

# DP5/22 – Artificial Intelligence and Machine Learning

Discussion Paper 5/22

## Content

### Privacy Statement

---

### Foreword

---

### Executive summary

Benefits and risks related to the use of AI in financial services

How existing legal requirements and guidance apply to the use of AI

Discussion questions for stakeholder input

---

### 1. Introduction

Background

---

### Box 1: UK government – establishing a pro-innovation approach to regulating AI

---

### 2. Supervisory authorities' objectives and remits

Why the supervisory authorities have an interest in AI

What is AI?

How AI is used in financial services

Questions: Supervisory authorities' objectives and remits

---

### 3. Potential benefits and risks

Consumer protection – FCA

Competition – FCA

Safety and soundness – PRA and FCA

Insurance policyholder protection – PRA and FCA

Financial stability and market integrity – the Bank and FCA

Questions: Benefits, risks, and harms of AI

---

### 4. Regulation

Introduction

Consumer protection – FCA

Competition – FCA and PRA

Safety and soundness: Data – PRA and FCA

Safety and soundness: Model risk management – PRA

Safety and soundness: Governance – PRA and FCA

Safety and soundness: Operational resilience, outsourcing, and third-party risk management – PRA and FCA

Questions: Regulation

---

## **5. Questions**

Supervisory authorities' objectives and remits

Benefits, risks, and harms of AI

Regulation

---

# Privacy Statement

---

By responding to this discussion paper, you provide personal data to the Bank of England (the Bank) (including the Prudential Regulation Authority (PRA)) and to the Financial Conduct Authority (FCA) ('we' or 'us'). This may include your name, contact details (including, if provided, details of the organisation you work for), and opinions or details offered in the response itself.

We will need to assess your response to inform our work as a regulator and central bank, both in the public interest and in the exercise of our official authority. We may also use your details if we need to contact you to clarify any aspects of your response.

The discussion paper will explain if responses will be shared with other organisations. If this is the case, the other organisation will also review the responses and may also contact you to clarify aspects of your response.

We will retain all responses for the period that is relevant to supporting ongoing regulatory policy developments and reviews. The Bank and PRA will redact all personal information from responses within five years of receipt. The FCA will retain all personal information in line with their retention schedule linked [here](#). To find out more about how we manage your personal data, your rights, or to get in touch, please visit the [Bank's privacy page](#) and the [FCA's privacy page](#).

Information provided in response to this paper, including personal information, may be subject to publication or disclosure to other parties in accordance with access to information regimes including under the Freedom of Information Act 2000 or data protection legislation, or as otherwise required by law or in discharge of the PRA's, the Bank's or the FCA's functions.

We may choose to publish certain responses to this discussion paper to help inform the public debate on artificial intelligence. In doing so, we will not publish your name or contact details or any information which may personally identify you.

Please indicate if you regard all, or some of, the information you provide as confidential. If we receive a request for disclosure of this information, we will take your indication(s) into account, but cannot give an assurance that confidentiality can be maintained in all circumstances. An automatic confidentiality disclaimer generated by your IT system on emails will not, of itself, be regarded as binding on us.

Responses are requested by Friday 10 February 2023.

**We prefer responses to be sent via email to:** [DP5\\_22@bankofengland.co.uk](mailto:DP5_22@bankofengland.co.uk).

**Alternatively, please address any comments or enquiries to:**

Reporting, Disclosure, Data Strategy and AI Team

Prudential Regulation Authority

Threadneedle Street

London

EC2R 8AH

## Foreword

---

The use of artificial intelligence (AI) and machine learning (ML) in financial services may enable firms to offer better products and services to consumers, improve operational efficiency, increase revenue, and drive innovation. All of which may lead to better outcomes for consumers, firms, financial markets, and the wider economy.

As our recent survey indicates, AI adoption within financial services is likely to continue to increase due to increased availability of data, improvements in computational power, and wider availability of AI skills and resources. Similarly, the Bank of England (the Bank), the Prudential Regulation Authority (PRA), and the Financial Conduct Authority (FCA) (collectively ‘the supervisory authorities’) also endeavour to leverage AI and benefit from this technology to help meet our respective statutory objectives and other functions.

Although the use of AI may bring a range of benefits, it can also pose novel challenges for firms and regulators as well as amplify existing risks to consumers, the safety and soundness of firms, market integrity, and financial stability. One of the most significant questions is whether AI can be managed through clarifications of the existing regulatory framework, or whether a new approach is needed. How to regulate AI to ensure it delivers in the best interests of consumers, firms, and markets is the subject of a wide-ranging debate, both here in the UK and in other jurisdictions around the world.

This Discussion Paper (DP) sits within the context of this wider debate and focuses on the regulation of AI in UK financial services. The Bank, PRA, and the FCA seek to encourage a broad-based and structured discussion with stakeholders on the challenges associated with the use and regulation of AI. We are keen to explore how best to address these issues in a way that is aligned with our statutory objectives, provides clarity, is actionable, and makes a practical difference for consumers, firms, and markets.

Beginning with our current regulatory framework, we have considered how key existing sectoral legal requirements and guidance in UK financial services apply to AI. This evaluation will allow us to consider which ones are most relevant, explore whether they are sufficient,

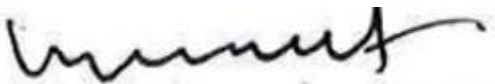
and identify gaps. The DP considers how such legal requirements and guidance apply to the use of AI in UK financial services to support consumer protection, competition, the safety and soundness of individual firms, market integrity, and financial stability. How can policy mitigate AI risks while facilitating beneficial innovation? Is there a role for technical and, indeed, for global standards? If so, what?

Given the extent of overlaps within the existing sectoral rules, policies and principles in UK financial services that apply to AI, the supervisory authorities' approach is largely limited to clarifying how the existing regulatory framework applies to AI and addressing any identified gaps in the regulatory framework. In particular, the supervisory authorities are interested in the additional challenges and risks that AI brings to firms' decision-making and governance processes, and how those may be addressed through the Senior Managers and Certification Regime (SM&CR) and other existing regulatory tools.

Given the wide-ranging implications of AI, we are keen to hear from a broad range of stakeholders. This includes firms regulated by the Bank, PRA and/or FCA, as well as non-regulated financial services firms, professional services firms (such as accounting and auditing firms), law firms, third parties (such as technology companies), trade associations and industry bodies, standard setting organisations, academics, and civil society organisations.

We note the importance of building, maintaining, and reinforcing the trust of all stakeholders, including consumers in AI. Engagement between the public and private sectors will facilitate the creation of a regulatory framework that enables innovation and mitigates potential risks.

We hope that this DP contributes to this process and look forward to hearing from you.



Victoria Saporta

Executive Director,

Prudential Policy Directorate,

Bank of England



Sheldon Mills

Executive Director,

Consumers and Competition,

Financial Conduct Authority

A handwritten signature in black ink that reads "Jessica Rusu". The signature is written in a cursive, flowing style with a large initial 'J' and 'R'.

Jessica Rusu

Chief Data, Information and Intelligence Officer (CDIIO),

Financial Conduct Authority

## Executive summary

---

Artificial intelligence (AI) and machine learning (ML) are rapidly developing technologies that have the potential to transform financial services. The promise of this technology is to make financial services and markets more efficient, accessible, and tailored to consumer needs. This may bring important benefits to consumers, financial services firms, financial markets, and the wider economy.

However, AI can pose novel challenges, as well as create new regulatory risks, or amplify existing ones. The Bank of England (the Bank), the Prudential Regulation Authority (PRA) and the Financial Conduct Authority (FCA) therefore have a close interest in the safe and responsible adoption of AI in UK financial services, including considering how policy and regulation can best support this.

The supervisory authorities are publishing this DP to further our understanding and to deepen dialogue on how AI may affect our respective objectives. This is part of the supervisory authorities' wider programme of work related to AI, including the [AI Public Private Forum](#), the [final report](#) of which was published in February 2022. This DP should also be considered within the context of the evolving wider national and international policy debate on AI, including the [UK government's policy paper 'Establishing a pro-innovation approach to regulating AI'](#) joint working between UK regulators through the [Digital Regulation Cooperation Forum \(DRCF\)](#) and international developments from other regulators and authorities, such as the [proposed AI regulation for the EU](#)

## **Benefits and risks related to the use of AI in financial services**

AI offers potential benefits for consumers, businesses, and markets. However, AI also has the potential to create new or increased risks and challenges. The benefits, risks, and harms discussed in this DP are neither exhaustive nor applicable to every AI use case.

The primary drivers of AI risk in financial services relate to three key stages of the AI lifecycle: (i) data; (ii) models; and (iii) governance. Interconnected risks at the data level can feed into the model level, and then raise broader challenges at the level of the firm and its overall governance of AI systems. Depending on how AI is used in financial services, issues at each of the three stages (data, models, and governance) can result in a range of outcomes and risks that are relevant to the supervisory authorities' remits.

### **Consumers**

AI may benefit consumers in important ways – from improved outcomes through more effective matching to products and services, to an enhanced ability to identify and support consumers with characteristics of vulnerability, as well as increasing financial access. However, if misused, these technologies may potentially lead to harmful targeting of consumers' behavioural biases or characteristics of vulnerability, discriminatory decisions, financial exclusion, and reduced trust.

### **Competition**

There may be substantial benefits to competition from the use of AI in financial services, where these technologies may enable consumers to access, assess, and act on information more effectively. But risks to competition may also arise where AI is used to implement or facilitate further harmful strategic behaviour such as collusion, or creating or exacerbating market features that hinder competition, such as barriers to entry or to leverage a dominant position.

### **Firms**

There are also many potential benefits for financial services firms including enhanced data and analytical insights, increased revenue generation, increased operational efficiency and productivity, enhanced risk management and controls, and better combatting of fraud and money laundering. Equally, the use of AI can translate into a range of prudential risks to the safety and soundness of firms, which may differ depending on how the technology is used by firms.



## **Financial markets**

AI may benefit the broader financial system and markets in general through more responsive pricing and more accurate decision-making, which can, in turn, lead to increased allocative efficiency. However, AI may also lead to risks to system resilience and efficiency. For example, models may become correlated in subtle ways and add to risks of herding, or procyclical behaviour at times of market stress.

## **How existing legal requirements and guidance apply to the use of AI**

In line with their statutory objectives and to support the safe and responsible adoption of AI in UK financial services, the supervisory authorities may need to intervene further to manage and mitigate the potential risks and harms AI may have on consumers, firms, and the stability and integrity of the UK financial system and markets.

It is important that the regulatory environment is proportionate and conducive to facilitating safe and responsible adoption of AI, so as not to act as a barrier to beneficial innovation. A first step towards this ambition is clarifying how existing legal requirements and guidance apply to the use of AI.

In addition to legal requirements and guidance targeted at particular risks (such as risks to consumers or effective competition), the supervisory authorities have identified sets of 'cross-cutting' legal requirements and guidance that encompasses multiple areas of risk.

