lay down uniform frameworks jointly.

**An inadequate governance framework in institutions can result in inefficiency and risks being overlooked.** The internal governance of financial institutions plays a vital role in mitigating the risks of AI. Proper strategic organisation of risk management is important for controlled operational management and ensures that the risks of AI systems can be identified promptly and managed. Standardised documentation of the AI models used in a financial institution, taking the potential risks into account, can be useful here. It is also important to use employees with adequate AI know-how, who are familiar with the content of the areas in which AI is being used.

**There are risks involved with using AI, e.g. as it can result in reputational damage and incorrect risk assessments, but there are also drawbacks if the institution does not take advantage of the potential of AI.** AI applications change financial institutions' business models and work processes. A risk of AI is that AI applications can undermine the profitability of an institution if they cause reputational damage or produce incorrect predictions and recommendations, e.g. by wrongly assessing credit risks. In addition, replacing staff who have specific expertise with AI models can result in human expertise no longer being available if the AI model does not perform satisfactorily or an AI model is banned or replaced. Even if an institution does not use AI, there are risks to financial institutions' business models: by not taking advantage of the innovation potential of AI, institutions can become less competitive.

**In addition to the risks of an institution implementing AI applications, there are risks due to the use of AI throughout the sector, including dependence on third parties.** AI systems implemented online can unintentionally come into contact with each other, resulting in unfair trading practices due to algorithmic collusion.[30] Another risk arises from multiple financial institutions using the same purchased AI software at the same time, making them all dependent on the same provider. Market concentration of this kind means that risks can occur on a large scale: for example, the massive purchasing of OpenAI GPT applications as a result of the heavy publicity around ChatGPT in late 2022. Certain potential shortcomings

29 DNB (2023), Insurers in a changing world
30 AFM (2023), Algorithmic collusion in capital markets, an in-depth analysis

or inaccuracies of Large Language Models can therefore affect many institutions at the same time, possibly resulting in system risks. In addition, dependence on a – currently small – group of AI technology providers can present risks. If the suppliers of this technology are no longer able for any reason to provide good services (or any services), it is not possible to simply switch to another supplier, and financial institutions using that technology will have problems, especially if they use AI in their core processes. The fact that this technology is only available from mainly non-European providers also makes for a geopolitical dimension.

**The use of generative AI, including LLMs such as ChatGPT, presents some additional risks due to the dissemination of inaccuracies and disinformation, among others.** The content generated can contain inaccuracies (referred to as 'model hallucinations') and thus result in the dissemination of wrong information. The use of online, publicly accessible AI applications can also result in confidential internal data ending up on external servers, hence it is important for generative AI to be used solely in secure environments. Other types of generative AI include deep fakes and voice generators that enable counterfeit images and audio to be used to cause reputational damage to institutions. From the point of view of the regulators, there is also the risk of the reports and forms submitted to them by institutions being AI-generated, making it impossible to determine how accurate or precise they are. It must be clear that the use of AI should not remove personal accountability for an institution's documentation or other actions.

**Cyber risks can also occur in various phases of the process.** Cyber attacks by hackers or penetration attacks can cripple whole systems, steal sensitive information or manipulate output from a model. AI models can also be trained to carry out more effective cyber attacks. The rise of generative AI has brought with it new types of crime and fraud: for example, scammers have started using AI voice distorters (e.g. to simulate a trusted person's voice) in order to ascertain customers' passwords and PIN codes. Employees of financial institutions need to have the requisite know-how to recognise false messages and incorrect information.

**The power and cooling water consumption of AI could in the future account for a substantial proportion of worldwide consumption and become a major cost factor for financial institutions.** This gives rise to both prudential concerns regarding dependence on volatile, unstable electricity prices and integrity concerns regarding climate policy and the sustainability of the sector.