Cross-cutting legal requirements and guidance are relevant primarily to the three key stages of the AI lifecycle – data, models, and governance. Data-related legal requirements and guidance are targeted at data quality, data privacy, data infrastructure, and data governance. Model-related legal requirements and guidance for managing capital risks may provide safeguards surrounding the model development, validation, and review processes for the firms to which these apply. Governance-related legal requirements and guidance (including, notably, the SM&CR, are focused on proper procedures, clear accountability, and effective risk management across the AI lifecycle at various levels of operations.

AI industry standards or codes of conduct may potentially complement the regulatory system by helping firms build trust amongst users that their systems meet widely accepted industry norms, which may extend beyond the minimum requirements for regulatory compliance.

The supervisory authorities encourage all relevant stakeholders to respond to this DP. This includes financial services firms regulated by one or more of the supervisory authorities (including both dual- and solo-regulated firms), non-regulated financial services firms, professional services firms (such as accounting and auditing firms), law firms, third parties (such as technology companies), trade associations and industry bodies, standard setting organisations, academics, and representatives from civil society.

## Discussion questions for stakeholder input

This DP seeks to explore whether stakeholders consider the existing sectoral legal requirements and guidance to be sufficient to address the risks and harms associated with AI, where there may be gaps in existing legal requirements and guidance, and/or how any additional intervention may support the safe and responsible adoption of AI in UK financial markets.

To help us in this aim, we are inviting responses to the questions listed in Chapter 5. The questions fall under three main categories:

- **Supervisory authorities' objectives and remits:** exploring the best approach to defining and/or scoping the characteristics of AI for the purposes of legal requirements and guidance.
- **Benefits and risks of AI:** identifying the areas of benefits, risks, and harms in relation to which the supervisory authorities should prioritise action.
- **Regulation:** exploring whether the current set of legal requirements and guidance is sufficient to address the risks and harms associated with AI and how additional intervention may support the safe and responsible adoption of AI in UK financial services. This includes understanding which areas of the current regulatory framework: (i) would benefit from further clarification with respect to AI, (ii) could be extended to better encompass AI, and (iii) could act as a regulatory barrier to the safe and responsible adoption of AI in UK financial services.

## 1. Introduction

---

1.1 The Bank and the FCA established the AI Public-Private Forum (AIPPF) to further dialogue on AI innovation and safe adoption within financial services. The AIPPF launched in October 2020 and ran for one year bringing together a diverse group of experts from across financial services, the tech sector, and academia.

1.2 The AIPPF published its final report in February 2022. The report explores the various barriers to adoption, challenges, and risks related to the use of AI in financial services. The report reflects the views of its members as individual experts, rather than the views of the Bank or the FCA.

1.3 The Bank and the FCA are publishing this DP in response to the AIPPF final report, which made it clear that the private sector wants regulators to have a role in supporting the safe adoption of AI in UK financial services, as well as a wider background of domestic and international developments regarding AI regulation.

1.4 The purpose of this DP is to share and obtain feedback on:

- the potential benefits, risks, and harms related to the use of AI in financial services;
- how the current regulatory framework could apply to AI;
- whether additional clarification may be helpful; and
- how policy can best support further safe AI adoption.

## Background

1.5 In 2019, the supervisory authorities conducted a [joint survey](#) to better understand the use of AI and ML in UK financial services. The survey identified that financial services firms were increasingly using AI and needed effective and evolving risk management controls if they were to use it safely and harness the benefits.

1.6 The Bank, in its [response](#) to the **Future of Finance Report**, announced in June 2019 that it would establish, with the FCA and firms, a public-private working group, the [AI Public-Private Forum](#), to further the dialogue on AI innovation and to explore whether principles and guidance could support the safe adoption of these technologies within financial services. The AIPPF launched in October 2020 and ran for one year, with four quarterly meetings and a number of workshops. It brought together a diverse group of experts from across financial services, the tech sector and academia, along with public sector observers from other UK regulators and government.

1.7 The AIPPF published its final report in February 2022. The report explores the various barriers, challenges, and risks related to the use of AI in financial services and potential ways to address them. The report reflects the views of AIPPF members in a personal capacity, rather than their institutions, the Bank, or the FCA.

1.8 The AIPPF has made clear to the supervisory authorities that the private sector wants regulators to have a role in supporting the safe adoption of AI in UK financial services.

1.9 At the same time, the UK government has sought to develop a national position on regulating AI across all sectors, with the publication, in July 2022, of a policy paper [‘Establishing a pro-innovation approach to regulating AI’](#) (July 2022 policy paper). Similarly, other governments and regulators around the world are developing their thinking and approaches to AI. For example, the [European Commission’s proposal for AI regulation](#) the [Principles to promote Fairness, Ethics, Accountability and Transparency in the Use of AI and Data Analytics in Singapore’s Financial Sector](#) (FEAT) and the [Veritas Initiative from the Monetary Authority of Singapore](#) the [Organisation for Economic Co-operation and Development \(OECD\) AI Principles](#) and the [Internet Information Service Algorithmic Recommendation Management Provisions](#) that came in force in March 2022 in the People’s Republic of China.

1.10 The supervisory authorities are publishing this DP in response to the AIPPF final report following their previous work on AI (including through the Digital Regulation Co-Operation Forum), and in the wider domestic and international context of emerging AI regulation, policies, and principles in order to facilitate a public debate on the safe and responsible adoption of AI in UK financial services.

## **Box 1: UK government – establishing a pro-innovation approach to regulating AI**

The UK government published the July 2022 policy paper which set out its emerging thinking on its approach to regulating AI.

The July 2022 policy paper highlighted several key challenges that cut across the regulatory landscape, including a lack of clarity, overlaps, inconsistencies, and gaps in the current approach. Many of these challenges may apply to the regulation of the UK financial services sector. The supervisory authorities hope the DP stimulates debate on how best to overcome these challenges within UK financial services.

The UK government aims to address these challenges by creating a framework for regulating AI that is context-specific, pro-innovation and risk-based, coherent, proportionate, and adaptable.

Moreover, the UK government's proposed approach to regulating AI is underpinned by a set of cross-sectoral principles tailored to the specific characteristics of AI:

- ensure that AI is used safely
- ensure that AI is technically secure and functions as designed
- make sure that AI is appropriately transparent and explainable
- embed considerations of fairness into AI
- define legal persons' responsibility for AI governance
- clarify routes to redress or contestability

The UK government stated that it will be considering the roles, powers, remits and capabilities of regulators, the need for coordination, and how this should be delivered across the range of regulators (statutory and non-statutory) involved in AI regulation as part of its next steps.

See Table A in Appendix 1 for a mapping of the UK government AI regulation principles against the DP chapters and sections.

## Discussion paper structure

1.11 Chapter 2 explains how the use of AI in UK financial services could impact the supervisory authorities' respective objectives. The chapter then explores the potential merits of providing a regulatory definition for AI, and how regulators in other jurisdictions are approaching this. The chapter also provides a brief overview of how AI is used in financial services.

1.12 Chapter 3 sets out some of the benefits and risks related to the use of AI in UK financial services. These are divided into different categories based on each of the supervisory authorities' objectives (eg consumer protection, competition, safety and soundness of firms, etc). More information on some of the potential risks and benefits associated with the use of AI in UK financial markets can be found in a [report by the Alan Turing Institute](#) which was commissioned by the FCA.

1.13 Chapter 4 provides an overview of how current key legal requirements and guidance apply to the use of AI in UK financial services. This includes how domestic regulation links to international regulation and how financial services regulations sit alongside a body of cross-sectoral legislation and regulation. The supervisory authorities are keen to understand stakeholders' views as to whether the current regulatory framework is sufficient to address the potential risks and harms associated with AI and/or how any additional intervention may support the safe and responsible adoption of AI in UK financial services.

1.14 Annex 1 sets out how the DP sits within the wider context of domestic developments that are relevant to the use and regulation of AI in UK financial services. This includes the [UK Government's National AI Strategy](#) and the July 2022 policy paper.

1.15 Annex 2 highlights the different international regulatory responses to AI in financial services and examines emerging approaches. The supervisory authorities are keen to explore if any of the international approaches or elements of any approaches may be helpful to support the safe and responsible adoption of AI in UK financial services.

1.16 Annex 3 sets out an overview of certain key existing legal requirements and guidance relevant to data and AI.

1.17 Annex 4 summarises existing PRA Supervisory Statements (SS) and certain other standards relevant to various aspects of model risk management for PRA-authorized firms.

1.18 Annex 5 provides a list of selected relevant publications on the use of AI in financial services.

## Responses and next steps

1.19 This DP closes on Friday 10 February 2023. The supervisory authorities invite feedback on the topics discussed in this DP. Please address any comments or enquiries to [DP5\\_22@bankofengland.co.uk](mailto:DP5_22@bankofengland.co.uk).

1.20 The responses to this DP will help inform both the supervisory authorities' thinking and any potential future policy proposals, which will sit within the wider domestic and international context of emerging AI regulation. As with other DPs, the supervisory authorities may choose to publish the anonymised and aggregated responses to this DP to help inform the public debate on AI.

1.21 Therefore, the supervisory authorities would encourage all relevant stakeholders to respond to this DP and engage in the discussion. This includes financial services firms regulated by one or more of the supervisory authorities (including both dual- and solo-regulated firms), non-regulated financial services firms, professional services firms (such as accounting and auditing firms), law firms, third parties (such as technology companies), trade associations and industry bodies, standard setting organisations, academics, and civil society organisations.

## 2. Supervisory authorities' objectives and remits

---

2.1 The supervisory authorities consider the use of AI in financial services to be relevant to their respective objectives, notably the below.

2.2 **The Bank's mission** is to maintain monetary and financial stability in the UK and, subject to maintaining price stability, the Bank is to support the UK government's economic policy (including its objectives for growth and employment). The Bank is also responsible for **supervising certain Financial Market Infrastructure firms (FMIs)** in the UK, including central securities depositories (CSDs), central counterparties (CCPs), recognised payment systems operators (RPSOs), and 'specified' service providers (SPs) to RPSOs.

2.3 The PRA has **two primary objectives**: (i) a general objective to promote the safety and soundness of PRA-authorised firms; and (ii) an objective specific to insurance firms, to contribute to securing an appropriate degree of protection for those who are or who may become policyholders. The PRA also has a **secondary objective** to facilitate effective competition in the markets for services provided by PRA-authorised firms in carrying on regulated activities.

2.4 **The FCA's strategic objective** is to ensure that the relevant markets (including financial markets) function well, and its operational objectives are to: (i) secure an appropriate degree of protection for consumers, (ii) protect and enhance the integrity of the

UK financial system and (iii) promote effective competition in the interests of consumers in the markets for regulated financial services. The FCA must, so far as is compatible with its consumer protection and market integrity objectives, also exercise its functions in a way that promotes effective competition in the interests of consumers.

2.5 Effective regulation relies on consent, trust, and confidence from the public, including consumers and regulated firms, and ensures that powers are used consistently, transparently, and proportionately. The [FCA Mission](#) sets out a framework for the way the FCA takes these decisions and serves the public interest.

## **Why the supervisory authorities have an interest in AI**

2.6 AI may bring important benefits to consumers, financial services firms, financial markets, and the wider economy. The promise of this technology is to make financial services and markets more cost-effective, efficient, accessible, and tailored to consumer needs. However, AI can pose novel challenges, as well as create new risks or amplify existing ones.

Therefore, the supervisory authorities have a close interest in the safe and responsible adoption of AI in UK financial services. In line with their statutory objectives and to support the safe and responsible adoption of AI technologies in UK financial services, the supervisory authorities may need to intervene further to manage and mitigate the potential risks and harms related to AI applications.

2.7 While the supervisory authorities generally take a technology neutral approach to regulation (see Box 2 for more details), they are aware that risks may relate to the use of specific technologies.

2.8 The supervisory authorities recognise the importance to consumers, financial services firms and markets of a regulatory environment conducive to facilitating safe and beneficial innovation and competition. The PRA and the FCA have competition-related statutory objectives and are required to have regard to the principle of proportionality when exercising their general functions (including when making rules). For instance, the supervisory authorities would wish to avoid introducing unnecessarily complex or costly rules that can act as a barrier to entry, especially for non-systemic firms, or unintentionally favour larger firms with more resource to manage compliance with complex rulebooks. A proportionate approach is critical to supporting the safe and responsible adoption of AI and other technologies across UK financial services.

2.9 It is also important to acknowledge that financial services firms are themselves subject to compliance obligations to other regulators when using AI. For example:

- where financial services firms use AI to process personal data, firms may have regulatory obligations under the [UK General Data Protection Regulation \(UK GDPR\)](#) and the [Data Protection Act 2018](#). The Information Commissioner's Office (ICO) has responsibility for enforcing compliance with data protection requirements; and

- financial services firms also need to ensure compliance with the [Equality Act 2010](#) (Equality Act), in respect of which the Equality and Human Rights Commission (EHRC) has enforcement powers. Where firms use AI systems in their decision-making processes, they will need to ensure that it does not result in unlawful discrimination based on protected characteristics.

## **Box 2: What do we mean by technology neutral?**

The supervisory authorities recognise that the use of new technologies (such as AI) in financial services may benefit consumers through new and better products and services, enhance the safety and soundness of firms through improved risk management informed by advanced analytics and enhanced decision making, and deliver a more efficient and effective financial system to the benefit of the wider economy.

Generally, the supervisory authorities adopt a technology-neutral approach in their supervisory and regulatory approaches. As such, the core principles, rules, and regulations do not usually mandate or prohibit specific technologies.

However, the supervisory authorities monitor and mitigate technology risks as they can equally have adverse implications for their objectives (including consumer protection, competition, the safety and soundness of firms, market integrity, and financial stability). The use of a certain technology may have an impact on the risks associated with firms' activities.

The supervisory authorities may therefore issue technology-specific rules and guidance where appropriate. Certain technologies may also raise novel challenges for firms and regulators, which may mean it is difficult for firms to understand how existing rules apply to that technology. In those cases, the supervisory authorities may issue guidance or use other policy tools to clarify how the existing rules and relevant regulatory expectations apply to those technologies.

For example, [FG16/5 Guidance for firms outsourcing to the 'cloud' and other third-party IT services](#) and paragraph 2.4 in [PRA SS2/21 'Outsourcing and third party risk management'](#). Firms could, for example, also consider outcomes and risk management as covered in the [FCA's 'Implementing Technology Change' paper](#)

This DP reflects the supervisory authorities desire to understand how they may best support the safe and responsible adoption of AI in UK financial services in line with their statutory objectives.



## What is AI?

2.10 While there is no consensus on a single definition, it is generally accepted that AI is the simulation of human intelligence by machines, including the use of computer systems, which have the ability to perform tasks that demonstrate learning, decision-making, problem solving, and other tasks which previously required human intelligence. Machine learning is a sub-branch of AI. AI, a branch of computer science, is complex and evolving in terms of its precise definition. It is broadly seen as part of a spectrum of computational and mathematical methodologies that include innovative data analytics and data modelling techniques.

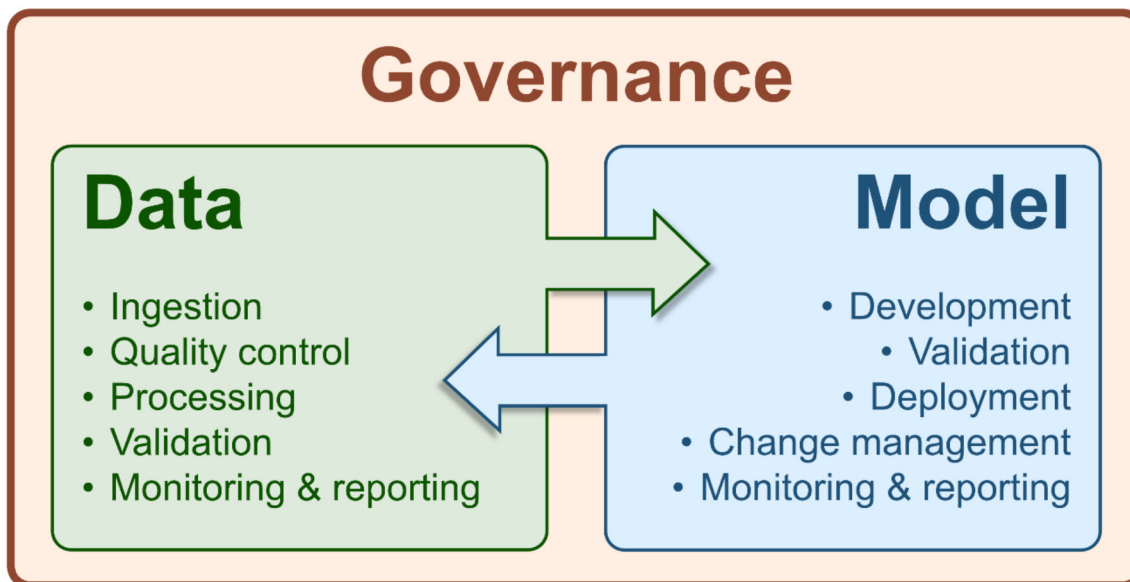
2.11 The Bank and FCA previously defined AI in non-legal terms as ‘the theory and development of computer systems able to perform tasks which previously required human intelligence’. The term AI system refers to the set of integrated computational elements and microservices that input, output, process, and store data and information. Therefore, one AI system may include multiple AI algorithms, models, and datasets.

2.12 It is important to distinguish between the terms ‘algorithm’ and ‘model’ as these have specific meanings within financial services regulation. For example, the Competition and Markets Authority (CMA) defines ‘algorithmic systems’ as shorthand for an automated system, at the intersection of algorithms, data, models, processes, objectives, and people. Although the supervisory authorities do not define algorithms, they do define ‘algorithmic trading’ in the context of the Markets in Financial Instruments Directive (MiFID) implementation and members of the Digital Regulation Cooperation Forum (DRCF) also defines ‘algorithmic processing’. The PRA defines ‘model’ in its recent CP6/22 ‘Model risk management principles for banks’. With all of these definitions, it is important to note that while both algorithms and models may be components of an AI system, they may not in themselves necessarily be deemed to be AI.

2.13 Many of the issues related to AI are not new and apply to other data analytics, modelling techniques, and technologies. For example, many traditional financial models (such as generalized linear models (GLM) used to classify risk in insurance and internal ratings based (IRB) approaches to modelling regulatory capital requirements in banking) that do not use AI may be as complex and difficult to understand or explain.

2.14 However, there are three key areas in the AI lifecycle (see Figure 1 below) that can introduce novel challenges and increase existing risks.

Figure 1: Stages of AI lifecycle



Source: AIPPF

2.15 **Data** – AI can analyse traditional data sources as well as unstructured and alternative data from new sources (such as image and text data). The increase in volume of data and data formats within the context of AI places added emphasis on data quality. Also, AI may pick up bias within datasets and may not perform as intended when exposed to issues excluded from the training/testing data, so new data quality metrics like representativeness and completeness may be needed.

2.16 **Model** – Whereas traditional financial models are usually rules-based with explicit fixed parameterisation, AI models are able to learn the rules and alter model parameterisation iteratively. The use of AI models also represents a step change for three other reasons: firstly, the speed and frequency at which the models update (with some AI models able to learn continuously); secondly, the scale in terms of the volume of data needed to train the models and the number of features that are used as inputs; and thirdly, the complexity of certain techniques, such as convolutional neural networks,[1] which can make them more opaque (the so-called '**black box problem**' ).

2.17 **Governance** – AI may also pose some novel challenges for governance, especially where the technology is used to facilitate autonomous decision-making and may limit or even potentially eliminate human judgement and oversight from decisions. Some of the data and model issues can also have implications for governance. For example, a lack of explainability or transparency in some AI models may mean extra care or actions are needed to ensure full accountability and sufficient oversight.

2.18 Given the kind of issues outlined above, regulators and authorities have generally found it useful to distinguish between AI and non-AI by either:

- providing a more precise legal definition of what AI is (for example, the definition of ‘artificial intelligence system’ in the proposed [European AI regulation](#) and proposed [Canadian AI and Data Act](#) and, thereby, what AI is not; or
- viewing AI as part of a wider spectrum of analytical techniques with a range of elements and characteristics. Such approaches may include a classification scheme to encompass different methodologies or a mapping of the characteristics of AI (such as, the [German financial services regulators’ consultation paper on machine learning in risk models](#) the [OECD Classification System](#) and the [UK government’s proposed framework for regulating AI](#) ).

2.19 Each approach aims to provide clarity of what constitutes AI within the context of a specific regulatory regime and, therefore, what restrictions and requirements may apply to the use of the technology. This, depending on the particular requirements applied, may help: (i) consumers to understand where and when AI is used in products and services they use; (ii) firms to use the technology responsibly and manage the relevant trade-offs; and (iii) regulators to assess whether firms are using the technology in a safe and responsible manner and establishing appropriate controls and procedures to mitigate against potential risks as well as adhering to any regulatory requirements.

2.20 For the supervisory authorities, there may be a number of benefits to providing a more precise definition of AI in their respective rulebooks. The benefits may include: (i) creating a common language for firms and regulators, which may ease uncertainty; (ii) assisting in a uniform and harmonised response from regulators towards AI; and (iii) providing a basis for identifying whether or not specific use cases might be captured under particular rules and principles.

2.21 At the same time, there are various challenges to the supervisory authorities providing a more precise definition as the basis for clarification or additional or alternative regulatory requirements. These include: (i) the difficulty to create a robust definition that would remain relevant in what is undoubtedly a rapidly developing field; (ii) challenges of a definition that is too broad, as this could cause regulation to lack specificity, precision, and proportionality; (iii) potential for certain use cases, products, and services to fall outside of scope when they

should not; and (iv) creating the ability and incentives for firms to misclassify AI to reduce regulatory oversight. The alternative approaches, as described above, may address some of these challenges and could potentially be more suitable for the regulation of AI in UK financial services.