## 1.5 Scenarios for the future of AI in the financial sector

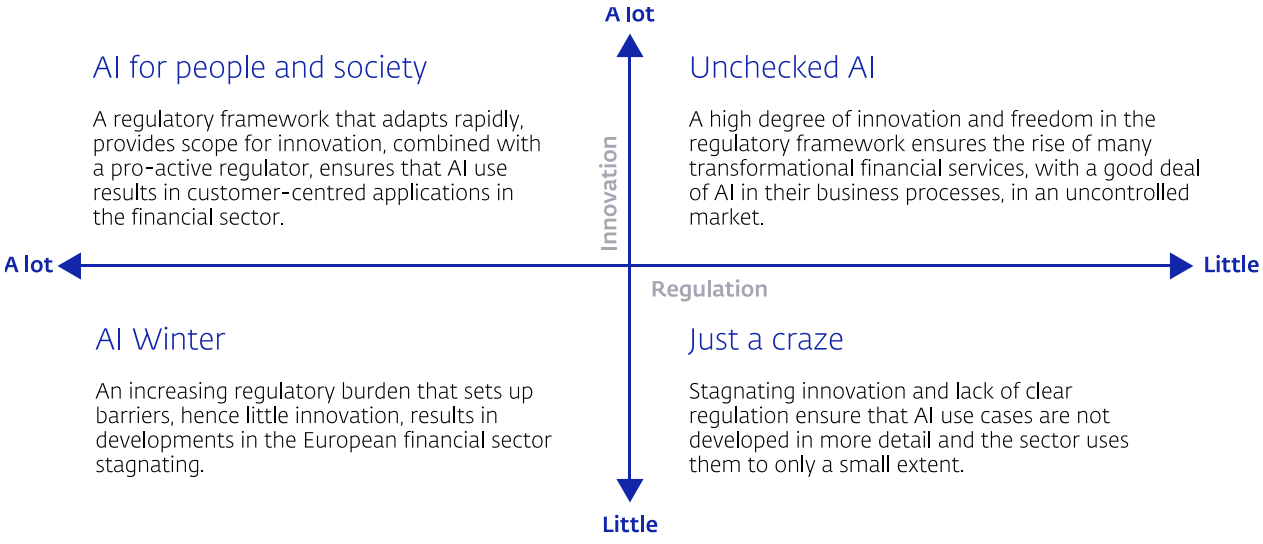
**A variety of scenarios for the future of AI in the financial sector have been drawn up to provide a picture of the possible directions in which this technological innovation could develop.** What impact AI will have on the financial sector in the years to come is still difficult to gauge. A scenario analysis can provide a more concrete outline of the possible changes that are expected in the financial sector and the circumstances that could lead to those changes. The scenarios are driven by two factors related to the use of AI: (a) the extent to which AI applications will be regulated and how that regulation takes shape, and (b) the extent to which innovative financial services are developed using AI. Regulation can keep AI on the right track, but it can also present barriers, possibly

preventing innovation; lack of innovative capacity in the financial sector could also be a factor. Innovation is brought about mainly by market forces, but regulation that inspires confidence can encourage adoption. The resulting four scenarios show possible futures for the financial sector (Figure 4).

**The scenario analysis underscores the importance of balance.** A scenario in which a balance is achieved between growing innovation and adaptive regulation seems to offer the best prospects. The scenario 'AI for people and society' would yield customer-centred applications and healthy growth for the sector. Conversely, scenarios in which there is an imbalance between regulation and innovation seem to produce less good results. The scenarios that combine heavy regulation with little innovation show that regulation can create barriers, resulting in an 'AI winter'. Another lopsided scenario is 'unchecked AI', with a lot of innovation and little regulation to manage the risks in an uncontrolled market. Both cases of imbalance could have an adverse impact on controlled operational management and customers' interests.

**The scenario analysis emphasises the importance of a pro-active, adaptive attitude on the part of policy-makers and regulators.** Regulation needs to create a framework for responsibility and transparency that at the same time provides enough scope for rapid developments in AI products and services. This is a challenge that needs to be tackled by both policy-makers and innovative institutions, so as to create a future in which AI benefits both the financial sector and society as a whole.

## Figure 4 Scenarios for the future of AI in the financial sector



# 2 AI supervision and legislation

This chapter discusses the supervision of financial institutions' use of AI. The financial soundness and integrity of institutions, safeguarding customers' interests, and the integrity of relationships between market players are paramount concerns in AFM and DNB supervision, and that is no different when AI is used (§2.1). The existing statutory financial supervision framework can therefore be applied to financial institutions' use of AI systems (§2.2). Cornerstones of financial supervision such as proper risk management and safeguarding customers' interests provide a basis for shaping AI supervision (§2.2.1). Existing standards need to be interpreted in the new context in which AI is being used. For the sake of things such as regulatory certainty, it may be desirable to clarify what is expected of an institution in regulations and policy statements, as in the requirements for algorithmic trading on the capital markets (§2.2.2). The use of AI also requires wider statutory protection of values and fundamental rights by the European Union, as we are seeing in the development of the European AI regulation. §2.3 explains how the AFM and DNB regard the application of the regulation to the financial sector. Given the relationship between AI and large data sets, the ethical issues and the risks to fundamental rights, we also consider the role of other regulators (§2.4).

## 2.1 Objectives of AI supervision

**The existing objectives of financial supervision apply in full to financial institutions' use of AI systems.** AI systems must not jeopardise the financial soundness and integrity of financial institutions, nor must the use of AI harm customers' interests or the integrity of relationships between market players. Innovative systems that use AI can support compliance with standards of prudence, integrity and conduct. Moreover, innovation is important for the competitiveness of financial institutions and thus for a healthy financial sector. Supervision should give sufficient room for innovation, provided the risks are adequately managed. There also needs to be clarity regarding the application of regulations and how institutions are required to demonstrate that their technologies are compliant.[31] This should preferably be done at European level, for the sake of regulatory convergence and a level playing field.