## How AI is used in financial services

2.22 Financial services firms are increasingly using AI across a range of business areas and the Covid-19 pandemic has accelerated the pace of adoption (see Annex 5 for a list of relevant publications).

2.23 In this rapidly evolving space, the following high-level trends in financial services may have implications for consumers, firms, markets, and the supervisory authorities:

- Greater AI adoption is occurring both within established financial services firms and also newer fintech and insurtech companies. This is being powered by improved software and hardware, including greater use of cloud computing which is often provided via third parties, as well as via partnerships and investment with smaller AI-specific vendors.
- Firms are using AI for more material business areas and use cases. From anti-money laundering (AML) functions and credit and regulatory capital modelling in banking; claims management and product pricing to capital reserve modelling in insurance; and order routing and execution to generating trading signals in investment management.
- firms are also using more complex AI techniques. This is because complexity often corresponds to improved performance and use of ever greater volumes of data. As firms begin to use more complex methods and data sets, and apply AI to more complex use cases, they will likely become more comfortable with the technology and adoption of even more complex techniques will increase further.

2.24 For more information and detail, please see the supervisory authorities' recent report on [Machine Learning in UK Financial Services 2022](#) (2022 ML Survey).

### Box 3: The supervisory authorities' use of AI

The supervisory authorities utilise AI, where appropriate, to support and enhance their capabilities. This can include RegTech, which refers to the use of technology by firms to support them to fulfil their regulatory obligations; and SupTech, which refers to the use of technology by regulators to facilitate their supervisory duties. There is also the potential to use AI to support central bank functions, such as economic analysis.

The Bank uses AI for predictive analytics, the study of non-linear interactions between variables, and analysis of larger and richer datasets, which can potentially help [forecast GDP growth](#), [bank distress](#) and [financial crises prediction](#). The Bank is

also exploring how [AI-enabled text analysis of newspapers](#) can help improve economic forecasting and how AI can create 'faster indicators', which may enable real-time economic analysis.

The PRA successfully introduced a cognitive search tool with AI capabilities that helps supervisors gain more insights from firm management information by extracting key patterns from unstructured and complex datasets. The PRA is also developing other AI tools for its staff to assist in their work.

The FCA has been exploring how AI can provide additional levels of insight across the organisation. Examples include using Natural Language Processing to generate insights from unstructured text documentation; predictive analytics to forecast where risks may lie across our supervisory waterfront; and simulation techniques to generate realistic synthetic data sets.

The potential for AI innovation to assist the supervisory authorities and regulatory activity depends significantly on access to data and technology, so may have implications for the data that firms are required to share.

## **Questions: Supervisory authorities' objectives and remits**

Q1: Would a sectoral regulatory definition of AI, included in the supervisory authorities' rulebooks to underpin specific rules and regulatory requirements, help UK financial services firms adopt AI safely and responsibly? If so, what should the definition be?

Q2: Are there equally effective approaches to support the safe and responsible adoption of AI that do not rely on a definition? If so, what are they and which approaches are most suitable for UK financial services?

## **3. Potential benefits and risks**

---

3.1 This chapter briefly summarises the potential benefits and risks of the use of AI in financial services. It also describes how the drivers of AI benefits and risks in financial services can occur at different levels within AI systems (data, models, and governance) and how these drivers of risk can result in a range of outcomes depending on how the AI system is used within financial services.

3.2 The benefits and risks are divided into different categories based on each of the supervisory authorities' objectives, namely consumer protection, competition, safety and soundness of firms, insurance policyholder protection, financial stability, and market integrity.

3.3 The benefits, risks, and harms discussed in this chapter are neither exhaustive, nor applicable to every AI use case. Moreover, while many of the benefits and harms presented in this chapter are neither new nor unique to AI, the use of AI in financial services may amplify existing risks and introduce novel challenges.

3.4 Previous work by the supervisory authorities has shown that the drivers of AI risk in financial services can occur at different levels within AI systems (see Figure 1), starting with the risks associated with the use of data to train, test, and run AI models; which can feed into risks arising from the design and use of AI models themselves; which can, in turn, lead to challenges for the governance structures necessary to manage those risks.

3.5 **Data** – Given that AI relies significantly on large volumes of data in its development (training and testing) and implementation, data-related risks can be amplified and have significant implications for AI systems. Drivers of data risk can include: errors in the training data, incomplete or unrepresentative data, significant outliers or noise, historical data biases, insufficient data, and more. Poor data preparation, validation, and management can also be drivers of risk.

3.6 **Models** – Poor AI model performance may result from data-related risks but also from a range of model-related risks. These could include inappropriate model choices, errors in the model design or construction, lack of explainability, unexpected behaviour, unintended consequences, degradation in model performance, model or concept drift, and more. As with data, poor model risk management (including validation, change management, and monitoring) can also be a driver of risk. The extent to which poor model performance leads to poor outcomes can depend on the degree of autonomy of a model.

3.7 **Governance** – At the governance level, the drivers of risk can include the absence of clearly defined roles and responsibilities for AI, insufficient skillsets, governance functions that do not include the relevant business areas or consider the relevant risks (such as ethical risks), a lack of challenge at the board and executive level, and general lack of accountability.

3.8 It is important to note that if the drivers of risks are mitigated, then AI may deliver a number of benefits. For example, the combination of high-quality data, appropriate model choices, and good governance can result in a well-performing AI system and accurate outputs. Moreover, both the benefits and risks that AI deliver are dependent on the context in which it is used and the purpose for which the technology is deployed. For example, issues relating to data quality within a consumer-facing AI system are likely to result in benefits and/or risks related to consumer protection. Whereas data quality issues within an AI trading system are likely to result in benefits and/or risks related to the safety and soundness of individual firms, and potentially even financial stability. Therefore, the following sections highlight some of the potential AI benefits and risks as they relate to the supervisory authorities' objectives and remits.

## Consumer protection – FCA

3.9 AI can harness the power of large volumes of data to identify characteristics about consumers and their preferences. This can be put to many uses, from providing access to financial services to consumers with non-standard histories, to identifying demographics with specific needs, or characteristics of vulnerability, and better product matching for consumers. At the same time, with a more granular understanding of individual consumers' characteristics, there may be a greater potential to identify and exploit consumer behavioural biases and characteristics of vulnerability– from exploiting inertia, to harmful price discrimination, to exploiting actual characteristics of vulnerability. Whether the use of AI is beneficial to consumers will depend on how it is used and for what purpose.

3.10 There is a risk that the use of AI could be associated with discriminatory decisions. Bias related to protected characteristics such as race or sex, could arise inadvertently during model development. This is because, even if these variables are excluded from the model, they could still be correlated from other data points and identify protected characteristics. Biased or discriminatory decisions can arise from bias in the underlying data on which the AI model is trained. Data may reflect historical biases in society, for instance where certain groups have been less able to access credit. Or the overall dataset may be unrepresentative. AI models can amplify the inherent historical biases in input data, potentially leading to biased decisions for consumers.

3.11 AI applications could be used in a way that excludes certain consumers. For example, AI-based insurance screening or credit provision could enable greater segmentation between 'low-risk' and 'high-risk' consumer groups. This may have implications for risk pooling and could lead providers to exclude or offer unaffordable premiums to 'high-risk' consumers. Personalisation (eg pricing specific to individual consumers) could also lead to some financial products not being offered to certain groups potentially resulting in unlawful discrimination based on protected characteristics.

## Competition – FCA

3.12 Consumer-facing AI systems such as those used in Open Banking might improve competition in a market through improving consumers' ability to assess, access, and act on information, which in aggregate, can increase competitive pressures on firms. Open Banking enables third parties to access consumers' transaction data (with their explicit consent) in a secured environment. Leveraging on these data to generate user-centric insights, firms can provide innovative and tailored solutions to end users. For example, enabling consumers to solve complex problems to optimise which financial product offers them best value or supporting consumers in making better financial decisions as well as providing access to financial markets with robo advisor tools, AI systems can remove some of the informational advantage which sellers have over buyers in markets with complex financial products.

3.13 **Academic research** has shown that by detecting price changes from rivals and enabling rapid or automatic response, AI systems could potentially facilitate collusive strategies between sellers and punish deviation from a collusive strategy.

3.14 Where AI is particularly relevant to a business' practice, the costs of entry (including both the staff and skills, as well as the data and technology itself) may be raised to a level that limits market entry with potentially harmful effects on competition. The CMA published a paper on **Algorithms: How they can reduce competition and harm consumers** and an accompanying **call for information** in January 2021.

### **Safety and soundness – PRA and FCA**

3.15 AI can help financial services firms create better decision-making tools, develop new insights, and new and/or better products and services for consumers. These benefits are largely due to the uplift in predictive power, more granular classification and segmentation, ability to capture non-linear relationships and the ability to analyse larger volumes of data and new data sources (such as unstructured and alternative data).

3.16 Financial services firms are already benefitting from the use of AI by **improving operational efficiency** across a range of areas, which can help reduce costs and processing times. AI models can also be more accurate in predicting default risk for **consumer** and **corporate** credit. These improved probability of default estimates could, in turn, lead to more appropriate modelling of regulatory capital.

3.17 There are a number of challenges and risks related to the use of AI, which may amplify prudential risks (credit, liquidity, market, operational, reputational, etc) and have implications for the safety and soundness of firms. For example, in credit risk, accuracy or consistency errors within the dataset will likely impact the ability of the AI model to accurately quantify the probability of default and risk of loss which could, in turn, lead to inaccurate capital modelling.

3.18 One of the strengths of some AI systems is their dynamic nature, including their ability to learn continuously from data. The nature of the hyperparameters,<sup>[2]</sup> functional form of the models, and their outputs can also adapt continuously. But this can also make such systems more susceptible to **data drift** and **concept drift** which in can in turn make the models and systems less stable. Moreover, the complexity of AI models is increasing, which can lead to a lack of explainability or interpretability (known as the 'black box problem').<sup>[3]</sup> The potential for drift, combined with the lack of explainability, can in turn lead to a range of prudential risks. For example, the use of complex and/or opaque AI systems to model probability of default and loss given default in IRB credit models could lead to firms having incorrect levels of regulatory capital.



## Insurance policyholder protection – PRA and FCA

3.19 AI in the insurance sector has the potential to improve the efficiency of data processing and decision-making in terms of both underwriting and claims processing. In life insurance, which includes an investment component, firms could leverage AI to support the investment choices of policyholders. In general insurance, AI could be used for automating claims management. Firms in the insurance sector can also use AI to analyse new unstructured data sources, like telematics or data collected from wearable devices, to provide more tailored products and/or pricing.

3.20 Insurers use AI across a range of business areas, which could pose risks to policyholder protection. Risks related to underwriting could lead to inappropriate pricing and marketing. For example, AI models trained on historical data may not account for a breakthrough healthcare treatment, which can lead to mispriced policies. Similarly, risks such as concept drift and lack of explainability in claims management AI systems could impact policyholders' ability to claim and their overall protection. Risks related to building AI models for cash-flow and capital reserve estimates could result in inaccurate predictions and reserve levels that could, in turn, impact insurers' ability to meet future liabilities.