**AI will have consequences for the way in which the AFM and DNB supervise the sector.** The AFM and DNB will need to expand their knowledge in this area to enable them to assess institutions' use of AI. Regulatory methods and procedures will need to be changed or developed in certain cases to take account of new technologies. Supervision will focus inter alia on deciding about and managing the risks of using AI applications, the modalities of those applications, and the

31 Bonnin Roca, J. & O'Sullivan, E. (2022), The role of regulators in mitigating uncertainty within the Valley of Death, Technovation, January.

output. Various aspects could be considered here, including HR, data and IT capacity, in order to determine whether financial institutions have sufficient in-house expertise to use AI. Model governance and model risk management are also areas of attention, as are policies on the scope for disregarding AI results, since staff (and ultimately senior management) remain accountable for the results and should be able to modify them whenever necessary. Increased attention will also be paid to strategic decision-making when competition shifts and (global) competitive pressure mounts, since this could affect the ethical use of AI. Regulators will also need to assess whether the regulatory framework gives them sufficient tools to exercise their supervision. The expectation is that standards will need to be clarified. With the entry into force of the AI regulation, supervisory authorities will need to be designated, and the interaction between the current regulatory framework and the regulation will need to be clarified (see also 2.3).

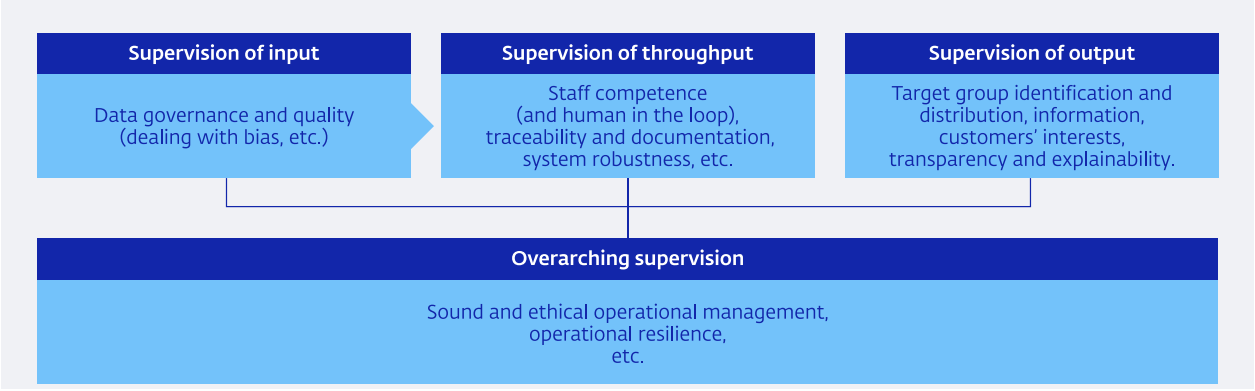
## 2.2 The statutory framework and AI supervision

**Financial supervisory requirements may be linked to the process and risks of an AI model, as introduced in Chapter 1.** Figure 5 shows the various elements in the AI process mentioned above and outlines how they relate to important elements of regulations. The ensuing subsections explain the supervisory requirements for various phases of AI applications in the statutory framework.

### 2.2.1 The regulatory framework in general

**The supervision of AI applications is being stepped up.** Financial institutions have used systems and models involving AI for some time now. The use of AI is part of existing processes and services supervised by the AFM and DNB under the legislation and regulations in force. How specific the supervisory activities are, and what requirements are laid down, depends on the particular application. Supervisory attention

Figure 5 Various phases of AI applications and supervision



is determined by the importance of the processes for which AI is being used and the risks that could occur. It is important here for the standards to be sufficiently clear. As the use of AI and its importance to the financial sector increases, the AFM and DNB will step up their supervision of the way in which institutions manage the risks of AI use. This year, for instance, DNB is carrying out a thematic examination of the use of AI by insurance companies.

**Cornerstones of financial supervision are provisions relating to integrity and risk management.** For instance, institutions must comply with the requirements of sound and ethical operational management, overarchingly applied to AI systems. Special attention needs to be paid to managing risks arising from shortcomings in the data (input) and the models used (throughput) when AI is being deployed. For the sake of sound and ethical operational management, institutions need to avoid laws being contravened and social norms being breached, which could jeopardise trust in the financial firm or the financial markets. It follows that institutions must ensure that the use of their AI systems does not result in discriminating or acting in breach of other fundamental rights or important social norms (output). They need to manage their business processes and business risks, as well as their financial risks and other risks that could affect the soundness of the institution. In other words, financial firms need to set up proper risk management. The systems used by financial institutions to gauge the risks of money laundering and terrorism financing, and which are increasingly using AI technologies, must also comply with the requirements of sound and ethical operational management.

**How risk management is organised will depend on the nature of the firm, and the details will often be specified in sectoral regulations,** which will often be based on European legislation and implemented in the Financial Supervision Act (Wet op het financieel toezicht – Wft): for example, the European Capital Requirements Directive (CRD) and Regulation (CRR). These list certain risks to which banks are exposed (e.g. credit risk and market risk) and lay down requirements for their management and the organisational embedding of risk management. Similar provisions on insurance companies are set out in the European directive for insurance companies (Solvency II). Another example is the rules for investment firms (MiFID II), including the consideration of particular market risks in order to ensure orderly trading. Generally speaking, when using AI, an institution needs to manage the associated risks in that context properly. Where necessary, what exactly is expected of institutions should be specified at EU level.

**Standards apply not only to operational management but also to safeguarding customers' interests.** These relate e.g. to the duty of care, product development and distribution, and to preventing excessive borrowing. Regardless of whether a product or service has been created using AI, the customer's interests need to be taken into account properly in the advice given, the customer must not be allowed to borrow excessively, and the target group must be defined appropriately before a product is distributed. It is important, *inter alia*, for AI applications to encourage consumers to make decisions (consciously or unconsciously) that benefit their financial wellbeing.[32] Also, dynamic pricing must not be used at the expense of market transparency.

32 AFM & DNB (2019), Aandachtspunten AFM en DNB bij artificiële intelligentie in de verzekeringssector (AFM and DNB areas of concern regarding artificial intelligence in the insurance sector), p. 28 (in Dutch).

**Financial institutions are expected to use AI responsibly.** The more important the use of AI becomes in a financial institution's decision-making, and the greater the potential consequences for the institution and its customers are, the higher the bar for the responsible, explainable use of AI is set. The AFM and DNB will take this into account when supervising financial institutions. They have stated in previous publications what aspects are relevant (see e.g. the joint report 'AI in de verzekeringssector' (AI in the insurance sector).[33] The forthcoming AI regulation (see §2.3) provides a framework of standards for the responsible use of AI.

### 2.2.2 Specific regulations for AI

**Specific references to algorithms, AI or AI systems in financial legislation are limited.** AI is explicitly included in sectoral financial legislation in several areas, for example, the rules on algorithmic trading on capital markets (MiFID II/MiFIR), specific requirements for the use of AI in consumer credit (in the new Consumer Credit Directive, CCD2), and rules on automated advice. These specific standards for AI show that particular requirements may be laid down, depending on the type of application. The regulatory framework for banks' internal models (under the CRR), and to some extent that for insurance companies (under Solvency II), contain many specific requirements with which AI models also need to comply. The expectation is that specific rules will be laid down or revised in other areas in the future.

#### Algorithmic trading

**The rules on algorithmic trading on the capital markets[34] lay down requirements for safeguards when using AI and prohibit it from being used in a way that contributes to or could result in an unorderly market.** These relate inter alia to the resilience of trading systems and adequate trading capacity, including appropriate trading thresholds and limits. The details of the requirements are set out in regulatory standards laid down at EU level. A selection from those standards shows that systems and trading strategies need to be tested thoroughly, staff need to have an adequate understanding of the systems, and pre- and post-trade standards apply, along with real-time monitoring requirements. Requirements of this kind relate to the responsible use of the AI system (see Figure 5).

#### Consumer credit

**The revised Consumer Credit Directive 2 (CCD2)[35] provides guidance on regulating the use of AI for creditworthiness assessment.** The requirements apply mainly to the data (input, see Figure 5) that is or is not used, and transparency regarding automated services (output). If the price of the product or service being offered has been set using AI (automated processes, including profiling), the consumer should be informed, as this will enable them to take the potential risks into account when making their purchase decision.[36] Creditors and intermediaries are also required to inform consumers clearly and comprehensibly if an offer has been personalised based on the automated processing of personal

33 AFM & DNB (2019), Aandachtspunten AFM en DNB bij artificiële intelligentie in de verzekeringssector (AFM and DNB areas of concern regarding artificial intelligence in the insurance sector) (in Dutch). See also DNB (2019) General principles for the use of Artificial Intelligence in the financial sector.
34 In this context see e.g. MiFID II, Articles 17, 18, 26, 27, 45, 47 and 48.
35 Directive - EU - 2023/2225 - EN - EUR-Lex (europa.eu) The final text of CCD2 was published on 30 October 2023. The directive must be transposed into national legislation by 20 November 2025, and the rules will be applied from 20 November 2026.
36 See also Recital 46 of the CCD2.

data (the transparency requirement, processes and safeguards). In addition, Article 18 of the CCD2 restricts the types of personal data that can be used (regardless of whether AI is involved) and lays down a right to human intervention if the creditworthiness assessment includes the use of automated processing of personal data.