## Financial stability and market integrity – the Bank and FCA

3.21 The benefits of AI for individual firms may also extend to the financial system. For example, more efficient processing of information by AI in credit decisions, insurance contracts, and customer interaction, may contribute to a more efficient financial system overall. Given the highly complex and non-linear nature of financial systems, AI can be a powerful tool for modelling macro financial risks and dynamics. AI's capacity to process and analyse large volumes of multidimensional data, means that it can be used to detect patterns in unstructured data and, for example, identify shifts in secular trends and narrative consensus.

3.22 Similarly, the use of AI by individual firms and within financial markets may amplify many of the existing risks to financial stability through various transmission channels (see Figure 2). For example, the use of similar datasets and AI algorithms may result in uniformity across models and approaches at multiple firms, which could amplify procyclical behaviour and lead to herding in certain use cases, such as algorithmic trading. Markets could also potentially become vulnerable to manipulation and prone to flash bubbles<sup>[4]</sup> or crashes if sentiment analysis and social media signals were used at scale in AI trading. The feedback loop<sup>[5]</sup> between the data and AI algorithm could exacerbate these effects.

3.23 Also, a further key challenge for firms lies in their ability to monitor operations and risk management activities that take place outside their organisations at third parties. Increased reliance on third parties, often outside the regulatory perimeter, for datasets, AI algorithms,

and other IT outsourcing (such as cloud computing) may amplify systemic risks. For example, operational failures and cyberattacks at critical third parties could result in disruption to certain AI services and therefore lead to a single point of failure that could impact multiple firms and markets. The supervisory authorities published [DP3/22 ‘Operational resilience: Critical third parties to the UK financial sector’](#) to explore these issues.

3.24 The potential risks associated with AI can also be assessed in relation to their outcome - the impact the technology has on markets, consumers or, indeed, the [environment](#). The [OECD](#) for example, has developed a baseline framework to advance a shared understanding of AI, including metrics for risks associated with the use of AI. Similarly, the [National Institute of Standards and Technology](#) is currently developing a framework to better manage risks to individuals, organisations and society associated with the use of AI. A key question is whether support for safe and responsible AI adoption is best delivered through a process-based or an outcomes-based framework, or indeed a combination of the two.

3.25 For an outcomes-based approach to AI, the following question arises: what are the most relevant metrics to measure impact, including what evidence is required to demonstrate good outcomes and how such evidence can be collected?

3.26 Finally, the use of AI by FMI firms, including in the clearing, settlement, and recording of financial transactions, may have implications for financial stability. Although the DP does not consider these issues in detail, FMIs are a vital part of the UK financial system and the supervisory authorities are keen to understand if there are any such benefits and risks specific to FMIs’ use of AI, so would encourage stakeholders with an interest in this sector to provide feedback.

## **Questions: Benefits, risks, and harms of AI**

Q3: Which potential benefits and risks should supervisory authorities prioritise?

Q4: How are the benefits and risks likely to change as the technology evolves?

Q5: Are there any novel challenges specific to the use of AI within financial services that are not covered in this DP?

Q6: How could the use of AI impact groups sharing protected characteristics? Also, how can any such impacts be mitigated by either firms and/or the supervisory authorities?

Q7: What metrics are most relevant when assessing the benefits and risks of AI in financial services, including as part of an approach that focuses on outcomes?

## 4. Regulation

---

4.1 This chapter discusses some of the current legal requirements and guidance that are considered by the supervisory authorities to be most relevant to mitigating the risks associated with AI. It also includes discussion of how financial services regulation sits alongside a body of cross-sectoral legislation and regulation, as well as the domestic and international context of emerging AI-specific legal requirements and guidance.

4.2 This chapter focuses on legal requirements and guidance that are applicable to PRA and FCA authorised firms, rather than Bank supervised FMI's. As with the previous chapter, the section headings relate to the supervisory authorities' objectives and remits.

### Introduction

4.3 As noted in the foreword, there is a wide-ranging debate about the regulation of AI. One of the central questions is whether the technology can be managed through extensions of existing regulatory frameworks, or whether a new approach is needed. This debate is happening across different sectors of the economy, and both here in the UK and other jurisdictions around the world.

4.4 Governments, regulators and other authorities have published numerous documents on the subject. These range from emerging regulations and laws (such as those in the [European Union](#), [People's Republic of China](#) and [Canada](#)) to cross-sectoral principles aimed at regulators (including the [UK government proposal](#) and [OECD principles](#)) and sector-specific principles for financial services firms (such as the ones issued by the [De Nederlandsche Bank](#), the [Hong Kong Monetary Authority](#) and the [Monetary Authority of Singapore](#)), as well as various policy papers and reports. This DP sits within the broader context of this global debate and focuses on a specific area – the regulation of AI within UK financial services.

4.5 At the same time, previous work by the supervisory authorities (including the [AIPPF](#) and [2022 ML survey](#)) have found that one of the challenges to the adoption of AI in UK financial services is the lack of clarity surrounding the current rules, regulations, and principles, in particular, how these apply to AI and what that means for firms at a practical level.

4.6 To help address this challenge, this chapter discusses some parts of the current regulatory framework that are considered by the supervisory authorities to be most relevant to the regulation of AI. The supervisory authorities are keen to gather feedback from stakeholders as to whether additional clarification of existing legal requirements and

guidance in respect of AI may be helpful, if the current regulatory framework could benefit from extension to better encompass AI, and how the supervisory authorities may best support the safe and responsible adoption of AI in UK financial services.

## **Box 4: Proportionality and the supervisory authorities' approach to regulation**

One of the regulatory principles under the Financial Services and Markets Act 2000 (FSMA) that the PRA and the FCA must have regard to in discharging their general functions is 'that a burden or restriction which is imposed on a person, or on the carrying on of an activity, should be proportionate to the benefits, considered in general terms, which are expected to result from the imposition of that burden or restriction'.<sup>[6]</sup> This principle of proportionality informs the supervisory authorities' thinking and approach to AI, including any potential future regulatory interventions.

Other financial services regulators and authorities have explicitly noted the importance of proportionality in relation to the regulation of AI. For example, the International Organization of Securities Commissions (IOSCO) guidance on the use of AI by market intermediaries and asset managers states the IOSCO members and firms should consider proportionality when implementing measures. It notes that firms and regulators should, in judging proportionality, consider the activity that is being undertaken, the complexity of the activity, risk profiles, the degree of autonomy of the AI applications, and the potential impact that the technology has on client outcomes and market integrity. Similarly, De Nederlandsche Bank (the central bank of the Netherlands) states that the applicability of their AI principles should be considered in light of the scale, complexity, and materiality of an organisation's AI applications.

## **Consumer protection – FCA**

4.7 The FCA's approach to consumer protection is based on a combination of the FCA's Principles for Businesses (the 'Principles'), other high-level rules, and detailed rules, and guidance. These include Principles and rules contained in the FCA Handbook. The Principles are general statements of the fundamental obligations of firms and other persons to whom they apply, who are liable to disciplinary sanctions if they breach one or more of the Principles.

4.8 The FCA has recently introduced new rules to raise standards for firms dealing with retail customers: the Policy Statement 22/9 'A new Consumer Duty' (the 'Consumer Duty'). These rules come into force for new and existing products and services that are open to sale

or renewal on 31 July 2023. The rules come into force for 'closed' products and services on 31 July 2024.

4.9 The Consumer Duty includes a new Consumer Principle, which sets a higher standard than the existing Principles 6 and 7 in terms of how firms need to treat retail customers. The Consumer Duty requires firms to play a greater and more positive role in delivering good outcomes for consumers, including (where a firm can determine or materially influence outcomes) those who are not direct customers of the firm. The Consumer Principle requires firms to play a greater and more positive role in delivering good outcomes for retail customers, including those who are not clients of the firm.

4.10 The Consumer Duty also includes cross-cutting rules requiring firms to act in good faith towards retail customers, avoid causing foreseeable harm to retail customers, and enable and support retail customers to pursue their financial objectives.

4.11 In addition, the Consumer Duty has rules relating to four key elements of the firm/consumer relationship which are instrumental in helping to drive good outcomes for customers:

- products and services designed to meet the needs of retail customers
- products and services to offer fair value for retail customers
- firms' communications to enable retail customers to understand products and services, their features and risks, and the implications of any decisions customers must make
- firms to provide ongoing support that meets their retail customers' needs.

4.12 The FCA also has the power to enforce the consumer protection requirements in the [Consumer Protection from Unfair Trading Regulations 2008 \(CPUTRs\)](#) These prohibit unfair commercial practices that involve misleading actions, misleading omissions of relevant information, or aggressive commercial practices.[7]

4.13 The subsections below outline at a high level how some of the Principles and the FCA's rules and guidance may be relevant to the AI risks to consumer protection discussed in Chapter 3.

## **Potential bias and vulnerability**

4.14 A number of the Principles may have particular relevance to this consumer protection risk, including:

- Principle 6: Customers' interests – '[a] firm must pay due regard to the interests of its customers and treat them fairly'.[8]

- Principle 7: Communication with clients – ‘[a] firm must pay due regard to the information needs of its clients and communicate information to them in a way which is clear, fair and not misleading’.
- Principle 9: Customers: relationships of trust – ‘[a] firm must take reasonable care to ensure the suitability of its advice and discretionary decisions for any customer who is entitled to rely upon its judgment’.
- Principle 12: Consumer Duty – ‘[a] firm must act to deliver good outcomes for retail customers’.

4.15 Principle 12 and the Consumer Duty apply from 31 July 2023 where firms deal with retail customers. Principles 6 and 7 continue to apply to conduct outside the scope of the Consumer Duty.

4.16 The Consumer Duty does not prevent firms from adopting business models with different pricing by groups (for instance risk-based pricing), but firms would need to ensure the price charged is reasonable, relative to the expected benefits (ie that products and services provide fair value to retail customers), including being able to justify the price offered to different groups of customers and considering those with characteristics of vulnerability or protected characteristics under the Equality Act. Certain AI-derived price-discrimination strategies could breach the requirements if they result in poor outcomes for groups of retail customers. As such, firms should be able to monitor, explain, and justify if their AI models result in differences in price and value for different cohorts of customers. Firms also have to take appropriate action under [PRIN 2A.4.25 R](#) to mitigate or remediate harm to existing customers and prevent harm to new customers if they identify that the product no longer provides fair value.

4.17 Prior to the Consumer Duty, the FCA had published its [Vulnerable Customer Guidance](#) which sets out what firms should do to comply with their obligations under the Principles and ensure they treat vulnerable customers fairly (including under Principle 6). The Vulnerable Customer Guidance notes that customers in vulnerable circumstances may be exposed to an increased risk of harm where firms do not understand the characteristics of vulnerability of their target market and main customer base, and so fail to ensure that products and services meet these needs.

4.18 The FCA notes in its Vulnerable Customer Guidance that it wants to see customers in vulnerable circumstances experience outcomes as good as those for other customers and receive consistently fair treatment across the firms and sectors it regulates – where firms’ use of AI does not take into account the differing needs and characteristics of such customers, those customers may be exposed to greater harm.