## Automated advice

**The rules on automated advice[37] also relate to the safeguards required in the environment in which a system generates automated advice.** They clarify how advice rules – originally designed for physical situations with the adviser and the customer sitting around a table – should be interpreted in an automated environment. In terms of Figure 5, the requirements include system supervision, incorporating professional expertise requirements in the system, and testing output with the aid of scenario analyses. The rules lay down inter alia that anyone designated as responsible for automated advice must also have the necessary professional expertise to advise on the financial product concerned themselves. The automated system must be tested before being taken into service. The financial service provider must also check regularly whether the automated advice is consistent with and in line with the information supplied by the consumer. Lastly, the rules lay down what action must be taken if automated advice is found not to be compliant with the statutory rules on advice.

## Banks' and insurance companies' internal models

**In order to determine their capital requirements, subject to certain conditions, banks can use the 'internal ratings-based (IRB) approach', e.g. to determine their credit risk.** Banks can use IRB models to gauge their credit risk, possibly yielding a lower capital requirement than would be needed using the standard method. AI, in particular machine learning, may provide a more accurate estimate of the risk than the traditional models. IRB models must comply with an extensive set of specific requirements, and institutions using AI must remain within those limits. Although not designed with the use of AI in mind, these requirements apply to the various phases of AI applications shown in Figure 5. One requirement, for instance, is that the classification of obligors in similar risk categories must be consistent and understandable to third parties (CRR, Article 171). Using AI, classification can in principle be highly refined, but it can be difficult to demonstrate consistency and ensure that the result is explainable to third parties, hence it is impossible to use AI as a matter of course. Another example is that projections made using models must be 'plausible and intuitive' (CRR, Article 179). AI can sometimes produce results that are not immediately intuitive but nevertheless turn out to provide a better estimate of credit risk exposure than a traditional approach.[38] The framework will therefore need to be clarified or possibly modified to enable particular types of AI to be used.

**Similar (albeit somewhat less detailed) provisions apply to insurance companies' use of internal models under Solvency II.** In particular, the use of advanced machine learning models can present problems, particularly in the areas of explainability and model validation. For instance, it must be possible to explain sensitivities to changes in underlying assumptions, and any independent outside expert must be able to

37 Decree on Business Conduct Supervision of Financial Enterprises (*Besluit gedragstoezicht financiële ondernemingen*), Section 32da This section has been incorporated in the Financial Markets (Amendment) Decree (*Wijzigingsbesluit Financiële Markten*) 2023 and enters into force on 1 July 2024.
38 For details see EBA (2021), Discussion paper on machine learning for IRB models

understand the design and operational details of the internal model.

## 2.3 The AI regulation

**The AI regulation is a cross-sector European regulation designed to protect important EU values and fundamental rights such as safety and non-discrimination.** The application of the regulation is based on the definition of AI and the distinction between 'unacceptable risks', 'high risks' and 'low or minimal risks'. There will also be rules on general-purpose AI models. Figure 6 shows this distinction between risk categories in diagrammatic form. Agreement was reached on the AI regulation in December 2023, and it is expected to be published before summer 2024, after which most provisions will have a two-year implementation period.

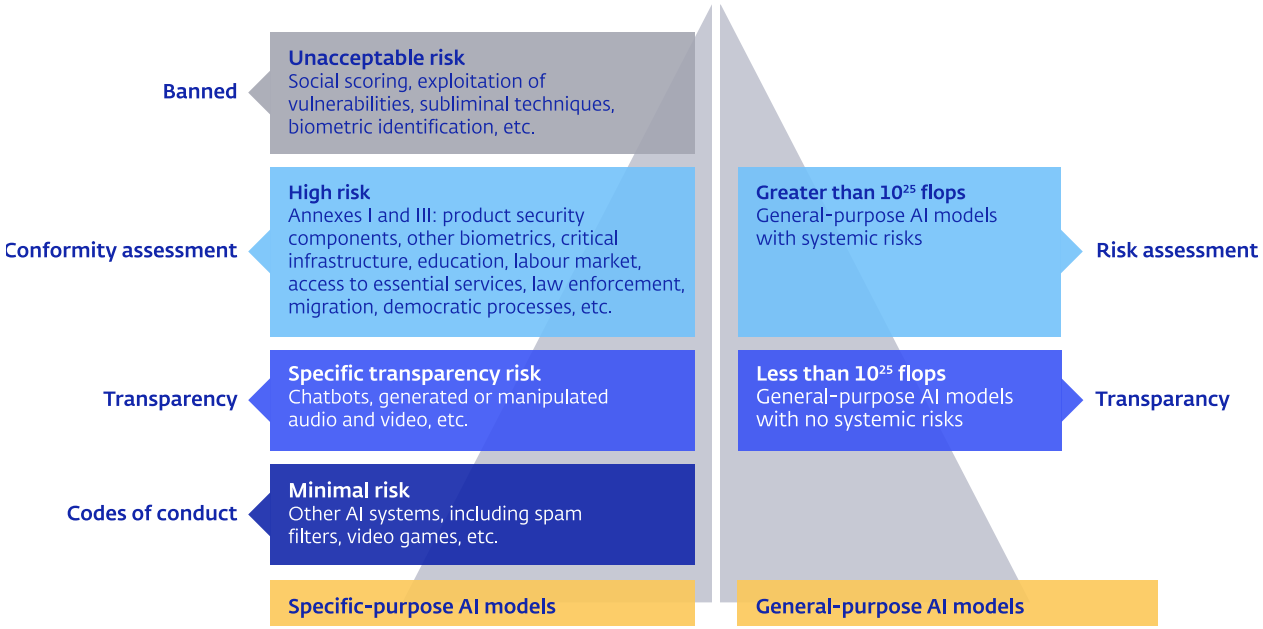
**The AI regulation designates AI systems intended to be used to evaluate the creditworthiness of natural persons and for risk assessment and pricing for life and health insurance as 'high risk'.** These applications must comply with requirements regarding inter alia risk management, data quality, technical documentation, human supervision, robustness and transparency for users. Deployers also need to assess the effects of the AI system on the fundamental rights of individuals or groups. Transparency requirements also apply to AI systems that are not classified as high risk but that interact with people (e.g. chatbots) or that generate content (generative AI), and these will also affect the financial sector.

**The European Commission and the member states will encourage codes of conduct to be developed for applications that are not high risk, so that they comply with the requirements for high-risk systems as far as possible.** These codes of conduct can be developed by the institutions themselves or trade associations. The AFM and DNB fully support this, in line with previous statements.[39] The regulation refers to the European 'Ethics Guidelines for Trustworthy AI',[40] on which the Dutch Association of Insurers has based its Ethics Framework. Although frameworks are often similar in important respects, it may be useful for the ESAs to play a coordinating role here in order to avoid fragmentation. The EIOPA has already published a report on AI governance principles, again based on the European Ethics Guidelines.[41]

**The AI regulation complements the existing sector-based legislation, and it should be applied in line with the financial sector legislation as far as possible.**[42] The requirements are based on existing product regulations, which relate mainly to tangible products under the 'New Legislative Framework'. The requirements apply to AI systems as products and place obligations upon providers and deployers of those systems. In short, high-risk AI systems may only be marketed if a conformity assessment has been carried out in advance, demonstrating compliance with those requirements. This is the responsibility of the system provider. There are also obligations upon an deployer[43] of high-risk AI systems, inter alia regarding the use of the system in line with the requirements, and the competences of persons

39 For details see EBA (2021), Discussion paper on machine learning for IRB models
40 High-Level Expert Group on AI (2019), Ethics Guidelines for Trustworthy AI.
41 EIOPA (2021), AI Governance principles
42 As indicated in Recital 158 of the AI regulation.
43 A user ('deployer') of an AI system for professional purposes.

## Figure 6 Risk-based requirements in the AI regulation




supervising the system, insofar as they have control over it. Financial institutions are deemed to be already partly compliant with this if they observe the existing financial supervision rules on internal governance. Deployers must also assess the impact on fundamental rights as mentioned. At European level, the sector-based regulatory frameworks need to link up with the AI regulation. This involves a coordinating role for the European supervisory authorities, which can provide further guidance where necessary.

**Under the financial sector legislation, a financial institution remains accountable for outsourced activities, and that does not change.** The AI regulation applies a layered approach, in which a different subset of obligations applies, depending on the role an operator has in the value chain (e.g. provider or deployer). Any uncertainties as to precisely who is responsible will need to be removed with regard to the concurrence of financial sector legislation and AI regulation. Financial institutions that themselves provide or deploy high-risk systems will need to comply with the requirements of the AI regulation in addition to the financial regulations. A financial institution that outsources activities is subject to the rules on outsourcing, and a provider to whom they are outsourced to the rules under the AI regulation. This provides better control over AI system providers for the financial sector: under the existing financial regulation they only need to comply with regulatory requirements indirectly, based on the outsourcing relationship. Particular attention needs to be paid to risk management if a small number of AI system providers come to dominate the market.

**In addition to specific-purpose AI systems, the AI regulation also regulates general-purpose AI models.** General-purpose AI models, including large generative AI models, can be used for all sorts of different purposes. Specific requirements apply to these models, inter alia regarding the provision of information enabling providers using these models to develop their own applications and to ensure that those applications are secure and in line with the requirements under the regulation. Providers must also ensure that they respect copyrights. Additional risk management and cyber security requirements apply to models that present system risks. This is gauged by the processing power with which they are trained, with a threshold of $10^{25}$ 'flops' ('floating point operations per second', a measure of computing power). Only the most advanced AI models exceed this.