## Potential bias and discrimination risk

4.19 Discriminatory decisions made through using AI systems could be a breach of the [Equality Act 2010](#) (Equality Act), which protects individuals from discrimination on the basis of nine protected characteristics.[9] The EHRC is the body with primary responsibility for upholding equality and human rights laws in the UK.

4.20 The supervisory authorities are subject to a public sector equality duty under the Equality Act, which requires them to have due regard, in the exercise of their functions, to the need to eliminate discrimination and other conduct prohibited under the Equality Act, and advance equality of opportunity and foster good relations between persons who share a relevant protected characteristic and others. They take an interest in equality issues that arise in a financial services context. For example, the FCA's Vulnerable Customer Guidance notes that firms may need to have regard to the Equality Act, and seeks similar outcomes to the Equality Act's anticipatory duty on reasonable adjustments.[10] Some of the characteristics of vulnerability described in the FCA's Vulnerable Customer Guidance may overlap with protected characteristics under the Equality Act. Discriminatory decisions by AI systems that could lead to a breach of the Equality Act may, therefore, also breach the Principles or FCA rules and be subject to action from the FCA.

4.21 The Consumer Duty also addresses discrimination harms by requiring firms to consider the diverse needs of their customers, including the fair treatment of customers with characteristics of vulnerability and those with protected characteristics. Firms will be required to monitor the outcomes their customers receive in practice and take action if they identify particular groups of customers are getting poor outcomes. Firms designing products or services will need to define a target market and ensure the product or service meets the needs, characteristics and objectives of the target market.

4.22 The Consumer Duty is aligned with the detailed product governance requirements contained in the [Product Intervention and Product Governance Sourcebook](#) (PROD). Where a firm's product or service is subject to the rules in PROD 3, 4 or 7, it must continue to comply with those rules in respect of that product or service and the rules under the products and services outcome of the Consumer Duty do not apply to the firm for that product or service. However, the Consumer Duty as a whole is broader than the existing rules in PROD; satisfying the PROD rules is unlikely to mean a firm meets all aspects of the Consumer Duty. For example, firms would still need to consider elements of the Consumer Duty such as the consumer support outcome for their product or service, and to pay appropriate regard to the nature and scale of characteristics of vulnerability that exist in the target market.

## Financial exclusion

4.23 The FCA's Strategy 2022 to 2025 sets out four overarching outcomes expected from financial services. One of these is access, which includes meeting diverse consumer needs through low exclusion. Principles 6 and 12 are also likely to have particular relevance to financial exclusion risks related to AI.

## Consent and privacy

4.24 Where financial services firms use AI to process personal data, firms will have obligations under UK data protection law. For the ICO's guidance on AI, see [Explaining decisions made with AI](#) and the [Guidance on AI and Data Protection](#)

4.25 Certain practices may also breach the Principles or the FCA Consumer Duty, for instance where a firm did not present the way they would use customer data in a way that was clear, fair and not misleading, and used their data in ways to which they had not consented, and which was potentially to their detriment.

## Competition – FCA and PRA

### Relevant legislation

4.26 [The FCA has functions under the Competition Act 1998](#) in relation to agreements, decisions, and concerted practices that prevent, restrict, or distort competition, conduct that amounts to abuse of a dominant position, and transferred EU anti-trust commitments and directions, as they relate to the provision of financial services in the UK or the provision of claims management services in Great Britain. The FCA may use these functions in relation to applications of AI in financial services in the UK.

### Market studies

4.27 Under FSMA, the FCA has powers to carry out market studies.<sup>[11]</sup> The FCA also has concurrent functions with the CMA to carry out market studies for financial services under the [Enterprise Act 2002](#) which may be used to assess whether features of a market, including the use of AI, have or may have effects adverse to the interests of consumers.

4.28 Following a market study (whether under the Enterprise Act or under FSMA), the FCA can introduce remedies that could include market-wide remedies, such as rulemaking,<sup>[12]</sup> or firm-specific measures, such as the imposition of requirements. This is particularly relevant to some of the potential AI competition risks highlighted in Chapter 3. The FCA could, after completing a market study, also make a market investigation reference (MIR) to the CMA for an in-depth market investigation.



## **PRA secondary competition objective (SCO)**

4.29 In addition to its primary objectives, the PRA also has a secondary competition objective to facilitate effective competition in the markets for services provided by PRA-authorized persons in carrying on regulated activities. The SCO is set out in FSMA and came into force in March 2014.

## **Safety and soundness: Data – PRA and FCA**

4.30 Data are a crucial element of AI. From sourcing large amounts of data and creating datasets for training, testing, and validating, through to the continuous analysis of data once the AI system is deployed, the safe and responsible AI adoption in UK financial services is underpinned by high-quality data. Similarly, the increasing volumes of data involved in the use of AI means that data security and related issues such as data protection and subject privacy are ever more important to ensuring safe and responsible AI adoption. To the extent that AI processes personal information, financial firms will need to comply with their data protection obligations. The ICO has [published guidance on this topic](#). The [Data Protection and Digital Information Bill](#) includes a provision on AI and processing special category data for the purposes of monitoring AI bias.

## **Data quality, sourcing, and assurance**

4.31 Poor quality or inappropriate data can potentially compromise any process that relies upon those data. The way in which data are sourced and aggregated can impact the intended outcome, and the overall quality of the data. The current regulatory framework aims to address these specific risk components of the data lifecycle. For example, the Basel Committee on Banking Supervision's (BCBS)<sup>[13]</sup> ['Principles for effective risk data aggregation and risk reporting'](#) (BCBS 239) contains principles aimed at strengthening prudential risk data aggregation such as ensuring the accuracy, integrity, completeness, timeliness, and adaptability of data. The PRA expects the UK's globally systemically important banks, in particular, to adhere to these principles.<sup>[14]</sup> With respect to insurance, Rule 12.1 of the [Technical Provisions](#) Part and Rule 4.3 of the [Conditions Governing Business](#) Part of the PRA Rulebook for Solvency II require firms to have internal processes and procedures in place to ensure the completeness, accuracy, and appropriateness of the data used in the calculation of their technical provisions.

## **Data privacy, security, and retention**

4.32 Data security is important in ensuring information is protected from malicious threats such as unauthorised access, theft, and corruption. The privacy of this data and use of this data in a responsible manner is also important. Most notably, UK data protection legislation applies standards for data privacy and security in respect of personal data. Firms also need to consider their obligations under [The Money Laundering, Terrorist Financing and](#)

**Transfer of Funds (Information on the Payer) Regulations 2017** **The Payment Services Regulations 2017** (PSRs) are focused on security and quality of data transfers to third parties.

### **Data architecture, infrastructure, and resilience**

4.33 Data architecture and infrastructure refers to the system, standards, and policies in which data are stored, arranged, and integrated. Data resilience refers to the ability for data to be preserved after a failure or disruption. Regulation within this area is more general, requiring firms to have strong data architecture and risk management infrastructure.

Examples include the **Risk Control Part of the Rulebook** **Fundamental rules 5 and 6**, BCBS 239 principles, and **BCBS's Guidelines on 'Corporate governance principles for banks'**

### **Data governance**

4.34 Data governance can be summarised as a set of principles and practices that ensure quality throughout the lifecycle of data. Regulation within this area is broader, with principles such as BCBS 239 highlighting the governance structures and/or board responsibilities, and the UK GDPR setting out the accountability of controllers and processors for compliance with data protection requirements.

4.35 While there appears to be some thematic overlaps (eg on data quality, most notably), the various regulations differ in focus and scope. For instance, the PSRs are focused on security and quality of data transfers to third parties; requirements contained within the Basel Committee's **Fundamental review of the trading book** revised market risk standard (FRTB) would only apply to market risk data;<sup>[15]</sup> UK GDPR and Data Protection Act 2018 only apply to personal data; BCBS 239 only applies to systemically important banks and is focused on the aggregation of risk data; rules implementing Solvency II apply to technical provisions and internal model data of insurers; and **Markets in Financial Instruments Regulation** (MiFIR) obligations apply to trade data with the aim of improving protections for investors.

4.36 Organisational and legacy issues make total data integration a significant challenge. The prudential regimes (eg Solvency II) tend to lay out high-level principles on data, without prescribing detailed guidelines on their implementation.

### **Safety and soundness: Model risk management – PRA**

4.37 As the **AIPPF final report** notes, model risk management (MRM) is becoming increasingly important as a primary framework for some firms to manage and mitigate potential AI-related risks. However, the current scope of MRM regulation in the UK is very limited, with only principles and regulation in place for the use of models in specific areas or

tasks (eg internal capital models or stress-testing models). At the same time, MRM regulation does not explicitly mention AI, and there is currently no explicit guidance on issues like model explainability/interpretability of AI models.

### **Box 5: PRA Consultation Paper on MRM**

The PRA published CP6/22 on Tuesday 21 June 2022, which includes a proposed set of principles which it considers to be key in establishing an effective MRM framework. The principles are proposed to be embedded as supervisory expectations, as set out in a proposed new SS, and include:

- model identification and model risk classification
- governance
- model development, implementation, and use
- independent model validation
- model risk mitigants

The proposed principles would apply to all regulated UK-incorporated banks, building societies, and PRA-designated investment firms. The PRA proposes that the principles should be applied by firms in a way commensurate with their size, business activities, and the complexity and extent of their model use.

Moreover, the proposed principles cover all elements of the model lifecycle and would be applicable to all types of models that are used to inform key business decisions, whether developed in-house or externally (including vendor models), and models used for financial reporting purposes. This includes AI models and the CP includes a question concerning the adequacy of the proposed principles to address the risks associated with AI models.

To submit a response, please go directly to [CP6/22](#). The consultation closes on Friday 21 October 2022.

### **Identification and classification**

4.38 While the supervisory authorities do not currently provide an explicit definition of a model (the PRA proposed definition is being consulted on in CP6/22), the PRA expects that banks should give consideration to what factors would constitute a model, and must establish their own definition of a model, and, furthermore, banks should maintain a model inventory with clear information regarding, but not limited to, owners, users, uses, and direct or material

dependencies (see PRA [SS3/18 'Model risk management principles for stress testing'](#)). Similarly, PRA [SS5/18 'Algorithmic trading'](#) expects firms to define the term 'algorithm' (section 2.7(b)) as used by the firms in the context of algorithmic trading.

## **Effective governance framework, policies, procedures, and controls to manage model risk**

4.39 Effective governance provides support and structure to MRM activity through policies which define relevant risk management activities, procedures that implement those policies and the allocation of resources, and mechanisms for testing that policies and procedures are being carried out as specified.

4.40 Beyond PRA rules, the BCBS [Corporate governance principles for banks](#) states that it is the responsibility of a bank's risk management function to ensure that the board and management are aware of the 'assumptions used in and potential shortcomings of the bank's risk models and analyses' (paragraph 119). The guidelines also stress that 'risk identification and measurement should include both quantitative and qualitative elements' (paragraph 114).

## **Robust model development and implementation process**

4.41 A robust model development process ensures models are developed to appropriate standards and use representative data, which are important considerations for AI models.

## **Model validation and independent review**

4.42 The validation and independent review of an AI model is important in order to ensure an objective view is given on the model, inclusive of the way in which it is developed and that it is suitable for the intended purpose.

4.43 The IOSCO Board, an international standards-setting body for securities regulation, has also developed guidance for regulators on supervising the use of AI and ML by market intermediaries and asset managers, specifically on the development, testing and ongoing monitoring of AI techniques. For example, [Measure 2 of the IOSCO guidance on AI](#) states that regulators should require firms to adequately test and monitor the algorithms to validate the results of an AI technique on a continuous basis, with any testing reflecting the underlying complexity and systematic risks posed by the use of AI. In particular, IOSCO recommends the implementation of 'kill switch' functionality in the control framework, which is similar to the existing expectations and requirements for algorithmic trading in [SS5/18](#) and [Commission Delegated Regulation \(EU\) 2017/589](#) [16]

## Safety and soundness: Governance – PRA and FCA

4.44 Good governance is essential for supporting the safe and responsible adoption of AI. This is because governance underpins proper procedures and effective risk management across the AI lifecycle, by putting in place the set of rules, controls, and policies for a firm's use of AI.

4.45 The supervisory authorities take a principles-based approach to governance. Principle 3 of the [FCA Principles for Businesses](#) requires states that '[a] firm must take reasonable care to organise and control its affairs responsibly and effectively, with adequate risk management systems'. Fundamental Rule 6 of the [PRA's Fundamental Rules](#) states that '[a] firm must organise and control its affairs responsibly and effectively'. Both the FCA Handbook and the PRA Rulebook contain provisions in respect of compliance, internal audit, financial crime, risk control, outsourcing, and record-keeping. These provisions include SYSC 4.1.1 and Rule 2.1 of the [General Organisational Requirements Part of the PRA Rulebook](#) which both state: '[a] firm must have robust governance arrangements, which include a clear organisational structure with well defined, transparent and consistent lines of responsibility, effective processes to identify, manage, monitor, and report the risks it is or might be exposed to, and internal control mechanisms, including sound administrative and accounting procedures and effective control and safeguard arrangements for information processing systems'. The general rules, guidance, and principles will be relevant to the firm's use of AI. There are also specific requirements of relevance to AI – for example, [MiFID Org Regulation](#) requires investment firms to store records in a way that it is not possible for them to be manipulated and altered (other than corrections or amendments) and which allows for IT or other efficient exploitation of the data when the analysis of the data cannot be easily carried out due to the volume and nature of the data.

### Board composition, collective expertise, and engagement

4.46 Concerning board composition, collective expertise, and engagement, [SS 21/15](#) sets out the expectations of the PRA in relation to how firms should comply with the rules in the [General Organisational Requirements](#) [Skills, Knowledge and Expertise](#) [Compliance and Internal Audit](#) [Risk Control](#) [Outsourcing](#) and [Record Keeping Parts](#) of the PRA Rulebook. Specifically, [FCA Handbook SYSC 21.1.2](#) details the role of the chief risk senior management function in challenging and ensuring the quality, quantity, and use of data.

4.47 As highlighted in the AIPPF final report, there may be a lack of understanding of the challenges and risks arising from the use of advanced technologies at firms' senior management and board levels, both individually and collectively, leading to a skills and engagement gap. This may be partly a cultural issue, and could lead to ineffective governance. There are requirements and expectations on firms to address this skill gap,

including the PRA expectations set out in PRA [SS5/16 'Corporate governance: Board responsibilities'](#) that boards should have the diversity of experience and capacity to provide effective challenge across the full range of the firm's business and boards should pay close attention to the skills of its members, and the FCA requirements for issuers to publish information on board diversity policies in their corporate governance statement.

4.48 The PRA and FCA could also use a range of supervisory tools to assess how firms are addressing their potential technology knowledge gap, including as part of board effectiveness reviews. Where appropriate, questions about the use of AI and relevant controls could be included in question banks for SMF interviews.

### **Who should be accountable for AI under the SM&CR?**

4.49 The supervisory authorities' existing rules and guidance, in particular, those implementing the SM&CR, emphasise senior management accountability and responsibility and are relevant to the use of AI. The SMR requirements are set out in the PRA rulebook eg for CRR firms [Allocation of Responsibilities](#) [Senior Management functions](#) etc and for SII firms some are listed under [Insurance – Allocation of Responsibilities](#) and [Insurance – Senior Manager Functions](#). The supervisory statements provide a source of guidance on the SM&CR eg PRA [SS28/15 'Strengthening individual accountability in banking'](#) and [SS35/15 'Strengthening individual accountability in insurance'](#) set out the PRA's expectations on strengthening individual accountability in banking and insurance. Specifically for international banks, PRA [SS5/21 'International banks: The PRA's approach to branch and subsidiary supervision'](#) states the PRA expectations on the accountability of Senior Management Functions (SMF) for branches and subsidiaries.

4.50 Within the SM&CR there is at present no dedicated SMF for AI. Currently, technology systems are the responsibility of the SMF24 (Chief Operations function – see the [PRA Rulebook Senior Management Functions Part, Rule 3.8](#) and [SUP 10C.6B 'Systems and controls functions: Other'](#)). Separately, the SMF4 (Chief Risk function) has responsibility for overall management of the risk controls of a firm including the setting and managing of its risk exposures. These functions apply to PRA-authorized SM&CR banking and insurance firms and FCA-authorized enhanced scope SM&CR firms, but not core or limited scope SM&CR firms.

4.51 PRA-authorized SM&CR banking and insurance firms and FCA-authorized enhanced scope SM&CR firms must ensure that one or more of their SMF managers have overall responsibility for each of the activities, business areas, and management functions of the firm. That means any use of AI in relation to an activity, business area, or management function of a firm would fall within the scope of a SMF manager's responsibilities.

4.52 For banks, the General Organisational Requirements (part 2.1) states that ‘firms must have robust governance arrangements, which include a clear organisational structure with well defined, transparent, and consistent lines of responsibility and effective control and safeguard arrangements for information processing systems’.

4.53 Part 5.1 lays out the board responsibilities stating that the board ‘oversees and is accountable for the implementation of governance arrangements, the prudent management of the firm, including the segregation of duties in the organisation and the prevention of conflicts of interest. And that the management body approves and oversees implementation of the firm’s strategic objectives risk strategy and internal governance’.

4.54 The [guidance from IOSCO](#) (Measure 1 – Governance and responsibilities) states that regulators should consider requiring firms to have designated senior management responsible for the oversight of AI development, testing, deployment, monitoring, and controls. This includes a documented internal governance framework, with clear lines of accountability. The guidance also states that firms should designate an appropriately senior individual (or groups of individuals) with the relevant skill set and knowledge to sign off on initial deployment and substantial updates of the technology. This measure ‘looks to embed accountability in all aspects of a firm’s use of AI and ML and helps ensure the technology (and its underlying data) is appropriately understood, tested, deployed and monitored’. Furthermore, ‘[t]his accountability extends to the actions and outcomes of AI and ML models, including externally sourced models’.

4.55 As the AIPPF final report noted, a key question for all firms is who should ultimately be responsible for AI and whether this should be a single individual or shared between several senior managers. A key question therefore will be how firms identify the relevant SMF(s) with responsibility for the use of AI in the business. Therefore, the most appropriate SMF(s) may depend on the organisational structure of the firm, its risk profile, and the areas or use cases where AI is deployed within the firm. This is without prejudice to the collective responsibility of boards and the respective responsibilities of each of the three lines of defence.

4.56 Looking ahead, there is a question as to whether there should be a dedicated SMF and/or a Prescribed Responsibility (PR) for AI under the SM&CR. Arguably, AI use may not have yet reached a level of materiality or pervasiveness to justify these changes. However, the AIPPF highlighted that the split in responsibilities is an area of uncertainty for firms and that more guidance on governance functions, roles, and responsibilities would help provide clarity.

4.57 Given the technical complexity of AI systems, it is important that staff responsible for developing or deploying them are competent to do so. One possibility for ensuring this could be a new certification function for AI, similar to the FCA’s certification function for algorithmic trading. The algorithmic trading certification function extends to persons who: (i) approve the

deployment of trading algorithms; (ii) approve the amendment to trading algorithms; and (iii) have significant responsibility for the management of the monitoring, or decide, whether or not trading algorithms are compliant with a firm's obligations.

4.58 PRA **SS5/18 'Algorithmic trading'** sets out expectations for governance (eg cross lines of defence coordination, SMFs, testing) with regard to the use of algorithms in the context of trading.

## AI Lifecycle

4.59 One useful approach to understanding firms' obligations, is to look at them from the perspective of the AI lifecycle. For example:

- pre-deployment: how should the quality of training data be assessed? How should AI models be tested before live deployment? Should AI models be 'compliant by design'? Who is the accountable SMF? Identification of and allocation of responsibility for new risks presented by AI.
- deployment: how should the performance of live AI systems be monitored? What safeguards should be introduced to monitor for, detect and stop potential harm eg kill-switch mechanisms?
- recovery and redress: if an AI system's performance leads to crystallised risks, should firms be required or expected to 'undo the damage' by (i) reversing decisions made by the model (where possible and appropriate); and/or (ii) compensating any relevant external parties who suffered damage as a result?

## AI and reasonable steps

4.60 The concept of 'reasonable steps' is a core element of the SM&CR. SMFs can be subject to enforcement action under S66A(5) and/or S66B(5) of FSMA if an area of the firm for which the SMF has responsibility breaches regulatory requirements and the FCA and/or PRA can demonstrate that they failed to take such steps as a person in the senior manager's position could reasonably be expected to take to prevent or stop these breaches.

4.61 One of the areas that could benefit the most from further discussion is what may constitute reasonable steps in an AI context and how, if at all, these steps differ from the reasonable steps that SMFs are generally required to take.

4.62 PRA SS28/15, SS35/15, and **DEPP 6.2 'Deciding whether to take action'** which are the key reference sources on the 'reasonable steps' criterion under the SM&CR, have detailed, but not exhaustive, expectations on what may constitute reasonable steps and on how firms and SMFs can document and evidence them. This guidance built on the PRA's,



FCA's, and FSA's prior enforcement activity and supervisory experience and was issued at a time when autonomous, decision-making technology such as AI was not as widespread and, as a result, the guidance does not explicitly refer to such technology.

4.63 A particularly useful approach could be to consider what may constitute reasonable steps at each successive stage of the lifecycle of a typical AI system.

### **Human-in-the-loop**

4.64 An important element in the operation of an AI system is the level of human involvement in the decision loop. Humans typically interact with a system in the design and training stages and may also be involved in operating the system and interpreting outputs. The human element can act as a valuable safeguard against harmful outcomes, providing contextual knowledge that may be outside the capability of a model, and identifying where an automated decision could be problematic, and therefore requiring further review.

4.65 Consumers and others affected by the decisions made by automated systems may feel uncomfortable where important decisions are made without a human-in-the-loop. In certain contexts human involvement can be a regulatory requirement, for instance under the UK GDPR, automated decisions are treated differently to human decisions. Article 22 of the UK GDPR restricts fully automated decisions which have legal or similarly significant effects on individuals to a more limited set of lawful bases and requires certain safeguards to be in place. The ICO has explained that the human input needs to be meaningful, and a decision does not fall outside Article 22 just because a human has 'rubber-stamped' it. The UK government is planning to reform these laws, including [amending the right to human intervention](#) in the [Data Protection and Digital Information Bill](#) which is currently at the second reading stage in the House of Commons.

4.66 It is worth noting, there are limits to the effectiveness of human involvement in an automated decision. Some reviewers may be subject to 'automation bias' where they simply accept automated recommendations or may be unable to effectively interpret the outputs of complex systems and falsely reject an accurate output. However, we think firms deploying AI systems need to have a sufficiently strong set of oversight and governance arrangements that make effective use of humans in the decision-making loop and review the accuracy of those arrangements.

### **Safety and soundness: Operational resilience, outsourcing, and third-party risk management – PRA and FCA**

4.67 Since 2018, the supervisory authorities have developed and implemented a coordinated regulatory and supervisory framework to strengthen the operational resilience of the UK financial services sector (see [joint covering document 'Operational resilience: Impact](#)

**tolerances for important business services'**). Operational resilience refers to the ability of firms and the financial services sector as a whole to prevent, adapt to, respond to, recover from, and learn from operational disruptions.

4.68 Operational resilience applies to the use of AI by firms and FMIs when it supports an important business service. This means firms and FMIs should set an impact tolerance for disruption for each of those important business services that involve AI, and ensure they are able to remain within their impact tolerances for each important business service in the event of a severe (or in the case of FMIs, extreme), but plausible disruption.

4.69 Many of the principles, expectations, and requirements for operational resilience may provide a useful basis for the management of certain risks posed by AI and support its safe and responsible adoption. For example, developing and implementing effective business continuity and contingency plans for AI systems that support an important business service. As the **AIPPF final report** noted, backup, and remediation actions being considered before the model is put into production can enable firms to respond appropriately and in a timely manner if an AI model performance or output deteriorates beyond the accepted risk threshold, which can help manage risks.

4.70 It is also worth noting firms and FMIs are expected to and/or required to meet applicable operational resilience requirements and expectations irrespective of whether the AI is developed in-house, or by third parties. Therefore, the supervisory authorities' requirements for outsourcing and third party risk management (see **SYSC 8.1** , **13.7** and **13.9** of the FCA Handbook (for CRR firms) the Outsourcing Part and (for Solvency II firms) Chapter 7 of the Conditions Governing Business Part of the PRA Rulebook, and PRA SS2/21) also apply to third party AI models used by firms.

4.71 Most recently, the supervisory authorities set out potential measures to oversee and strengthen the resilience of services provided by critical third parties ('CTPs') to firms and FMIs. The discussion in **DP3/22 'Operational resilience: Critical third parties to the UK financial sector'** explored measures that could complement, not replace, firms' and FMIs' existing expectations and/or requirements to manage potential risks to their operational resilience, including as a result of the impact of the failure or disruption of a third-party. The supervisory authorities note that certain third parties providing data and AI models could emerge as future potential CTPs as a result of the increasing use of these data and models.

## **AI Standards**

4.73 Industry technical standards (such as those issued by the **International Organization for Standardization and the International Electrotechnical Commission** ) can form part of a route to establishing common best practice for the development, deployment, and use of

AI systems, as well as a way for firms to signal their AI systems meet a certain benchmark of quality.

4.74 Standards can also potentially complement the regulatory system by helping firms build trust among users that their systems meet widely accepted standards, which may go above satisfying minimum requirements for regulatory compliance.

### **Questions: Regulation**

Q8: Are there any other legal requirements or guidance that you consider to be relevant to AI?

Q9: Are there any regulatory barriers to the safe and responsible adoption of AI in UK financial services that the supervisory authorities should be aware of, particularly in relation to rules and guidance for which the supervisory authorities have primary responsibility?

Q10: How could current regulation be clarified with respect to AI?

Q11: How could current regulation be simplified, strengthened, and/or extended to better encompass AI and address potential risks and harms?

Q12: Are existing firm governance structures sufficient to encompass AI, and if not, how could they be changed or adapted?

Q13: Could creating a new Prescribed Responsibility for AI to be allocated to a Senior Management Function (SMF) be helpful to enhancing effective governance of AI, and why?

Q14: Would further guidance on how to interpret the 'reasonable steps' element of the SM&CR in an AI context be helpful?

Q15: Are there any components of data regulation that are not sufficient to identify, manage, monitor, and control the risks associated with AI models? Would there be value in a unified approach to data governance and/or risk management or improvements to supervisory authorities' data definitions or taxonomies?

Q16: In relation to the risks identified in Chapter 3, is there more that the supervisory authorities can do to promote safe and beneficial innovation in AI?

Q17: Which existing industry standards (if any) are useful when developing, deploying, and/or using AI? Could any particular standards support the safe and responsible adoption of AI in UK financial services?

Q18: Are there approaches to AI regulation elsewhere or elements of approaches elsewhere that you think would be worth replicating in the UK to support the supervisory authorities' objectives?

Q19: Are there any specific elements or approaches to apply or avoid to facilitate effective competition in the UK financial services sector?

## 5. Questions

---

### **Supervisory authorities' objectives and remits**

Q1: Would a sectoral regulatory definition of AI, included in the supervisory authorities' rulebooks to underpin specific rules and regulatory requirements, help UK financial services firms adopt AI safely and responsibly? If so, what should the definition be?

Q2: Are there equally effective approaches to support the safe and responsible adoption of AI that do not rely on a definition? If so, what are they and which approaches are most suitable for UK financial services?

### **Benefits, risks, and harms of AI**

Q3: Which potential benefits and risks should supervisory authorities prioritise?

Q4: How are the benefits and risks likely to change as the technology evolves?

Q5: Are there any novel challenges specific to the use of AI within financial services that are not covered in this DP?

Q6: How could the use of AI impact groups sharing protected characteristics? Also, how can any such impacts be mitigated by either firms and/or the supervisory authorities?

Q7: What metrics are most relevant when assessing the benefits and risks of AI in financial services, including as part of an approach that focuses on outcomes?

### **Regulation**

Q8: Are there any other legal requirements or guidance that you consider to be relevant to AI?

Q9: Are there any regulatory barriers to the safe and responsible adoption of AI in UK financial services that the supervisory authorities should be aware of, particularly in relation to rules and guidance for which the supervisory authorities have primary responsibility?

Q10: How could current regulation be clarified with respect to AI?

Q11: How could current regulation be simplified, strengthened and/or extended to better encompass AI and address potential risks and harms?

Q12: Are existing firm governance structures sufficient to encompass AI, and if not, how could they be changed or adapted?

Q13: Could creating a new Prescribed Responsibility for AI to be allocated to a Senior Management Function (SMF) be helpful to enhancing effective governance of AI, and why?

Q14: Would further guidance on how to interpret the 'reasonable steps' element of the SM&CR in an AI context be helpful?

Q15: Are there any components of data regulation that are not sufficient to identify, manage, monitor and control the risks associated with AI models? Would there be value in a unified approach to data governance and/or risk management or improvements to the supervisory authorities' data definitions or taxonomies?

Q16: In relation to the risks identified in Chapter 3, is there more that the supervisory authorities can do to promote safe and beneficial innovation in AI?

Q17: Which existing industry standards (if any) are useful when developing, deploying, and/or using AI? Could any particular standards support the safe and responsible adoption of AI in UK financial services?

Q18: Are there approaches to AI regulation elsewhere or elements of approaches elsewhere that you think would be worth replicating in the UK to support the supervisory authorities' objectives?

Q19: Are there any specific elements or approaches to apply or avoid to facilitate effective competition in the UK financial services sector?

- 
1. A convolutional neural network is a type of AI/ML deep learning technique used in image recognition, classification, and processing that is specifically designed to process pixel data.
  2. Hyperparameters are parameters used in AI/ML to control the model selection and training process. They are specified prior to training iterations unlike parameter values calculated as part of model fitting.
  3. See also the [FCA's article titled 'Explaining why the computer says 'no''](#) for a discussion of the 'black box' problem in relation to consumer protection and potential approaches to AI explainability for financial consumers.
  4. A bubble is an economic cycle that is characterised by the rapid escalation of market value, particularly in the price of assets. This fast inflation is followed by a quick decrease in value of the assets, or a contraction.
  5. A feedback loop refers to the process by which an AI/ML model's predicted outputs are reused to train new versions of the model.
  6. [Section 3B\(1\)\(b\) of the Financial Services and Markets Act 2000](#)

7. In order to enforce these Regulations, the FCA must show that the action has contravened the requirements of professional diligence and materially distorted or is likely to materially distort the economic behaviour of the average consumer with regards to a product, or it must be a misleading action, omission or aggressive sales practice, or a commercial practice listed in Schedule 1 of the CPUTRs. Schedule 1 of the CPUTRs includes a list of practices that are deemed to be automatically unfair regardless of their actual or likely effect on consumer behaviour.
8. Principle 6 and Principle 7 are replaced by Principle 12 from 31 July 2023 for retail businesses.
9. Protected characteristics under the Equality Act 2010 are: age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion and belief, sex, and sexual orientation.
10. See the [Memorandum of Understanding \(MoU\) between the EHRC and the FCA](#)
11. The FCA can investigate markets where competition may not be working well for consumers and intervene where appropriate. More information about FCA market studies can be found in [FG 15/9: Market studies and market investigation references: A guide to the FCA's powers and procedures](#)
12. For example, in relation to the ways in which firms engage with consumers, the reduction of barriers to entry and expansion, and measures that directly control outcomes such as price or service standards.
13. BCBS is a standard setting in the banking sector and its membership comprises of 45 central banks and banking supervisors from 28 jurisdictions, including the UK. Since the BCBS does not possess any formal supranational authority, its decisions do not have legal force. The BCBS, however, expects its members and their internationally active banks to implement standards in a full, timely and consistent manner.
14. PRA [CP17/16 'Regulatory reporting of financial statements, forecast capital data and IFRS 9 requirements'](#).
15. The PRA will [consult](#) on the UK implementation of Basel 3.1, including FRTB, in quarter 4 of 2022. As part of the CP, a planned implementation date of 1 January 2025 will be consulted on.
16. For further detailed regulatory requirements relating to algorithmic trading, see [MAR 5.3A](#) for multilateral trading facility (MTF) operators, [MAR 5A.5](#) for organised trading facility (OTF) operators, [REC 2.5](#) for recognised investment exchanges (RIEs), [Commission Delegated Regulation \(EU\) 2017/584](#) for operators of trading venues (MTFs, OTFs, and RIEs) and [MAR 7A.3](#) for firms engaged in algorithmic trading.

Appendices

Spreadsheet