**Alongside managing risks and protecting fundamental rights, the AI regulation is designed to foster innovation.** It provides a number of measures to encourage innovation, including the setting up of regulatory sandboxes. Here innovatory organisations can obtain advice from regulators on how to interpret the AI regulation in preparation for market launch. There will also be ways of testing applications under real-world conditions outside the sandboxes. In general, the rules under the regulation are risk-based, which means that a strict framework applies to high risk and fewer rules apply to limited and low-risk categories. Applications that are high risk can be provided as long as they have adequate safeguards. Rules and supervision must also foster trust: this is by no means superfluous, as a 2023 survey of Dutch households showed (see Box 2). Only a quarter of respondents were positive about the use of AI by financial institutions. Almost two-thirds of respondents were more positive if they knew that the use of AI was being supervised.

**Coordination mechanisms will be set up for supervision.** Member states will designate national authorities to carry out supervision. In the case of high-risk AI systems provided or used by financial institutions for financial services, the regulation assigns the supervisory role to the existing financial regulators; these are the AFM and DNB in the case of the Netherlands (member states are able to derogate from this). An AI Board will be set up in which the national regulators will work together. The European Commission will also set up a European AI Office to perform a coordinating role and supervise general-purpose AI models. The European authorities will be able to call upon the assistance of an Advisory Forum of stakeholders' organisations and a Scientific Panel of independent experts. AI applications can cross national boundaries (e.g. in algorithmic trading), hence global coordination is required. Given its prudential supervisory role, the ECB is responsible for the AI applications of banks of systemic importance, but it does not assume a supervisory role under the AI regulation.[44] Coordination between the ECB and the national financial regulators will be required to supervise these banks' use of AI.

**Broad coordination of AI regulation is also desirable at national level.** Uses of AI can be found throughout the public and private sectors. Also outside the financial sector, existing regulatory frameworks are already being applied, e.g. to AI applications in health care. The risks of AI systems are best assessed in their specific contexts: this requires sector-based expertise, which limits the extent to which the details of standards can be harmonised. Collaboration between the regulators is desirable for the sake of effective and efficient supervision, while respecting those boundaries. The Dutch Data Protection Authority (AP), with its Department for the Coordination of Algorithmic Oversight in the role of supervisor of algorithms, and the Dutch Authority for Digital Infrastructure (RDI) play a coordinating role here. With regard to the increasing cyber risks, close collaboration with the National Cyber Security Centre (NCSC) will be required.

44 For the background to this see the ECB Opinion: Publications Office (europa.eu).

## Box 2 Opinions differ on AI use in the financial sector

In DNB asked a panel of Dutch households about their impressions of AI in 2023. The survey was presented to 3,054 panel members, and 2,256 respondents completed it in full (response rate 73.9%).

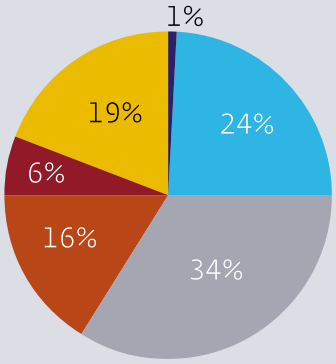
The survey findings show that opinions on AI differ substantially. One-third of respondents were generally positive about AI, 15% had a negative impression. Half of the respondents did not really have an opinion on AI yet or were neutral, perhaps partly because they do not yet have a clear idea of precisely what AI involves: only four in ten said they have a good or reasonably good idea of what AI is.

A small majority of respondents had no opinion yet or were neutral about the use of AI by financial institutions. A quarter were positive. A striking point is that the group of respondents who were negative about financial institutions' use of AI (22%) was larger than the group who were negative about AI in general (15%). Only a few respondents had a good idea of the extent to which their bank (6%), insurance company (5%) or pension fund (4%) was using AI.

As to which were greater, the benefits or risks of financial institutions' AI use, opinions differed: 28% thought that the benefits were greater, 29% that the risks were greater. The remainder were neutral or did not know. The majority (62%) of people were more positive about financial institutions' use of AI if it was supervised.

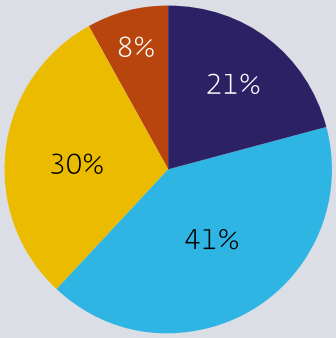
### Figure 7 What is your opinion of financial institutions' use of A

What is your opinion of financial institutions' use of AI?




Very positive
Positive
Neutral
Negative
Very negative
Don't know/no opinion

What is your opinion if your bank, insurance company or pension fund uses AI and you know that its use of AI is being supervised?




Much more positive
Slightly more positive
The same

Note: the survey was held between 13 and 28 March 2023.

## 2.4 AI in the financial sector and the protection of fundamental rights

**Special attention needs to be paid to the protection of fundamental rights when AI is used.** There is a risk of citizens' fundamental rights being infringed when AI is used, as mentioned in §1.4. In the case of financial institutions' use of AI, the important fundamental rights aspects are the prohibition of discrimination and the protection of personal data, which require collaboration between the financial and other regulators. The use of AI in the EU is also subject to the overall EU fundamental rights framework, which comprises the Charter of Fundamental Rights of the European Union (hereinafter referred to as 'the Charter') and the European Convention for the Protection of Human Rights (ECHR). Secondary EU legislation such as the GDPR and EU non-discrimination legislation sets out in detail how particular fundamental rights should be protected. The AI regulation fits within this context, as it lays down detailed rules for the protection of fundamental rights and Union values. At national level the fundamental rights are set out in the Constitution, and the prohibition of discrimination in Article 1 of the Constitution is set out in detail in the Equal Treatment Act. For the purpose of the implementation of the AI regulation it is important for the Dutch legislature to clarify how the regulation's provisions relating to the protection of EU fundamental rights and values relate to the existing roles of regulators (including the AFM and DNB).

**Institutions need to ensure that the AI systems they use do not discriminate.** Models used to estimate risks of transactions – whether e.g. credit risks, insurance risks or money laundering risks – are likely to take customer characteristics into account as well as features of the transaction itself. Distinguishing between customers in this way is permitted, provided it is not on the grounds of religion, beliefs, political affiliation, race, gender, nationality, sexual orientation or marital status. This applies to both direct and indirect distinctions. An indirect distinction is an apparently neutral arrangement that disadvantages a particular group: for instance, requiring an applicant for a bank account to have a home address in the Netherlands is regarded as an indirect distinction based on nationality.[45] Indirect distinction is permitted if it is objectively justified, appropriate and proportional. In the case cited, for instance, applying a residence requirement could be a legitimate way for a bank to take anti-money laundering measures, but it is not necessary to couple failure to comply with the residence requirement with automatic rejection, as the applicant could demonstrate in some other way that the money-laundering risks are mitigated. Models that use self-learning algorithms could unintentionally make an indirect distinction in a way that is prohibited. Institutions need to prevent this, or if it does occur, take steps to remove discriminatory effects from the algorithm.[46]

45 Judgment of the College for Human Rights 2023-22, 22/2/2023 (in Dutch).
46 Judgment of the College for Human Rights 2023-82 ('Breeze Social'), 1/8/2023 (in Dutch).

**Data protection is an area of concern, as developing, training and using AI involves processing large quantities of data.** The more information and data an AI model can use as input, the better it will be able to predict e.g. a customer's creditworthiness. The GDPR lays down requirements for the processing of personal data and gives customers the right not to be subjected to a decision based solely on automated processing. These rules also apply to the training and use of AI models by financial institutions. The processing of sensitive personal data such as health, political preference and religion is prohibited, unless an exception is applicable, e.g. explicit consent or an exception in the law. For a financial institution this means that all the available information cannot automatically be used to develop or use an AI model.

**Dilemmas can arise between the interests of data protection and the prevention of discrimination.** Even if sensitive personal data is expressly not used to develop an algorithm, that algorithm could nevertheless obtain certain sensitive personal data. For instance, based on a customer's payment behaviour, an AI model could develop proxies for sensitive personal data to which the algorithm did not originally have access. For example, an AI model could use payments for medical treatment as a proxy to gauge a customer's health. Another example is using a customer's home address as a proxy to estimate their income. Using an AI model could thus – inadvertently, and possibly unnoticed – result in a type of discrimination. In order to ensure that an AI model 'acts fairly', checks need to be carried out – using specific models if necessary – to ensure that it does not structurally discriminate against particular groups. It is important for financial institutions – taking the guidance provided in various statutory frameworks into account – to develop AI models in such a way that effective checks on discrimination can be carried out, so that appropriate action can be taken. The AI regulation provides for the use of sensitive personal data for this purpose in exceptional cases and subject to strict conditions. This enables a model to be validated and modified if necessary; if this is not possible, it can be decided not to use it.

## Method

**The content of this report was developed in collaboration between various AFM and DNB experts.** Multidisciplinary teams conducted interviews with institutions under supervision, trade associations, European regulators, technology suppliers and academics. Various studies and analyses (both internal and external) were also consulted.

De Nederlandsche Bank N.V.
PO Box 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl

Follow us on:

Autoriteit Financiële Markten
PO Box 11723, 1001 GS Amsterdam
020 797 20 00
afm.nl

Follow us on: