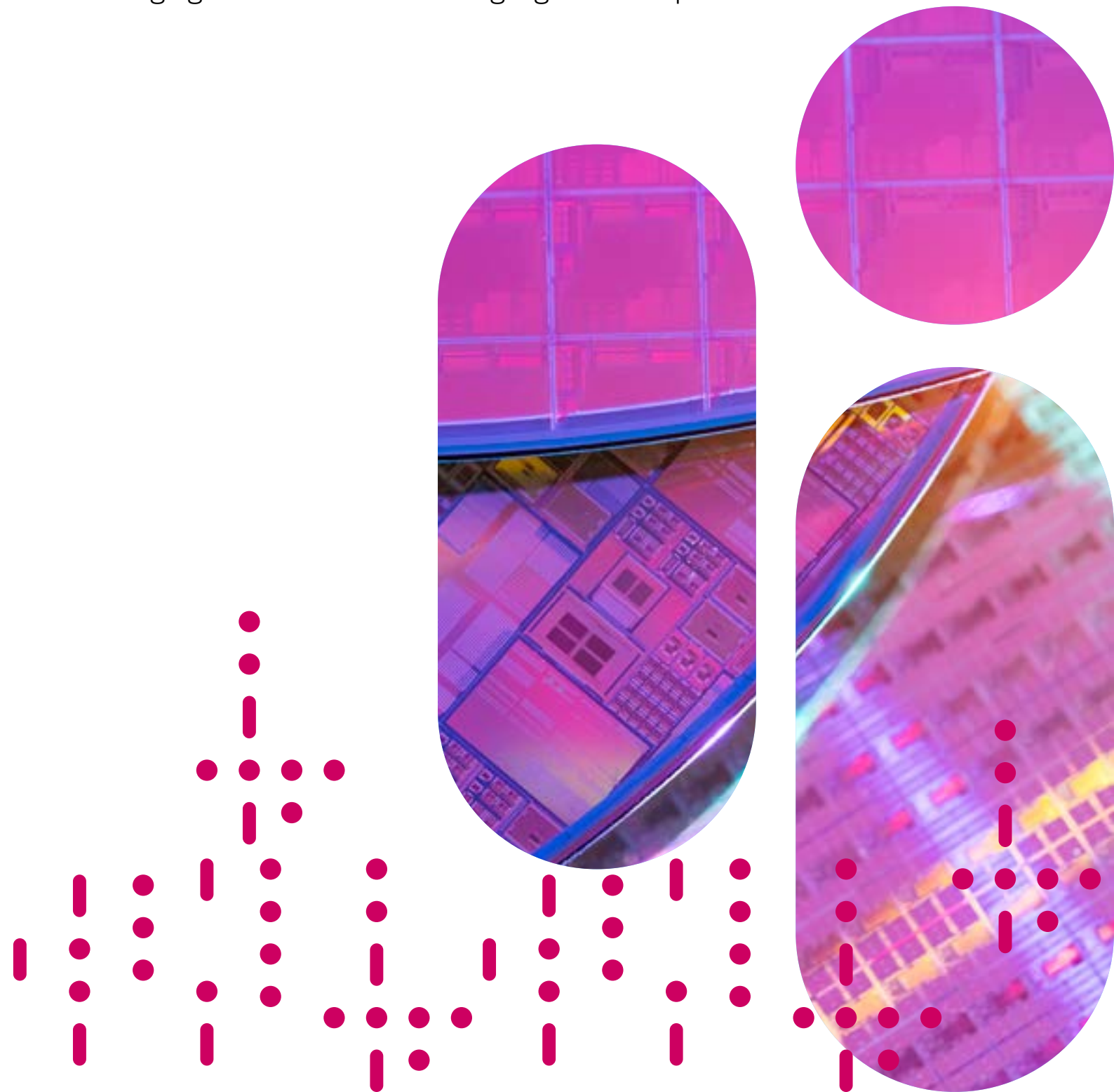
Linklaters

AI in Financial Services **4.0**

Managing machines in an evolving legal landscape



Contents

1 The AI revolution in finance – opportunities and challenges

2 The global legal landscape for AI

3 AI and financial services regulation

4 Ensuring AI is compliant with global data protection laws

5 Regulating AI through competition law – key issues to consider

6 Practical guidance on managing legal risk in AI

Introduction

The explosion of generative AI models and developments in agentic AI are driving a new wave of digital transformation. Regulators across the globe are striving to keep up, particularly in the highly regulated financial services sector. In the fourth edition of this report we partner with IBM, whose research demonstrates how AI-first institutions are already outperforming peers. We highlight both the opportunities and challenges facing businesses deploying AI in core financial markets, offering a broad overview of the complex and fast-changing regulatory landscape, along with practical guidance for managing risk.

The popularity of consumer and increasingly enterprise platforms, such as ChatGPT, Google Gemini, Perplexity and Claude, has accelerated adoption across sectors, making AI widely accessible. Use cases are developing in retail finance while AI developments are also unlocking key B2B, wholesale and operational efficiencies in financial services. Firms are scaling AI beyond pilots to enterprise-wide rollouts with measurable business impact. AI models are now being integrated into agentic workflows, involving teams of 'digital workers' working autonomously, enhancing decision-making and improving regulatory responsiveness well beyond simple responses to instructions.

The surge in interest and investment in AI has brought a renewed focus on its risks: from existential threats to humanity to more immediate issues like misinformation, biased outputs, and privacy and intellectual property protection.

There are ongoing calls for specific legal guardrails to safeguard consumers – and societies at large – from AI's potential harms. Meanwhile, rising geopolitical tensions between AI superpowers are intensifying the focus on national AI supremacy and hindering efforts at international cooperation on global standards.

Deployments of AI in financial services raise complex legal and regulatory issues, compounded by the rapidly evolving global landscape. Legal teams must now help to anticipate, mitigate and embed AI-related legal, operational and reputational risks into enterprise risk frameworks. As AI embeds into core business functions, Chief AI Officers (CAIOs) are emerging to align strategy, governance and execution. This report offers a high-level approach to identifying relevant issues, and practical guidance on managing AI deployment risks in the financial services context.

We are delighted to partner with IBM, an AI-first organisation aiming to become the world's most productive company through agentic AI. IBM is uniquely positioned as both a technology innovator and a strategic adviser, helping clients scale AI to achieve measurable financial returns. As noted by the Wall Street Journal, IBM's approach to agentic AI is reshaping how businesses orchestrate workflows, reduce operational friction and accelerate value creation. Gartner has also recognised IBM's leadership as a Leader in the 2025 Magic Quadrant for Data Science and Machine Learning Platforms and Generative AI Consulting.

For tailored advice or to explore how these insights apply to your organisation, please reach out to any of the Linklaters and IBM contacts listed.

Read more: Tech Legal Outlook 2025 Mid-Year Update: US, Asia, Europe

1 The AI revolution in finance – opportunities and challenges

AI adoption is accelerating in financial services, a data-rich industry with a long history of machine learning. With advanced deployments of AI increasing, and significant process mapping already in place, the industry is primed to lead in generative AI, notably AI agents. But, given the highly regulated nature of the sector where consumer protection and market stability are regulatory priorities, firms must understand and manage the associated legal risks.

A pivotal juncture for financial markets and services

The term 'artificial intelligence' is a broad one and often defined in slightly different ways. Perhaps the most internationally ubiquitous definition of an AI system is by the OECD, adopted in 2023 and updated to include the generating of content, clarify the role of input, account for autonomous operation and align with international regulations like the EU AI Act (discussed in Chapter 2): "*a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments*".

A 2023 report by Tony Blair and William Hague heralded AI as "*the most important technology of our generation*" with "*a level of impact akin to the internal combustion engine, electricity and the internet*" which is a view now widely accepted in IBM's own research and client engagements.

In 2025 some trends are emerging:

- **AI use is on the rise**: Just as the internet once brought "*a computer on every desk and in every home*", generative AI tools such as ChatGPT, Claude, and Perplexity are widely accessible to individuals. Meanwhile, enterprise-grade assistants like Microsoft Copilot or IBM's watsonx Orchestrate are being selectively adopted within organisations, depending on licensing, governance and readiness. This mainstreaming of large language models (LLMs) is unlocking productivity across industries, but also raising concerns around data security as individuals use these tools beyond business guardrails.
- **The pace of change continues to be high**: Despite claims of diminishing returns, AI capabilities continue to evolve quickly with new, more capable models being released faster than they can be thoroughly evaluated. Determining the 'best' is now more than a full-time job. The future of AI usage is likely hybrid: individuals and organisations will mix models and tools for specific needs, eg, using enterprise-grade assistants like IBM's watsonx Orchestrate for regulated internal workflows, Claude Code and other coding tools for software development, and consumer platforms such as Perplexity or Claude for exploratory tasks outside the enterprise environment. Even software providers are adopting this approach: Microsoft has announced Claude integration for certain applications in Office 365 while retaining GPT for others.
- **Generative AI is becoming democratised**: The creation of LLMs is no longer limited to the tech giants. DeepSeek's R1, released in January 2025, was reportedly trained at a fraction of the cost of major LLM providers, yet delivered competitive ranking seventh on a Hugging Face's leaderboard within eight months. This reflects broader democratisation of AI development, enabling financial institutions to build domain-specific capabilities without relying solely on proprietary platforms.
- **Open-source models are gaining ground**: Open-source models like Meta's Llama and IBM's Granite series are evolving faster than some proprietary LLMs, offering competitive capabilities with greater transparency. IBM's Granite series delivers enterprise-grade performance while supporting broader innovation. Designed for trust, enterprise-grade performance and adaptability, these systems are well-suited to regulated sectors like financial services.

- **Smaller models are making AI use economically viable**: Newer compact LLMs like Qwen-14B-Chat outperform larger predecessors such as Llama 2-70B in key tasks, offering similar or better results with lower compute demands. This enables significant cost and energy savings, making AI more viable and sustainable – especially for ESG-focused institutions. IBM highlights that smaller models not only cut costs but also support environmental goals by reducing energy use and carbon footprint.

IBM research on potential productivity gains – and talent gaps

AI agents (discussed below) are transforming enterprise and delivering real productivity gains. IBM's research shows that operating profit improvements attributable to AI have grown steadily from 2.4% in 2022 to 7.7% in 2024.

This automation revolution is particularly transformative in financial services, where AI agents must seamlessly access complex enterprise systems, comply with regulatory requirements, and adapt to real-time data inputs. Emerging standards like Model Context Protocol (MCP) could accelerate adoption by enabling seamless integration – like a USB-C for AI – without custom connectors.

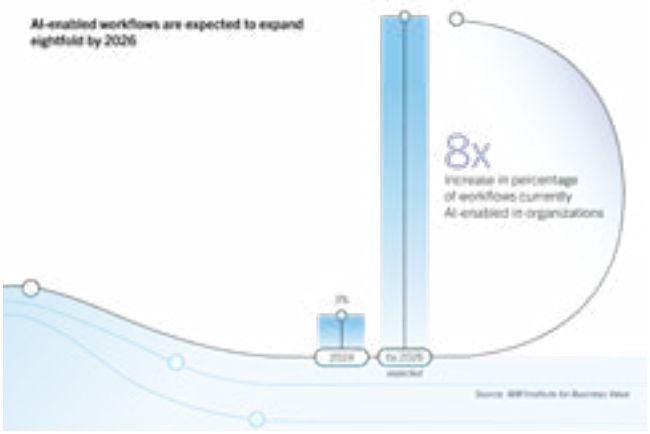
Read more: What is Model Context Protocol (MCP)? | IBM

However, innovating with AI must be balanced with strict regulatory requirements. Reflecting on risk and compliance is a prerequisite to scaling AI across the enterprise. This includes the methods for validating AI outputs and improving the skills of the workforce, where all of us will become AI users. This shift requires a rethinking of workforce roles and skills.

IBM research highlights talent gaps in model validation and risk control as key barriers to safe deployment. Legal and compliance professionals must develop fluency in AI risk management, validation and governance to both mitigate risk and shape responsible AI use across the enterprise.

AI-enabled workflows – many driven by agentic AI – are poised to expand from 3% in 2024 to 25% by 2026, and investment in AI continues to rise.

The eightfold increase in AI-enabled workflows suggests not just incremental improvement but a fundamental rethinking of business processes across functions – as reflected in the sustained rise of AI investment, which has grown to about 12% of IT spend. This suggests that business and AI senior executives remain convinced of AI's long-term transformative potential.



Read more: The Impact of AI | IBM

From AI projects to profits | IBM

AI adoption trends in financial services

Moving through the AI hype cycle in financial services

Following the extreme hype around generative AI as a potentially transformational technology, with huge impacts for business across all sectors, we are naturally experiencing more realism and pragmatism as organisations start really testing the technology's capabilities and its limits. Adoption appears more modest than originally anticipated but, that will also be impacted by the investment required in the infrastructure – technological, governance and data – for applications at scale.

Existing areas of deployment

A wide range of AI-related activity is reported across the financial services sector, with adoption varying significantly across institutions, sub-sectors and jurisdictions. Examples include:

- **Risk-management**: One of the earliest areas of adoption, this includes tools to monitor, detect and manage a variety of risks such as operational, market, credit or regulatory risk.
- **Customer on-boarding and engagement**: AI is used in the verification of know-your-customer information and in customer communications, including through chat-bots.
- **Insurance**: AI is used to support sales (eg, to improve the risk-sensitivity of pricing) as well as claims management (eg, in streamlining pay-outs triggered by real-world events).
- **Asset management**: AI techniques are used to support portfolio management, historically analysing past performance data but increasingly using other data sources and techniques.
- **Algorithmic trading**: Rules-based algorithms have long been used in market trading. AI techniques are now being used, including in 'algo-wheels' that select between alternative trading strategies.
- **Advisory**: Many robo-advisers use rules-based algorithms. Where AI techniques are used, they can create outputs to inform decision-making ultimately taken by humans.
- **Customer service**: The capability of AI agents to comprehend customer queries and make calls on relevant knowledge bases and utilise access to customer records means that many customer-related tasks are now performed with little to no human intervention.

Industry examples of AI in financial services

- **IBM**: has worked with a number of leading financial institutions in AI deployments, including Lloyds Banking Group, Nationwide, NatWest Bank and Virgin Money
- **Citi**: entered a strategic multi-year partnership with Google Cloud aimed at accelerating its digital transformation through cloud computing and AI, with use cases including developer toolkits, document processing and customer services (October 2024)
- **ING**: collaborated with McKinsey to build and scale a GenAI-powered customer services assistant projected to impact 37 million customers across 10 markets (2024 – 2025)
- **UBS**: using Open AI and Synthesia to create AI-generated scripts and avatars of its analysts in order to scale its video offering for clients. This is intended to achieve both efficiencies and to respond to greater customer demand for research in video format in the TikTok era (January 2025)
- **Mastercard and Visa**: introduced, respectively, Agent Pay and Visa Intelligent Commerce payment infrastructure designed for agentic commerce, where autonomous AI agents can execute transactions on behalf of their respective customers (April 2025)
- **Stripe**: launched its Payment Foundations Model, trained on tens of billions of transactions and designed to detect hundreds of nuanced signals per payment to enable improvements in fraud detection, authorisation rates and overall payment performance (May 2025)
- **Revolut**: is reported to be actively exploring the use of AI agents to automate customer service and sales processes. It has established an AI department and is hiring for roles to build AI products, including AI agent for call centres, outbound sales and voice assistants (September 2025)

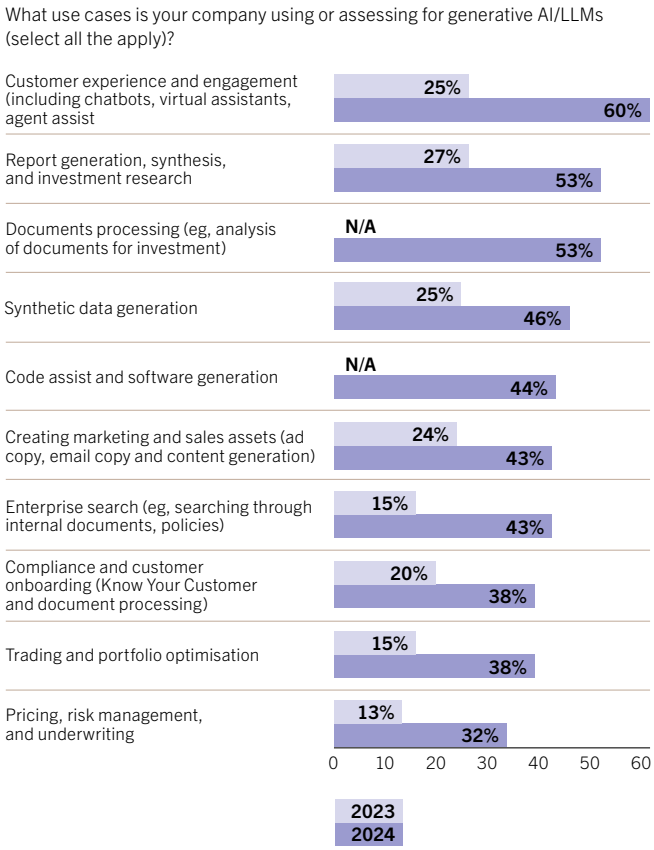
Key findings of industry surveys

75% of UK firms responding to a 2024 Bank of England survey said that they use AI, with a further 10% planning to use AI over the next three years. Foundation models (including LLMs) now make up 17% of AI use. The same report gave the following details on generative AI use cases in financial services:

Read more: AI in financial services survey results shine light on third party risks and AI governance, (November 2024)

At a global level, NVIDIA has conducted a survey of 600 global financial services professionals in their State of AI in Financial Services Survey Report (2025) which shows that, globally, generative AI is "standing out as a pivotal technology", with over half of their respondents currently using generative AI, up 40% from last year. They cite the example of PayPal, which has updated its AI infrastructure to achieve a "reduction of up to 70% in cloud costs and 35% in runtime".

Top generative AI use cases



State of AI in Financial Services, Survey Report (2025), Nvidia

AI use cases with the greatest impact in financial services

IBM's global research across financial institutions also shows that AI is rapidly evolving from isolated pilots to enterprise-wide transformation. As firms seek to unlock measurable value, five use cases consistently emerge as delivering the greatest impact across operational resilience, regulatory compliance, and client engagement.

1. **Fraud detection and cyber security**: AI systems continuously monitor transactions and network activity to detect anomalies, flag suspicious behaviour and prevent financial crime. **Benefit**: Enables real-time threat identification, reduces financial losses and strengthens institutional trust. 61% of banking executives cite fraud detection as the top driver of business value.
2. **KYC and AML compliance with agentic AI**: Agentic AI orchestrates autonomous agents to streamline customer due diligence, document verification and risk assessment. **Benefit**: Reduces review timelines from weeks to hours, improves accuracy and ensures regulatory alignment, which is especially critical in high-risk jurisdictions.
3. **Customer service and engagement**: Generative and agentic AI agents deliver personalised, context-aware support across channels, accessing customer records and knowledge bases. **Benefit**: Enhances customer satisfaction, reduces operational costs and enables 24/7 service with minimal human intervention.
4. **Asset and portfolio management**: AI models analyse structured and unstructured data to optimise portfolio strategies, pricing and risk exposure. **Benefit**: Improves decision-making speed and precision, supports dynamic asset valuation and unlocks new investment insights.
5. **Operational resilience and IT automation**: AI agents diagnose and resolve IT issues, predict maintenance needs and automate root cause analysis across infrastructure. **Benefit**: Reduces downtime, improves system reliability and cuts cloud and runtime costs as demonstrated by firms like PayPal and IBM's own observability platform.

The evolution from Assistants to Agents to AGI

From generative AI to agentic AI

Agentic AI is the next step in the proliferation of AI across both enterprises and consumer usage, moving beyond generative AI's content creation to autonomous task execution. While generative AI *responds* to prompts, agentic AI *acts* toward goals with minimal human input. Users provide a goal and operating space, but agents independently devise strategies, refine them through feedback and execute tasks via tools, APIs or GUIs – often using LLMs as their reasoning core.

Agentic AI is already being applied in Web2 environments – desktops, browsers and enterprise systems – with Web3 use cases emerging. Companies like Anthropic, Google and OpenAI are leading the way: Claude can interact with desktops. Google's Project Mariner navigates websites and OpenAI's Operator executes multi-step tasks online.

This evolution builds on generative AI's promise of faster ROI, enabled by foundation models and techniques like retrieval augmented generation (RAG). Just as generative AI democratised access through intuitive interfaces, agentic AI represents a major step forward in capability and autonomy.

However, scaling agentic AI requires careful attention to risk and compliance. Validation methods and workforce upskilling are essential, as AI becomes embedded in daily roles. Legal and compliance professionals must develop fluency in AI governance to ensure responsible deployment and mitigate risks.

What makes agents different from AI assistants?

AI assistants and AI agents serve different functions and interact with users in distinct ways. AI Assistants, like Siri or Alexa, are designed for individuals and are reactive (although becoming more proactive) responding to user prompts to perform tasks.

In contrast, AI agents operate with greater autonomy – given a goal and environment, they plan, reason and act independently, refining strategies based on feedback. Agents can operate in complex environments and handle more sophisticated workflows, like supporting fraud detection in financial transactions or trading execution decisions under human oversight.

As technology advances, the line between assistants and agents is blurring. Modern systems can switch between reactive and proactive modes depending on task complexity, although prompting strategies still differ. Assistants need precise instructions, while agents determine optimal paths themselves. OpenAI's ChatGPT Agent, for example, was trained using reinforcement learning on final outputs, reducing the need for detailed prompt engineering.

How can AI agents be used?

A confined AI agent, without access to external sources of data, is limited in much the same way as an assistant. The key benefit of an agent is its use of tools which allows it to operate within an ecosystem and, with the right tools, frameworks and protocols in its environment, use those capabilities to take action and use outside knowledge. This is often in the form of company policies, financial data, personnel records, client data or perhaps the Internet.

Credit card providers Visa and Mastercard have announced agentic shopping experience where users may tell an agent what they want and the agent will search for that item and pay for it. Agents are becoming commonplace in handling customer service inquiries. 71% of executives anticipate that the process of handling customer support inquiries will be fully automated by the end of 2026.

Anthropic has developed Claude for Financial Services, a financial analysis agentic AI solution able to call on research agents and a plethora of financial data from third-party suppliers. Such solutions are being utilised by leading financial services organisations in the delivery of their services to customers, allowing them to cut research costs and respond more quickly.

With the launch of protocols such as Google's Agent Payments Protocol (an open, payment-agnostic standard for agent-led transactions) and OpenAI and Stripe's Agentic Commerce Protocol (enabling users to buy products directly within ChatGPT), the prospect of consumers being able to fully delegate financial decisions to AI agents is quickly becoming a reality.

With agentic commerce set to transform e-commerce, and the role of agentic payments key to this, there are a number of legal and regulatory issues to be addressed.

Read more: Agentic payments: what are they, what are the legal risks and what's next? (August 2025)

Agents with far reach

AI agents operate around the clock and the globe, so automated operational functions never take a break.

The development of Model Context Protocol (MCP) is allowing AI agents to connect seamlessly with external tools, databases, the Internet or APIs and legacy systems, dramatically expanding their capabilities. Acting like a universal interface (akin to a USB-C port), MCP allows agents to discover and interact with resources without custom integrations. In financial services, for example, an AI agent handling loan processing could use MCP to:

1. retrieve applicant data from customer databases;
2. access credit scoring APIs;
3. check regulatory compliance databases; and
4. generate reports using document tools,

all through standardised MCP servers, eliminating the need for bespoke connections. These tools can be reused across departments, enhancing interoperability and productivity.

By combining MCP with agent-to-agent protocols and workflow orchestration, organisations can build networks of agents that automate complex, multi-system tasks – from HR to customer service. For instance, a host agent could co-ordinate with HR and scheduling agents to book holidays or meetings, reducing administrative overheads.

This shift is why executives increasingly see HR and customer service as prime candidates for automation. As agents become more capable, they will transform how businesses operate, freeing employees to focus on higher-value work.

However, much of the concern around widescale adoption of rapidly evolving AI stems from this shift in capabilities toward independence. Just as the work of humans is overseen by those with more experience and guided by training, the same will be true in scaling agentic AI, where strong AI governance will be critical (see Chapter 6).

Read more:
Understanding Agentic AI: The power of autonomy. (March 2025)

Looking ahead to AGI

The next frontier may be Artificial General Intelligence (AGI), which are AI systems capable of performing any intellectual task that a human can, across domains. While AGI has not yet been achieved, advances in multi-modal models, agentic systems and reasoning capabilities are accelerating progress. AGI could one day autonomously interpret regulations across jurisdictions in real time, or design financial products without human input, raising new legal, ethical and operational considerations.

Recent advances in LLMs and multi-modal systems have prompted renewed speculation about timelines with much of the artificial intelligence community discussing whether AGI could emerge as early as 2027. Although timelines remain speculative, the majority of AI researchers surveyed place AGI emergence somewhere between 2030 and 2070, reflecting significant uncertainty.

Achieving AGI will require substantial breakthroughs in several critical areas: persistent memory systems that can retain and build upon long-term experiences, robust reasoning capabilities that generalise across diverse domains, effective transfer learning mechanisms that allow knowledge to apply broadly and comprehensive solutions for AI safety and alignment with human values. These challenges raise fundamental questions about intelligence, consciousness and the safe development of transformative technology.

Technological limitations and risks

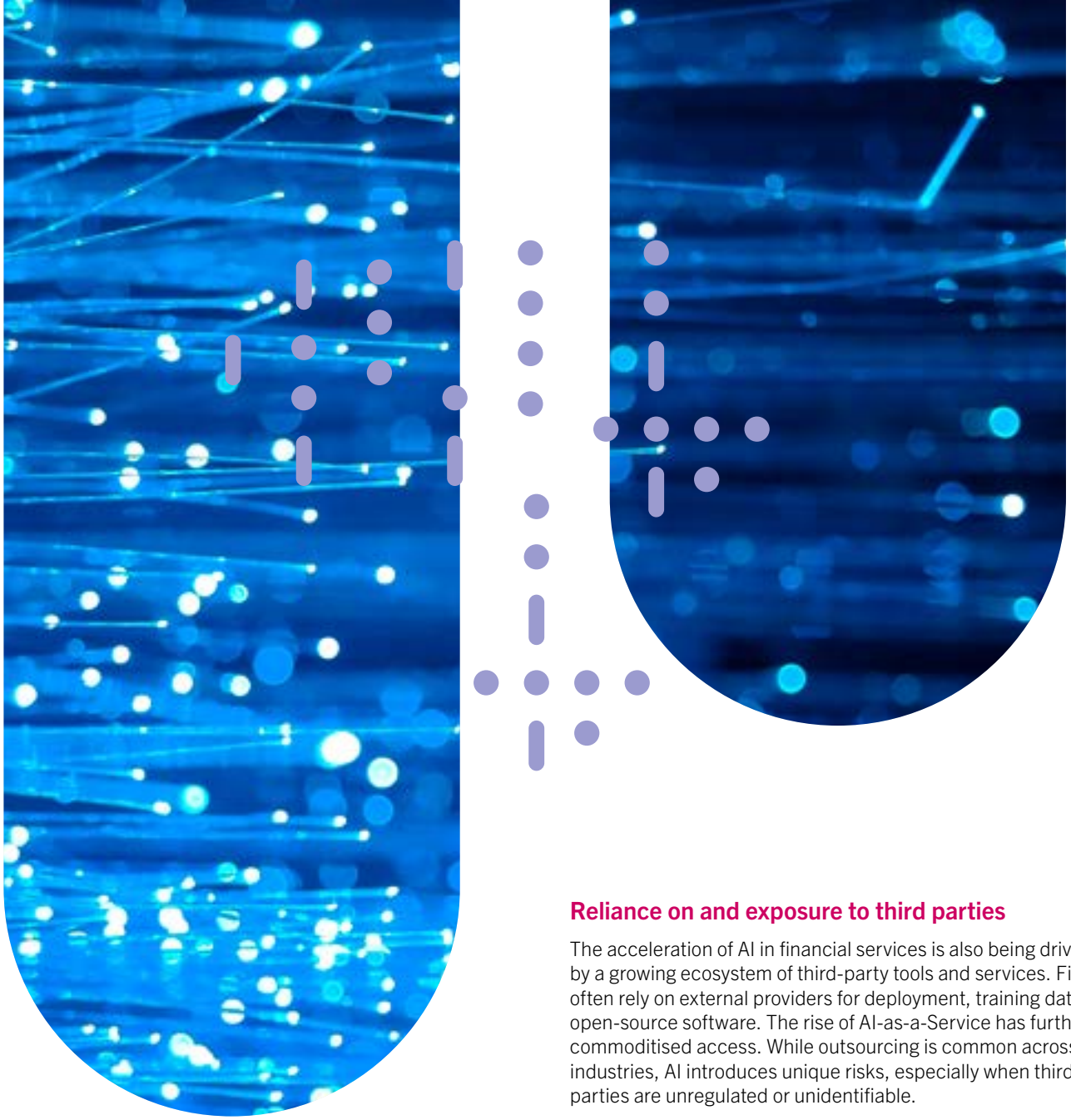
Features of machine learning techniques

There are certain technological features of machine learning techniques and deployments that raise particular legal and regulatory challenges, including when it comes to achieving regulatory compliance and identifying where liabilities fall. Addressing these challenges is not merely a 'nice-to-have': a failure to do so can have very real legal and regulatory consequences.

The nature of the challenges will depend on the precise application, but many of them stem from the following features of machine learning techniques:

- **Reliance on training data**: Unlike rules-based algorithms, machine learning algorithms are dynamic and their efficacy is highly dependent on the quality of the data through which they are trained. Such data may be obtained from a variety of sources and applied over a period of time. This may drive a need for new processes and controls to ensure that data quality is maintained to acceptable standards.
- **Predictability**: Whereas the outcomes of rules-based algorithms are predetermined, machine learning algorithms are designed to achieve a certain degree of accuracy and may also deliver different outputs in response to the same inputs over time (as the model is retrained with new data). These features may be incompatible with regulations that apply an absolute standard of compliance or require consistent results.
- **Explainability**: Advanced AI models may sacrifice explainability for performance: some outputs cannot be clearly traced to inputs, and future models may develop even more opaque reasoning. While chain-of-thought methods offer partial insight, reverse engineering of black-box models provides only limited transparency.

Firms deploying AI solutions will need to consider these novel features in determining the adequacy of their existing governance, oversight and risk-management frameworks from a legal and regulatory compliance perspective (see Chapter 6).



Reliance on and exposure to third parties

The acceleration of AI in financial services is also being driven by a growing ecosystem of third-party tools and services. Firms often rely on external providers for deployment, training data or open-source software. The rise of AI-as-a-Service has further commoditised access. While outsourcing is common across industries, AI introduces unique risks, especially when third parties are unregulated or unidentifiable.

This reliance raises concerns around control, compliance and liability. The Bank of England's Artificial Intelligence Consortium (AIC) has flagged concentration and accountability risks. Firms must ensure robust governance, including contractual and transparency measures, particularly when using third-party training data to avoid harmful biases.

As AI skills become more widespread, firms may shift toward customisation. Platforms like IBM's InstructLab allow organisations to fine-tune models such as Granite using proprietary data, offering competitive advantage and reducing third-party dependence.

See Chapter 3 for how operational resilience rules impact third-party relationships in financial services.

Read more: A new way to collaboratively customise LLMs – IBM Research

Risks amplified by generative and agentic AI

There are also risks specific to the use of AI, **which may be amplified by the use of generative and agentic AI**, which could lead to financial loss and reputational damage as well as legal sanctions:

Risks with generative AI

- **Employee experimentation**: With widespread access, employee experimentation can lead to unintended consequences. Iterative prompting may improve results, but without prompt governance – such as version tracking – those risks increase.
- **Unreliable outputs**: Models like ChatGPT can produce plausible but false 'hallucinated' information, which may be misused for harassment, defamation or misinformation, even potentially impacting market confidence.
- **Bias and discrimination**: Given the sheer size and variety of the datasets on which generative AI can be trained, and the fact that it generates content autonomously, the potential for bias and discrimination in generative AI outputs may be greater/harder to manage than with traditional AI.
- **Breach of copyright**: The use of copyrighted training data and the reproduction of protected content in outputs (eg, image regurgitation litigation, such as Getty v Stability AI) can constitute an infringement. IBM's Granite models, trained on proprietary data, offer IP indemnity, allowing firms to safely leverage their own data for competitive advantage.
- **Misuse of personal data**: Using public data to train AI models raises legal concerns, particularly around GDPR's accuracy principle and data erasure rights. If personal data is embedded in a model, removing or correcting it can be technically challenging (see Chapter 4).
- **Cyber security risks are evolving**: Generative AI enables more sophisticated malware, phishing, deepfakes and chatbot poisoning that can bypass traditional defences. Despite recent high-profile attacks, IBM's 2025 data breach report shows a 9% drop in breach costs (to $4.4M), thanks to faster detection and containment.

Risks with AI agents

- **Malicious use risks**: AI agents can be used for cyber offences, disinformation and manipulation of public opinion. Mitigating these risks can be challenging, especially for open-source products.
- **Malfunction risks**: AI agents may not always function as intended, with agent-to-agent dynamics increasing the chance of malfunction (eg, issues with control over simultaneous failures across different servers). LLMs also often have inherent biases that may impact agents' function and result in disparate performance.
- **Systemic risks**: Systemic risks associated with the introduction of AI agents, particularly at the enterprise level, range from environmental and intellectual property risks to considerations for privacy and labour market impacts.
- **Lack of transparency**: The decision-making processes of AI agents, especially those using complex algorithms like deep learning, can be opaque or 'black boxes', making it hard to understand why a particular decision was made. This lack of transparency can undermine trust and hinder debugging when errors occur.
- **Bias and fairness**: Like other AI systems, AI agents can inadvertently perpetuate or even amplify existing biases if trained on biased data. This can lead to unfair outcomes in critical areas like employment, lending or law enforcement.
- **Security vulnerabilities**: AI agents, especially those connected to the internet or internal networks, can be targets for cyber attacks. Malicious actors might manipulate inputs, extract model data or cause the agent to behave in harmful ways, leading to security breaches and resulting damage.
- **Dependence and over-reliance**: Over-reliance on AI agents can lead to skill loss among human operators, and create systemic vulnerabilities. If agents fail, supported systems may struggle. Ongoing training on AI capabilities and limitations is essential to maintain resilience.
- **Ethical dilemmas**: AI agents can face ethical dilemmas, especially when navigating conflicting ethical principles or values with no clear right answer. This is particularly challenging in advisory roles where understanding client context is critical.
- **Scalability and resource intensity**: Advanced AI agents may struggle to scale in resource-limited environments due to high compute demands. Without clear governance, predicting running costs is difficult, making deployment risky where efficiency is critical.
- **Legal and regulatory compliance**: The rapid evolution of AI technologies can outpace legal and regulatory frameworks, leaving organisations uncertain about how to ensure that their AI agents comply with existing laws and standards.

Weighing the opportunities and the challenges – a boardroom issue

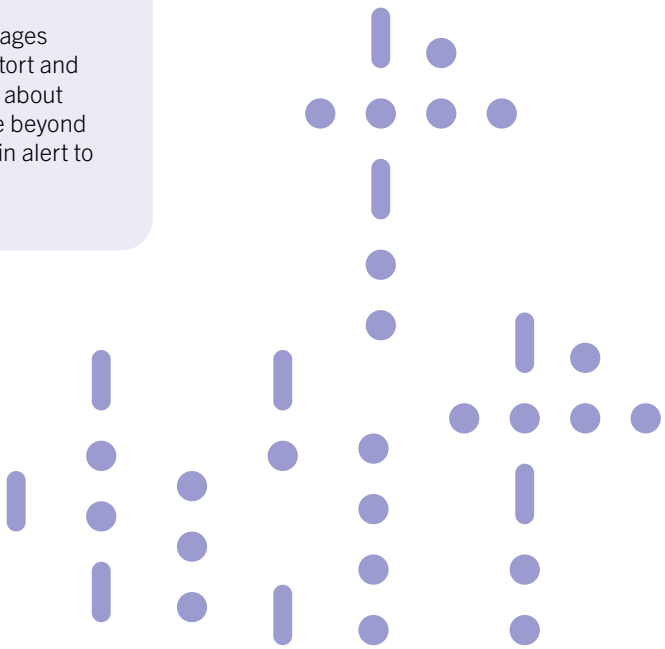
AI technologies are presenting firms with a wide range of new opportunities, both in terms of cost- and risk-reduction, as well as revenue generation. Firms may need to embrace these types of opportunities to remain competitive. However, the considerable new challenges these technologies present warrant careful consideration at board level. In particular:

- **Heightened risk of failure**: AI failures in financial services can be severe, potentially leading to consumer discrimination, market abuse or breaches of regulatory capital requirements. These risks are amplified by the novel and complex nature of AI, which can make issues harder to detect and control.
- **Evolving regulation**: Regulatory frameworks are continuing to evolve in response to the novel features of machine learning discussed above and as AI technologies continue to develop. Approaches to regulation differ across different regions and may change regularly. In addition, there are various soft law standards to consider (see Chapter 2).
- **Compliance with existing financial regulation**: Existing financial services regulation already sets high standards for governance, risk management and control, outsourcing and cyber security. However, AI's novel features can test the adequacy of these frameworks and, in some cases, may be incompatible with current compliance requirements (see Chapter 3).
- **Senior manager accountability**: Under senior manager accountability regimes, individuals may bear regulatory responsibility for the AI that is deployed in their remit. Senior managers must be proactive in overseeing AI-related risks (see Chapter 3).
- **Cross-sectoral regulation**: As discussed in Chapters 4 and 5, depending on the application of AI, various cross-sectoral laws and regulations may come into play, such as data protection and competition law (see Chapters 4 and 5).
- **A broad spectrum of legal risk**: AI use engages multiple legal regimes, including contract, tort and product liability, raising complex questions about accountability when things go wrong. While beyond this report's scope, legal teams must remain alert to these issues (see Chapter 6).

As early as 2019, the Bank of England's Executive Director for UK Deposit Takers, James Proudman, directly addressed the governance implications of adopting AI and machine learning technologies within the financial services sector in his Managing Machines speech. He highlighted the following three principles to guide regulatory focus – matters that still deserve board attention in 2025 and beyond:

- **Data usage**: Since AI poses challenges to the proper use of data, boards should attach real priority to the governance of data – what data should be used, how it should be modelled and tested and whether the outcomes derived from the data are correct.
- **The role (and responsibilities) of people**: Boards should continue to focus on the oversight of human accountabilities within AI-centric systems.
- **Transition risks**: Boards should reflect on the skills and controls that are necessary to oversee the transition. Many of the challenges raised by this transition can only be brought together at, or near, the top of the organisation.

AI's greatest potential lies not in making existing processes marginally more efficient, but in enabling entirely new approaches to value creation and delivery.



2 The global legal landscape for AI

The challenge for lawyers advising business is both to understand how existing legal and regulatory frameworks apply to AI and to keep up with new AI-specific law and regulation as it develops at an international, regional and national level.

Evolving legal and regulatory framework

AI-specific law has arrived

Many governments see great potential for AI to drive economic development and to solve societal challenges. They want to provide the legal framework needed to encourage innovation, attract investment and enable growth. At the same time, they recognise that there is a need to protect their citizens and to address the ethical, legal, societal and economic issues associated with AI.

There is pressure to move quickly, but not rashly. Policymakers face the challenge of keeping up with technological development and establishing a sustainable legal framework that neither dampens innovation nor quickly becomes outdated or unenforceable as new technologies emerge.

While principles-based regulation is more adaptable, commentators have argued that technology-specific regulation could be clearer and more effective. With the adoption of the EU's AI Act (read more below), South Korea's AI Basic Act (due to take effect in January 2026 – read more) and various state-level AI laws in the US (read more below), the era of AI-specific legislation has arrived.

Regulation of AI in financial services

Financial services is a key sector for AI (see Chapter 1). Use of AI in the financial sector has primarily been governed by the application of existing laws and regulation (see Chapter 3). Firms have also applied self-regulation by adhering, for example, to the voluntary Responsible AI principles first published by Microsoft in 2023, which have since been operationalised through the Microsoft Responsible AI Standard in shaping Microsoft's own AI products.

With AI-specific laws and rules now emerging, financial services providers must account for an additional layer of regulation in their risk management frameworks.

How regulators have been responding to the disruptive effect of AI in finance

An active response

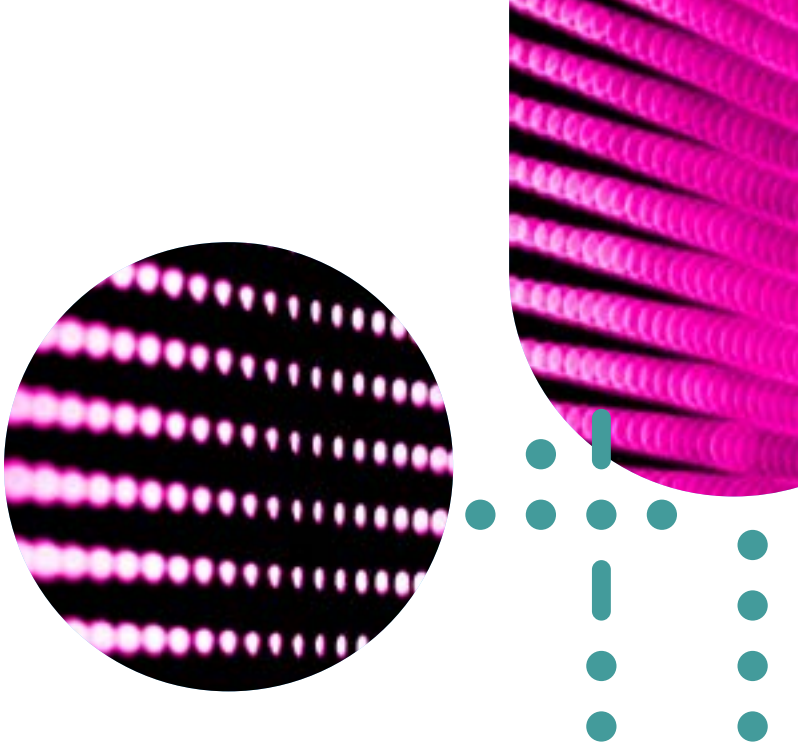
Even within the tech industry itself, regulation has for some time been considered necessary to both support and encourage AI and to manage associated risks. There has been much concern around bias in the underlying technology of machine learning algorithms (in respect of training data, testing and the decision model) and the issue of explainability in machine-led decision-making, as well as potential impacts on the stability of the wider financial ecosystem and anti-competitive effects (see Chapters 3 and 4).

We have seen an active response from regulators in **financial services, data protection and antitrust** as they attempt to balance innovation against the protection of consumers, wider society and the financial system's role therein.

Moving to experimentation

Regulators across the world have been moving from the more traditional 'wait and see' approach to what has been described by the Alan Turing Institute as a 'test and learn' experimentation approach to learn about various new technologies. They have created various testing environments such as test beds, living labs and digital and regulatory sandboxes which provide a means to both support innovation and test what regulatory guardrails are needed.

In the UK, for example, in June 2025 the FCA in collaboration with NVIDIA launched a Supercharged Sandbox which is open to any financial services firm looking to innovate and experiment with AI. The Supercharged Sandbox is one of the components of the FCA's AI Lab initiative. As directed under the framework of the EU AI Act (see below), EU Member States (for example, Spain) are also actively engaged in developing and implementing national regulatory sandboxes for testing AI systems in a controlled environment.



Competing policy objectives

While its flagship AI Act is now the principal tool to regulate AI in Europe (see below), the EU is also developing a swathe of new regulations specifically aimed at the digital economy, building on the General Data Protection Regulation (GDPR) which focused on the protection of personal data (see Chapter 4).

These interventions are primarily driven by consumer protection concerns, something which has been at the top of the regulatory agendas in recent years. Antitrust concerns about the dominance of Big Tech, including their foray into finance via payments, has also driven both enforcement action and bold regulatory proposals in several key jurisdictions.

However, there can be competing policy objectives from different regulators in the mix – particularly tension between competitive markets and the protection of personal data – and there is a growing recognition that we need greater regulatory cooperation between these supervisory authorities (see UK approach below).

International regulatory initiatives – and the issue of fragmentation

There have been numerous initiatives across the globe at international, intergovernmental, regional and national levels addressing the issues presented by AI.

A long way from international consensus

The OECD published the first intergovernmental standard for AI policies in the form of the OECD Principles on AI (2019), which have been endorsed by 49 countries. Regulators looking to support innovation therefore need to focus not only on national concerns but also on maintaining close dialogue with other jurisdictions, with a view to aligning AI regulation and ideally to developing global standards for AI governance.

Achieving regulatory harmony is challenging where there are variations in public acceptance and use of technologies across different countries and cultures. This includes, for example, different societal attitudes to the balance of privacy versus convenience. Differences in regulatory philosophy, economic priorities, and geopolitical interests continue to shape the landscape, leading to divergent regulatory approaches.

The coming years will likely see further convergence on technical standards and best practices, but greater harmonisation of AI regulation across the global will require sustained diplomatic effort and ongoing dialogue among governments, industry, and civil society.

Efforts at international collaboration

There have been various multi-lateral and often overlapping initiatives seeking to drive consensus around the approach to the safe use of AI. In 2023, the G7 leaders confirmed the Hiroshima AI Process, aiming to bring OECD countries together in a forum to start defining internationally acceptable principles to apply to the ongoing development of AI.

The UK also hosted the first global AI safety summit in 2023 which was followed by the Paris AI Action Summit in February 2025, where more than 100 nations and stakeholders launched major initiatives like the EU's €200 billion InvestAI programme and France's Public Interest AI Platform. The next global AI summit, in Seoul (November 2025), will focus on AI mega trends, industry applications and cutting-edge technologies like generative AI and robotics.

The United Nations has taken notable steps to foster inclusive global dialogue on AI governance. In 2025, it launched two new AI governance bodies – the Global Dialogue on AI Governance and the Independent International Scientific Panel on AI – aimed at bringing together all 193 member states, industry, and civil society to shape international rules and standards.

Divergence between Europe, the US and China

In 2025, however, we are still a long way from reaching any form of international consensus on the rules for AI. The rise in geopolitical tensions between the major AI superpowers is making that consensus harder to achieve than ever. As a result, diverging regulatory approaches towards AI give rise to material compliance challenges, particularly for multi-national businesses.

In 2023, we saw the EU and US working on voluntary AI code of conduct, but that has seemingly stalled. The EU has moved ahead with its own agenda, issuing a swathe of guidance to support its AI Act and hoping to set a global benchmark for trustworthy and rights-respecting AI. In 2025, the US released a US focused AI action plan focussed on promoting AI innovation, infrastructure, national security and US global leadership. Meanwhile, China proposed a global AI cooperation organisation and released a Global AI Action Plan, noting that global AI governance is still fragmented and emphasising multilateralism and support for the Global South. This three-lane approach will continue to play out in a way which is almost impossible to anticipate.

Regional and national AI strategies

At the time of publishing, according to the OECD, Governments in over 70 countries and territories (including the EU) have published national strategies on AI, and over 900 AI policy initiatives between them. Many involve consulting experts and industry, proposing ethics and principle-based guidelines, and identifying changes needed to existing law and regulation to enable the use of AI.

We have selected six prominent examples of jurisdictions with AI-specific strategies: the EU, the UK, Mainland China, Hong Kong SAR, Singapore, and the US, on which we will focus in this report.

EU approach

From soft principles to hard law

After several years of focus on ethical issues in AI, and a white paper in 2020 stressing the need to avoid fragmentation in the uptake of AI across the EU economy, the EU enacted its flagship AI Act in 2024. This aims to ensure consistency of rules for AI across the EU and support the development and adoption of AI across the whole EU economy. Its provisions are being phased in over time:

2 February 2025
Banned practices + literacy

2 August 2025
GPAI + governance

2 August 2026
General application (incl. standalone high-risk AI systems)

2 August 2027
Embedded high-risk AI systems

A risk-based approach

The AI Act establishes new obligations for both developers and deployers of AI systems, with significant requirements for 'high-risk' AI models. In-scope businesses are preparing for compliance with substantive provisions as they become applicable.

A global leader

As part of its broader digital regulation agenda, the EU was the first jurisdiction to adopt a cross-sector AI-specific regulatory framework. Since 2018, the EU's GDPR has governed the use of personal data which is central to development and running of AI systems, and already anticipated the automated decision-making it enables. But while the AI Act overlaps conceptually with the GDPR (with respect to personal data) and other digital regulation such as the Data Act, it introduces distinct requirements for implementing and managing AI systems within organisational frameworks.

AI liability rules

Amendments to the Product Liability Directive, which came into force in December 2024, modernises the EU's liability rules to address new technologies, including AI, and provide better protection for victims of defective products. It also introduces provisions for software, AI and digital services, and alleviates the burden of proof for victims. A proposed AI Liability Directive aimed to harmonise national rules on non-contractual civil liability for AI-related harm. However, its claimant-friendly approach drew strong opposition from the tech industry as overly burdensome, leading to its withdrawal in early 2025.

The EU AI Act – Impact for in-scope businesses

The AI Act introduces sweeping changes with a comprehensive and mandatory regime. It has the potential to impact a large number of businesses based in the EU as well as those established in a third country where the AI system is located in the EU or where the outputs produced by systems in a third country are used in the EU.

Most notably, all businesses using AI may need to assess the risk category in which their AI systems fall, to determine what their relevant obligations are under the Act.

Risk-based approach

The new rules aim to regulate AI systems according to the level of risk they present: five risk tiers set proportionate requirements and obligations depending on the role in the AI value chain. Systems that present unacceptable risks are banned and strict requirements are imposed on those considered to be high-risk, with lower-risk systems still subject to transparency requirements (see diagram on p19).

Significant and extensive compliance obligations for use of 'high-risk' AI:

The rules are cross-sector and not primarily focused on financial services but impact all aspects of the use of AI – in particular, where that use could infringe human rights and freedoms.

Of most interest to the financial sector is the use of AI for evaluating the **creditworthiness of individuals** and AI systems for risk assessments and pricing of life and health **insurance** – are within the high-risk regime. Other 'high-risk' systems include those used for **employment-related** purposes, such as advertising vacancies and screening applications in recruitment, and for making decisions on promotion and the termination of work contracts, the allocation of tasks and monitoring and evaluating performance and behaviour.

Firms deploying 'high risk' AI systems will be subject to **significant and extensive compliance obligations**, including requirements to take steps to mitigate harm; to use high-quality data sets for training; to keep records and logs of decisions and detailed technical documentation on the system and its purpose; to be transparent to users by providing clear and adequate information (eg, that they are interacting with an AI system); to have appropriate human oversight measures; and to ensure a high level of robustness, security and accuracy. For **regulated financial institutions**, however, many of these obligations align with existing financial regulatory controls (see Chapter 3), effectively easing some of the burden for those with mature governance processes.

Failure to comply with the legal requirements set for high-risk AI systems could result in fines of up to EUR 15 million or 3% of annual global turnover, whichever is greater.

Potential simplification

The European Commission has recently launched a simplification agenda for the EU rulebook. This is expected to comprise several proposals to amend existing legislation, including the AI Act. The aim is to "ensure a predictable and effective application of the AI Act, aligned with the availability of all the necessary support and enforcement structures". This could include postponing the application of some provisions of the Act, such as those on high-risk AI rules which are currently due to apply from August 2026.

What and who is in scope?

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

This is in line with the definition used by the OECD, which is expected to provide harmonisation and acceptance at international level. Recitals also state that reference to 'predictions' includes content, which intends to ensure that generative AI models, such as ChatGPT, fall within the scope of the AI Act.

The AI Act applies to a range of actors involved in the development, distribution and use of AI systems: providers and deployers (including third-country providers and deployers where the output of the system is used in the EU), importers and distributors, product manufacturers, authorised representatives of providers and affected persons that are located in the EU.

Provider

Develops AI systems or general purpose AI models **and** places them on EU market **or** puts them into service under own name or trademark

Product manufacturer

Embeds AI system in product under own name or trademark

Deployer

Uses an AI system under its authority

Importer

Places on the EU market AI system with name or trademark of person outside EU

Distributor

Person other than provider or importer making AI system available on EU market

Guidance for compliance

> **Code of Practice for General Purpose AI**: The EU has published a voluntary framework to help providers comply with the AI Act.

Read more: EU AI Act General Purpose AI model rules now apply, with new Code of Practice designed to support compliance (August 2025)

> **Guidance on AI literacy requirement**: The AI Act mandates that providers and users of AI systems must ensure that their staff have a clear understanding of the AI they employ.

Read more: EU guidance suggests firms may need to train clients and service providers on AI (May 2025)

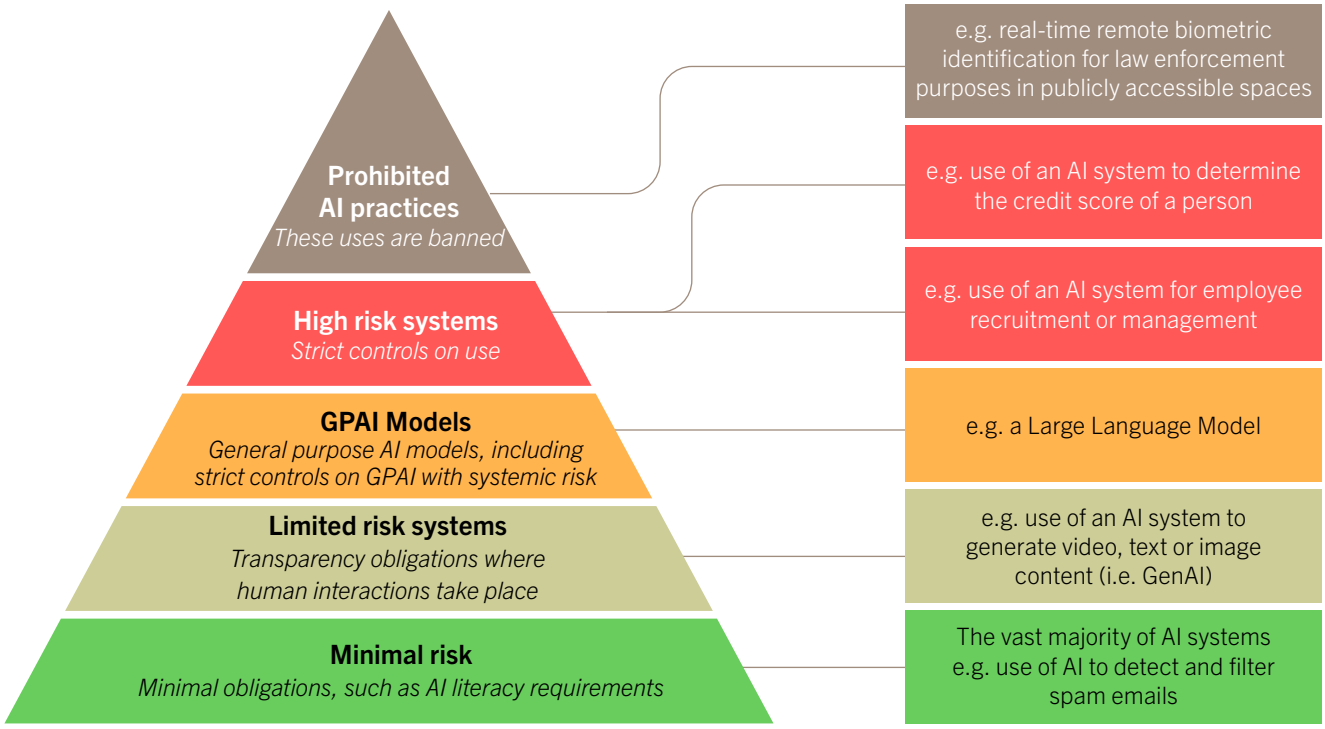
> **Guidance on high-risk AI**: Following a survey seeking input on high-risk AI rules, guidance is expected in 2026 to clarify how high-risk AI systems should be classified and regulated under the AI Act.

Read more: EU seeks input on high-risk AI rules (June 2025)

> **Guidance on the definition of 'AI system'**: The Commission has published guidelines regarding the definition of an 'AI system' under the EU's AI Act.

Read more: AI know it when AI see it: EU tries to clear up what an AI system is (February 2025)

EU AI Act – which tier applies?



UK approach

"Pro-innovation" approach

The UK continues to take a "pro-innovation" (ie, lighter touch), sector-driven, principles-based approach to AI regulation. Key digital regulators like the Financial Conduct Authority (FCA), Information Commissioner's Office (ICO) and Competition and Markets Authority (CMA) are engaged in dialogue and sandbox initiatives with no new rules anticipated. Any forthcoming AI-specific legislation is expected to focus narrowly on high-risk models and copyright issues, with limited direct impact on financial services.

Regulators' approach

UK regulators are taking a collaborative, non-prescriptive approach to AI. The Bank of England is assessing financial stability implications and the FCA has shared AI themes and launched an AI test service. The ICO, CMA and Ofcom continue to issue guidance focused on the risks associated with AI. HM Treasury is facilitating dialogue and live experimentation through structured engagement, sandbox initiatives and real-world testing pilots allowing firms to test AI in areas such as credit assessment, fraud and market monitoring within a supervised environment.

Impact of EU AI Act

The UK is keen to differentiate itself from the EU's more top-down model with a less prescriptive approach to AI regulation. However, given how close the two markets are, firms in the UK should be aware that they may still be caught by the broad scope of the EU AI Act which applies where: (1) the system is placed on the market in the EU; (2) the providers or users are physically present in the EU; or (3) the output of the system is used in the EU.

Private sector guidance

The Alan Turing Institute has published several key pieces for private sector guidance on AI in the UK focused on explainability, ethics and responsible development. This includes a collaboration with the ICO on best practice guidance for "Explaining decisions made with AI". The ICO has also published useful Guidance on AI and data protection (see Chapter 4).

Regulatory framework

The previous Conservative government favoured a 'pragmatic, proportionate' approach to AI regulation, proposed a principles-based context-specific framework in its AI white paper (2023) focusing on the use rather the technology itself. The succeeding Labour government's Modern Industrial Strategy (2025) builds on this, placing AI at the heart of its growth agenda promoting adoption in key sectors, supporting SMEs and startups in accessing AI tools and infrastructure and investing in national upskilling initiatives.

High level principles

The AI white paper set out five high-level overarching principles, which remain voluntary, but are being actively implemented and supported. The Department for Science, Innovation and Technology (DSIT) published initial guidance for regulators to help them interpret and apply these principles in their sectors. These are not legally binding, but regulators are encouraged to determine how to work the principles into their existing regimes and provide guidance specific to their sectors.

Regulatory cooperation in digital regulation

Financial regulators have been engaging actively with industry in considering whether AI-specific regulation is needed in the sector (see Chapter 3), while competition authorities are also increasingly looking at AI as they flex their regulatory muscle, particularly with respect to Big Tech in financial services (see Chapter 4). The FCA has also partnered with the Competition and Markets Authority, Ofcom and the ICO in the UK's Digital Regulatory Cooperation Forum (DRCF) on AI (see below).

Direction of travel

While promised regulation for frontier AI has yet to materialise, the policy direction is clear: regulators and businesses must self-regulate by adhering to the agreed-upon principles and preparing for increased oversight in the future. CEOs and compliance teams should view the current framework as a transition: maintain a rigorous, documented approach to AI governance now, stay engaged with regulators and ensure that all AI-driven services meet existing UK standards.

Work of the UK's DRCF on AI

This unique forum aims to collaborate in order to advance a coherent approach and to inform policy-making in digital regulation. Many of the decisions that must be made in the digital regulatory space involve trade-offs, for example balancing transparency and fair competition against protecting consumers' data privacy. The DRCF aims to unpack these types of overlapping issues and to articulate the trade-offs that must be made so that regulators can better advise policymakers.

The DRCF has conducted a cross-regulatory review of AI focusing initially on the benefits and harms of algorithms, and shared regulatory concerns. It has considered consumer use of Generative AI and has piloted the 'AI and Digital Hub' to offer coordinated advice from the four regulators to innovators on new products and services. Building on insights from the Hub, it is developing a 'one-stop' digital regulatory tool to streamline access to digital policy and regulation and has turned its attention to the regulatory challenges associated with the adoption of agentic AI.

APAC approach

In the APAC region, the situation is evolving quickly and countries are adopting a range of approaches, including: (1) a prescriptive regulatory framework similar to the EU's approach (eg, Mainland China and Australia); (2) high level guidance (eg, Singapore and Hong Kong SAR); or (3) combining regulation and guidance (eg, Japan and South Korea).
We focus on three key Asian jurisdictions in the journey to regulate AI: Mainland China, Singapore and Hong Kong SAR.

Mainland China approach

Acceleration in the development of AI law and regulation

While the launch of China's DeepSeek R1 model created ripples in the West which disrupted the markets and brought down the value of some of the major tech companies, in Asia it sparked significant attention for the AI sector. In China, the trade tensions with the United States have spurred a renewed drive towards advancing its strategy of achieving self-sufficiency in the technology sector, particularly in the AI space. Recently, President Xi Jinping advocated for "safe, reliable, and controllable" AI, and called for an acceleration in the development of AI laws and regulation.

Proliferation of AI policies and initiatives

Since launching the Next Generation AI Development Plan in 2017, China has introduced a wave of AI regulations and policies at both central and local levels to support AI R&D, industry, and applications. From 2019, the focus shifted from support to governance, marked by the release of AI Governance Principles followed by the Outline for Establishing a Rule-of-Law-Based Society (2020-2025). In 2024, a joint directive established a national AI standardisation system. Most recently, the State Council's "AI Plus" action plan (August 2025) aims to promote AI use across the economy, society and government.

Extensive AI legislative framework:

China has already implemented an extensive AI legislative framework, which includes various sector-specific AI rules for industries like finance, healthcare and automotive, addressing growing risks and ethical concerns. This includes:

- **Algorithm recommendations:** The Cyberspace Administration of China's (CAC) provisions on the algorithm recommendations management (2021) focus on internal ethics and governance mechanisms for tech services providers. They introduce a classification system based on algorithm properties (such as their content type, number of users and their impact on public opinion), and impose filing and assessment requirements for algorithms with recommendation functions which can 'shape public opinion' or 'mobilise society'. Read more: China's Algorithm Regulation: reshaping the Tech Sector (May 2022); China's first batch of algorithm filings revealed – What can we learn? (August 2022).
- **Deep synthesis:** In response to concerns that advances in AI technology could be exploited for illegal purposes or identity impersonation, the CAC issued deep synthesis provisions in November 2022. The provisions mainly address deepfakes, targeting AI applications used to generate text, video and audio, prohibiting generation of 'fake news' and requiring synthetically generated content to be labelled. It also creates an 'algorithm registry' to be used in future regulations. Read more: Regulating ChatGPT and the Metaverse for social good: How do China's first-ever deepfake rules affect AI governance (July 2023).

Generative AI measures

The world's first mandatory generative AI-specific measures were introduced in July 2023: interim measures on the management of generative AI services. These measures regulate providers of generative AI where there are 'public-facing' services, covering all aspects of generative AI, from how it is trained to how users interact with it. They also contain rules unique to China, such as that AI must observe socialist values and have extraterritorial effect. Read more: What y'all ChatGPTing about? China enters the debate on generative AI with new draft rules (April 2023); China's first generative AI regulation unveiled: Are there positive signals for the emerging technology under global scrutiny? (July 2023).

GenAI labelling

China's first AI generated content labelling measures took effect in March 2025 together with mandatory national standards on AI generated content labelling rules. With effect from 1 September 2025, these labelling rules set obligations for responsible parties across the entire value chain of AI-generated content (from generation to distribution to users) to apply explicit and implicit labelling on their AI-generated content. Read more: China: dual-track AIGC labelling and latest AI regulatory development (September 2025).

Status of proposed AI Law

Although a standalone Artificial Intelligence Law was under legislative review, the latest status of China's highly anticipated comprehensive AI Law (expected to adopt a risk-based approach similar to the EU Act) remains uncertain. More recent government legislative plans call for promoting legislation to support the healthy development of AI, indicating a legislative shift to a broader, more flexible approach focused on guiding principles and regulatory frameworks, rather than enacting a single, comprehensive AI law at this stage.

Increasing AI standards

China's national standardisation committee released its AI Safety Governance Framework in September 2024 and the first version of its AI Safety Standardisation System in January 2025 for public consultation. Following these higher level plans, national and industry standards covering security requirements for generative AI and AI technology have been proposed or implemented, such as service security requirements, data annotation security specifications and pre-training and fine-tuning training data security guidelines. Read more: China: New AI standards for what your ChatGPT-like tools can and cannot say (June 2024).

Strengthening ethical governance

Governance of AI in China places a strong emphasis on ethics. Most recently, draft AI Ethics Management Service Measures were released in August 2025 by 10 ministries, including China's National Financial Regulatory Administration. These measures propose to establish an ethical review mechanism targeting high-risk autonomous decision-making systems.

Global engagement

China also wants to lead on global standards for AI: most recently, China launched the Global AI Governance Action Plan in July 2025 – a 13-step roadmap for safe, inclusive AI (see Chapter 2).

AI misuse enforcement on the way

The CAC ran a three-month nationwide special campaign to regulate AI services in 2025. The campaign targets issues including non-compliant AI products, improper training data management, weak security measures, AI-generated rumours, fake information and violations of minors' rights. Otherwise, China saw its first enforcement actions in respect of AI output copyright infringement and AI-generated voice rights in 2024.

Singapore approach

Guidance-based, risk-focused

Singapore continues to adopt a risk-based approach to AI regulation, focusing on specific use cases and existing laws to address potential risks. While Singapore does not have specific statutory laws on the use of AI, it has promulgated numerous voluntary guidelines and initiatives to keep pace with AI developments. This approach ensures that AI-enabled technologies are regulated appropriately without stifling innovation.

Key developments

Singapore has made significant strides in its soft law approach to AI regulation over the past couple of years:

- **AI Verify Framework**: Singapore launched its voluntary AI governance testing framework, AI Verify, in 2022, expanding it to cover generative AI systems through Project Moonshot in 2024. This initiative aims to address the security and safety challenges of LLMs. The AI Verify Foundation, established in 2023 with members including IMDA, Microsoft, Google, AWS and IBM, supports the development and use of AI Verify through industry collaboration.
- **Model AI Governance Framework for Generative AI**: Following public feedback, Singapore published the Model AI Governance Framework for Generative AI in 2024 to provide guidelines for the responsible use of generative AI technologies.
- **Starter Kit for Safety Testing of LLM-Based Applications**: The Infocomm Media Development Authority (IMDA) published the draft Starter Kit for Safety Testing of LLM-Based Applications for public consultation in May 2025, which is the first-of-its-kind set of guidelines for safety testing of LLM-based applications and sets out practical steps for testing beyond high-level frameworks.
- **International alignment**: Singapore has worked to ensure that its AI governance frameworks are interoperable with international standards. For example, the AI Verify framework has been mapped to the USA's National Institute of Standards and Technology AI Risk Management Framework and ISO/IEC 42001:2023.
- **Global AI Assurance Pilot and Sandbox**: The IMDA and the AI Verify Foundation launched the Global AI Assurance Pilot in February 2025 for the testing of generative AI applications, with the aim of bringing together specialised testing companies and generative AI application builders or deployers. Following its success, the Pilot was developed into an ongoing Global AI Assurance Sandbox and launched in July 2025.
- **Agentic AI accelerator**: The Digital Industry Singapore partnered with Microsoft to launch a agentic AI focused accelerator in August 2025, aiming to create "Frontier Firms", organisations powered by hybrid teams of humans working with AI agents. The accelerator is part of the Enterprise Compute Initiative, established to support the implementation of AI across the economy and sits within the Government's broader National AI Strategy 2.0.

Hong Kong SAR approach

Guidance only

There are no specific statutory laws on the use of AI and regulations mainly rely on guidelines issued by government bodies and more specifically by the data protection regulator (see Chapter 4), Securities and Futures Commission and banking and payments regulator to assist regulated entities to comply with general regulatory requirements which may apply to the use of AI (see Chapter 3).

Hong Kong government voluntary guidelines

The Digital Policy Office released the Ethical Artificial Intelligence Framework in 2024 which outlines ethical AI principles, a governance model for AI, a lifecycle guide for AI and an impact assessment template – providing voluntary guidance rather than mandatory rules. It then introduced the voluntary Generative Artificial Intelligence Technical and Application Guideline in 2025, aiming to offer a best practice framework for the ethical and responsible development, deployment and use of generative AI technologies.

Government policy

Executive Order 14179 Removing Barriers to American Leadership in Artificial Intelligence

Returning President Trump held true to his promise to rescind former President Biden's landmark AI Executive Order and replace it with his own Executive Order on AI in January 2025, marking a clear deregulatory shift focussed on encouraging economic competitiveness, national security and AI innovation.

AI Action plan

President Trump released his AI Action Plan to sustain and enhance America's AI dominance in July 2025. "Winning the AI Race: America's AI Action Plan" is a comprehensive effort aimed at solidifying United States leadership in AI and on focussing regulatory efforts on harmonisation across existing state laws (see Executive Order 14319 on Preventing Woke AI in the Federal Government and Executive Order 14318 on Accelerating Federal Permitting of Data Center Infrastructure). Read more: White House's AI Action Plan: Winning the Race in a Patchwork Regulatory Era (July 2025).

NIST standards for trustworthy AI

The National Institute for Standards and Technology released its AI Risk Management Framework in 2023 (see Chapter 6) to help public and private sector companies assess and manage AI-related risks. Like many NIST standards, this non-binding framework provides voluntary guidelines and recommendations but is increasingly recognised as an industry standard. NIST is also researching managing AI biases and its National Cybersecurity Centre of excellence is looking into mitigation of AI bias in credit underwriting.

NIST expands AI Security Framework

NIST also advanced its efforts to create a Cybersecurity Framework Profile for AI in August 2025 (see Guidelines to manage misuse risk from advanced AI (draft) and Taxonomy of AI attacks and mitigations) by proposing to develop a series of control overlays for securing AI Systems. Using SP 800-53 controls, the overlays will offer implementation-focussed guidelines for security controls tailored to different AI types, components and use cases – aiming to protect confidentiality, integrity and availability of information.

FTC enforcement approach

In its AI Compliance Plan (2024), the FTC outlined its strategy for responsible AI use, emphasising transparency, accountability and public benefit. However, with the change in Presidential administration and to the composition of the FTC in 2025, the shorthanded commission appears less likely to expansively leverage Section 5 of the FTC Act (which prohibits unfair and deceptive acts and practices in commerce generally) to pursue enforcement actions related to how companies deploy AI. This reflects a broader deregulatory trend under the current administration.

US Copyright Office

The US Copyright Office's three-part AI policy report analyses copyright elements of digital replicas, copyrightability and generative AI training. Part 1 (July 2024) discussed whether current legal frameworks adequately protect against the unauthorised distribution of digital replicas (and concluded that they do not). Part 2 (February 2025) reaffirms that copyright protection requires human authorship, meaning that AI-generated materials, content and prompts cannot be copyrighted (even if there is human effort, time, or cost involved). When AI-generated content is part of a larger, human-authored work, the overall work's copyrightability remains intact, but the AI-generated elements themselves may not be protected. Part 3 (pre-publication version, May 2025) concludes that training AI models on copyrighted works may raise legal concerns, but leaves the question of infringement to courts and Congress, calling for greater transparency and legal clarity.

Federal legislation

No comprehensive federal legislation to date

While there is general consensus among US lawmakers and industry on the need for comprehensive federal AI legislation diverging opinions on the scope of its restrictions and applicability have stalled progress. Congress recently rejected a proposed 10-year moratorium on new state AI regulations, and states continue to legislate independently (see below). To date, only sector-specific bills covering areas like financial services, healthcare and transportation have emerged, tackling discrete issues such as privacy and discrimination. Bi-partisan proposals (such as the Unleashing AI Innovation in Financial Services Act and the Validation and Evaluation for Trustworthy AI Act) have failed to gain traction.

Take It Down Act

The Take It Down Act (S.146) (May 2025) addresses harms caused by AI-generated content. It requires platforms hosting user-generated content to remove non-consensual intimate visual depictions, including "deepfakes", establishes criminal penalties for violators who intentionally disclose these depictions, and mandates a process for victims to request such content's removal. The Act also criminalises making threats to publish such content.

State legislation

Colorado AI Act

In the absence of federal regulation, US states have begun legislating on AI. Colorado's AI Act (currently due to be effective in June 2026) leads these state efforts targeting high-risk AI systems that are a substantial factor in making consequential decisions relating to areas such as education, employment, lending, government services, healthcare, housing, insurance and legal services. It will require developers and deployers of high-risk AI systems to prevent algorithmic discrimination, mitigate risk and function transparently.

Patchwork of state law

In 2025, over 550 AI-related bills were introduced across 45 states and Puerto Rico with at least eight states to date proposing comprehensive bills similar to Colorado's. California has been especially active in its piecemeal legislative approach to AI. This growing patchwork of state-led AI regulation, together with applicable privacy law (see Chapter 4) has created compliance challenges for companies operating nationally.

Facial recognition

Some states have taken steps to regulate AI use through AI-specific laws within their jurisdictions, although they are limited in scope, eg, facial recognition regulation in California and Washington:

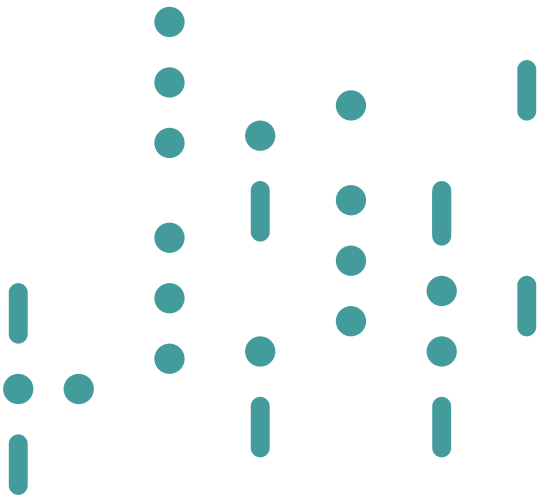
- Colorado's AI Act (see above) further regulates facial recognition to the extent that it serves as a part of a high-risk AI system that makes or is a substantial factor in making consequential decisions about Colorado residents.
- Connecticut, Delaware, Florida, Indiana, Iowa, Montana, New Jersey, Oregon, Tennessee, Utah and Virginia have each enacted comprehensive privacy laws that include biometric data and/or facial recognition as a protected category. Other states and local cities or counties have established limits on law enforcement usage of facial recognition technologies (see, eg, Alabama, Boston, Baltimore and Vermont).

Opting out of automated decision-making

- Out of the 19 current comprehensive state privacy laws, 17 provide a consumer right to opt out of fully automated decision-making or 'profiling' based on automated decision-making,which is often made by technologies incorporating AI. See the discussion below of recent CCPA rulemaking activity.

Employment use cases

- The most progressive regulations about automated decision-making and machine learning apply to hiring, promotion and termination of employment decisions.
- Illinois's Artificial Intelligence Video Interview Act (2020) regulates the use of AI to analyse video interviews of job candidates.
- New York City's Local Law 144 (2023) requires employers using automated employment decision tools in hiring and promotions to notify candidates and employees accordingly and to conduct annual bias audits, which must be publicly available.
- Maryland's House Bill 1202 (2020) prohibits employers from using facial recognition technologies during pre-employment interviews unless the candidate consents in writing.
- California's employment regulations (amended June 2025) now require employers to retain automated decision system data for four years to demonstrate anti-discrimination law compliance. They may also need to provide reasonable accommodation to applicants disadvantaged by automated decision systems, and evidence of anti-bias testing or similar proactive efforts may be used to help defend discrimination claims.



United Kingdom – Light touch, industry led
The UK government previously announced that it intends to adopt a light-touch and industry-led approach, meaning that there won't be specific legislation like the EU AI Act. However, in July 2024 government proposed a set of binding measures on AI and its intention to establish appropriate AI legislation.

European Union – Extensive AI regulatory regime
The EU AI Act was adopted by the EU Parliament in March 2024 and entered into force in August 2024. It is the first AI-specific regulation across the globe and has extensive regulatory and liability regime with extra-territorial reach. It will become fully applicable in two years with various provisions being implemented at different intervals.

AI specific regulation – snap shot of key jurisdictions

International – Efforts at coordination
The OECD has updated its principles to which 49 countries are committed, including the G20 and the EU– this addresses issues related to the emergence of general purpose and generative AI. There have been a variety of international co-ordination efforts including the G7 Hiroshima AI Process; US-EU Trade and Technology Council and global AI Safety Summits.

Voluntary, light-touch approach

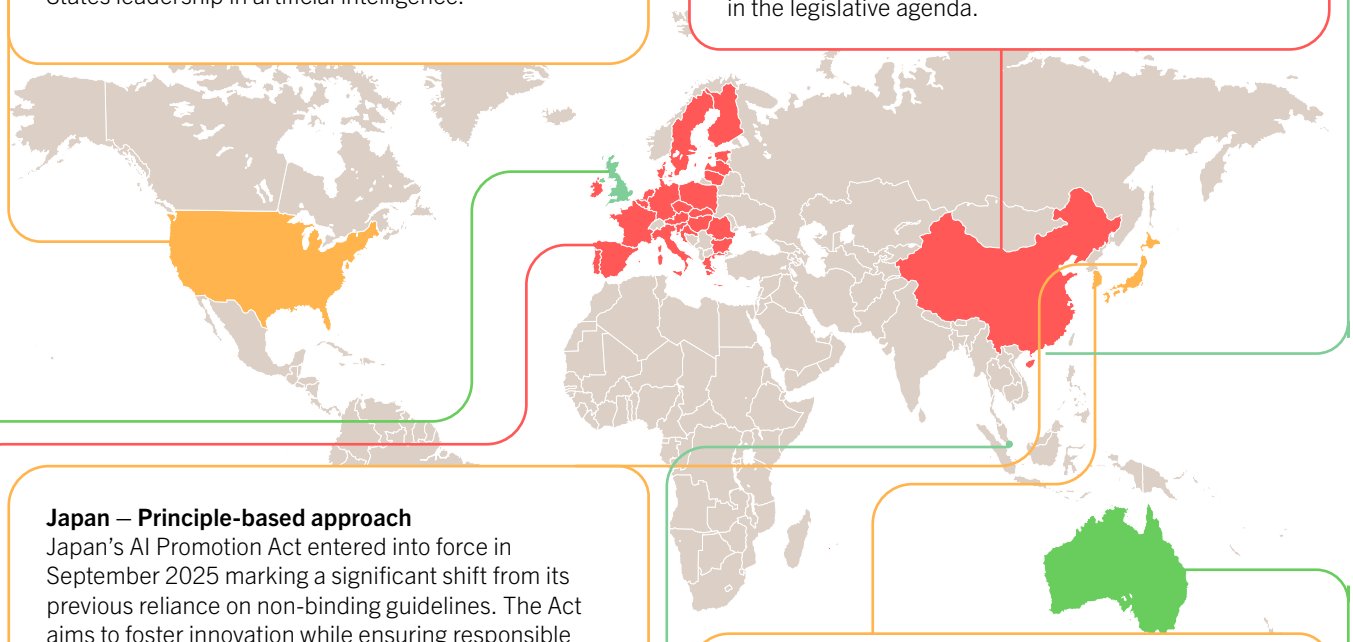
Incremental legislation

Prescriptive, risk-based regulations

Hong Kong SAR – Currently no AI specific laws
No statutory laws on the use of AI, although the HKMA and SFC have issued guidance on the use of AI in financial services. The Privacy Commissioner has also published the Artificial Intelligence: Model Personal Data Protection Framework.

United States – Growing body of AI guidance and law
The US is characterised by a complex regulatory landscape which is mainly state-led. There is various sector specific regulator guidance and various state level laws with respect to AI specific activities. The US has also produced a comprehensive AI Risk Management Framework. May and September 2024 saw AI legislation in Colorado and California. Acknowledging this uniquely complex regulatory landscape, the Trump administration published in July 2025 the AI Action Plan which is a comprehensive effort aimed to solidify United States leadership in artificial intelligence.

Mainland China – Prescriptive and risk-based
Rules effective as of August 2023 cover all aspects of generative AI. Several rules were published to regulate other AI-related service/products (e.g. Algorithm Recommendation and Deep Synthesis Services). A comprehensive official draft AI law has been listed in the legislative agenda.



Japan – Principle-based approach
Japan's AI Promotion Act entered into force in September 2025 marking a significant shift from its previous reliance on non-binding guidelines. The Act aims to foster innovation while ensuring responsible development and use of AI.

South Korea – Upcoming overarching AI legislation
South Korea's AI law is aimed at supporting innovation rather than imposing restrictions. The Act on Promotion of the AI Industry and Framework for Establishing Trustworthy AI was passed in January 2025 and will take effect on 22 January 2026.

Singapore – Mostly self-regulated
The general regulatory approach is to foster AI innovation through the responsible use of AI. The financial regulator has issued AI-friendly guidelines and best practice for regulated firms, while the data regulator has also issued guidance for the use of personal data in AI. In May 2024, the Singapore government published a comprehensive Model AI Governance Framework for GenAI.

Australia – There are currently no AI-specific laws or regulations. In 2024 the government published the Voluntary AI Safety Standard providing voluntary AI guardrails alongside technology-neutral regulations, including consumer protection, online safety, and privacy laws. There were proposals to make these guardrails mandatory for "high risk" AI, although the Productivity Commission has since cautioned against this.

Regulatory and legislative milestones in key jurisdictions

MC Mainland China | HK Hong Kong | SG Singapore | SK South Korea

2017

- **MC:** Next Generation AI Development Plan *Effective July 2017*

2018

- **SG:** MAS FEAT Principles *Published November 2018*

2019

- **MC:** Next Generation AI Governance Principles *Effective June 2019*
- **HK:** HKMA High level Principles and Guidance for Finance sector *Published November 2019*

2021

- **MC:** Next Generation AI Code of Ethics *Effective September 2021*

2022

- **MC:** Measures for Internet Information Recommenders *Effective March 2022*
- **US:** Blueprint for an AI Bill of Rights *Published October 2022*

2023

- **MC:** Measures for Internet Information Deep Synthesis *Effective January 2023*
- **MC:** Interim Measures Management of Gen AI *Effective August 2023*
- **SG:** Proposed Guidelines re personal data in AI systems *Consultation published July 2023*
- **UK:** AI White Paper on AI regulation *Published March 2023*
- **EU/US:** Developing voluntary AI code of conduct *Announced June 2023*
- **US:** NIST AI Management Framework *Published January 2023*
- **US:** Biden's AI Executive Order *Published October 2023*

2024

- **EU:** AI Act enters into force *August 2024*
- **UK:** Government Response To AI White Paper feedback *Published February 2024*
- **UK:** Digital regulators respond to government's AI White Paper *Published April 2024*
- **SG:** Model AI Governance Framework for GenAI *Published May 2024*
- **SK:** AI Basic Act *Published December 2024*
- **HK:** SFC Use of generative AI language models' *Published November 2024*
- **HK:** HKMA Consumer Protection in respect of Use of Generative Artificial Intelligence *Published August 2024*

2025

- **US:** Trump's Removing Barriers to AI Executive Order *Published January 2025*
- **EU:** GPAI Code of Practice *Published July 2025*
- **US:** White House's AI Action Plan *Published July 2025*

3 AI and financial services regulation

As important as it is to keep one eye on future developments, when implementing AI systems today, financial services firms must do so within the constraints of the existing financial services regulatory framework.

Existing regulatory considerations

- **General compliance**: Financial institutions must ensure that their approach to AI meets the regulatory requirements placed on them by sectoral regulation. When rolling out AI, firms will need to continue meeting their general obligations in relation to governance, systems and controls and risk management.
- **Oversight and validation**: A well-designed AI-based model can reduce regulatory risks, eg, by analysing huge volumes of transaction data to identify fraudulent payments. Even so, firms will need to ensure that they maintain appropriate oversight of the activities of the AI. For example, in the context of robo-advice, firms should ensure that they can validate the suitability of the advice provided by the robo-adviser in the same manner as they would for human advisers.
- **Outsourcing, supply chain and third-party liability**: Financial services firms cannot outsource responsibility for meeting their regulatory obligations. It is the responsibility of firms that use AI systems provided by a third party to ensure that their use of those systems is compliant. Firms should question whether their arrangements with third parties qualify as a regulated outsourcing. The outsourcing of critical or important functions may attract additional requirements. Some of these requirements are now also being extended to non-outsourcing arrangements. Increasingly, the expectation is that firms should manage third-party risks more holistically rather than focusing on arrangements with third parties that are classified as outsourcings for regulatory purposes.
- **Consumer protection**: Firms should carefully consider how their adoption of AI could impact customers. For example, the FCA's Consumer Duty requires UK firms to prove that they are acting fairly and delivering good outcomes for retail customers. Customer-facing AI, such as chatbots, would need to be consistent with this principle and the associated rules on how firms support their customers, especially vulnerable customers.
- **Algorithmic trading**: More specific regulation may also apply. For example, the rules on algorithmic trading and high-frequency trading in the EU Markets in Financial Instruments Directive aim to avoid the risks of rapid and significant market distortion. These rules would apply to AI tools that are intended to make high-frequency trading decisions.

Particular challenges presented by AI in financial services

The OECD asserted in a 2021 report that: "*Policy makers and regulators have a role in ensuring that the use of AI in finance is consistent with the regulatory aims of promoting financial stability, protecting financial consumers and promoting market integrity and competition*".

However, it is not always easy for financial institutions and their supervisors to apply novel concepts to established regulatory regimes. AI presents particular challenges in this regard. We expect regulators to focus on:

- **the resilience of firms** as they rely on AI to trawl massive data sets or communicate with customers;
- **the financial stability implications** of widespread AI adoption; and
- **higher-level ethical and related compliance questions**, including accountability for machine-made decisions and the transparency and explainability of machine-led decision-making processes.

We will explore each of these in turn below.

Resilience

What is operational resilience?

Building the resilience of the financial system to operational disruption has been a long-standing policy aim and is now a focus area in jurisdictions with major financial centres and at an international level. Some jurisdictions have introduced operational resilience regimes (eg, the EU, via the Digital Operational Resilience Act (DORA) and the UK, via an operational resilience framework), while others have issued additional guidance on how firms should strengthen their operational resilience (including Mainland China, Singapore, Hong Kong SAR and the US).

At an international level, the Basel Committee on Banking Supervision has also issued principles for operational resilience to help coordinate national approaches in this area. The Financial Stability Board has consulted on a toolkit for enhancing third-party risk management and oversight.

Read more: Operational Resilience | Guide

Learn more: AI in financial services: Applying operational resilience rules to AI | Webinar (September 2025)

How does AI threaten resilience?

There are three key threats:

- **Third-party failure**: In many cases, financial services firms will not develop their own AI systems but instead will work with technology companies and other third-party service providers. If any critical system relies on a third-party provider, failure of that provider is a key threat. From the regulators' point of view, this threat is amplified if there is concentration in the market around a small number of providers, especially if those providers are unregulated and so not subject to direct supervision. Some regulators (including in the EU and UK) are now able to directly supervise third parties that provide 'critical' services to the financial services sector.
- **Challenge in substituting systems**: AI systems today tend to operate as black boxes. One result of this is that it may not always be clear how the system operates and what dependencies it has. In the midst of a system failure, eg, a 'black swan' event, it may be very difficult to maintain business continuity by substituting systems if it is not clear how the AI system operates.
- **Big data**: AI relies on huge quantities of data. Put simply, more data processing means a greater risk of data breaches.

Operational resilience compliance

In addition to identifying critical operations, setting tolerance for disruption, mapping interdependencies and carrying out testing, firms will need to consider the impact of the use of AI on their operational resilience and whether this introduces any additional risks to consider. How firms use contracts to manage third-party technology risk, including supply chain risks, is going to be a key focus for regulators.

Financial stability

Several central banks and regulators have drawn a connection between AI, concentration risk and financial stability risk. Risks may arise if the same handful of "best in class" algorithms are used by a large number of systemically important financial institutions. In stress situations, for example, herding behaviours could lead to undesirable outcomes.

It remains to be seen whether macroprudential concerns will lead to AI-specific regulatory interventions (see, for example, the UK approach to financial services below). In the meantime, regulators will at least expect firms to ensure that their scenario testing programmes keep pace with their adoption of the latest technologies and the risks associated with them.

Ethical deployment of AI

It must not be assumed that AI can be programmed to act ethically in its own right. Any system that is complex enough to be considered 'intelligent' will likely also be complex to control. Applying an AI system to provide financial services can result in unpredictable consequences. In combination with other AIs, very complex behaviours could develop.

Multiple algorithms interacting and competing with one another can result in undesirable outcomes. In principle, AI should 'respect' human autonomy and human rights, and should abide by basic ethical concepts such as prevention of harm, fairness and accountability, as well as avoiding biases and protecting vulnerable groups (including children and people with disabilities).

How do ethics apply to AI?

Taking the EU ethical guidelines as a paradigm, there seem to be many limbs to the meaning of 'ethical' AI:

- **Human agency and oversight**: AI systems should enable equitable societies by supporting human agency and fundamental rights, and not decrease, limit or misguide human autonomy.
- **Robustness and safety**: "Trustworthy AI" requires algorithms to be secure, reliable and robust enough to deal with errors or inconsistencies during all life-cycle phases of AI systems.
- **Privacy and data governance**: Citizens should have full control over their own data, while data concerning them will not be used to harm or discriminate against them.
- **Diversity, non-discrimination and fairness**: AI systems should consider the whole range of human abilities, skills and requirements, and should ensure accessibility.
- **Societal and environmental well-being**: AI systems should be used to enhance positive social change and to promote sustainability and ecological responsibility.
- **Accountability**: Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes.
- **Transparency**: The traceability of AI systems should be ensured.

Many similar concepts are seen in China's principles on developing responsible AI. These principles set out broad requirements of wellbeing, fairness, accessibility, reliability, safety, universality and governability with which all parties involved in AI development should comply. A consistent through-line can be found in the US Federal Trade Commission's Artificial Intelligence Plan and in the Hong Kong Monetary Authority's high level principles on AI.

We will focus on accountability and transparency, which have particular relevance in financial services and have been receiving attention from financial regulators.

Accountability

Governance and accountability

With the industry looking to scale its application of AI and machine learning technologies rapidly, many regulators are focusing on board-level engagement and strong governance principles that will enable regulated firms to deal with challenges posed by these new technologies.

One idea that may gain further traction in the market is for firms to appoint a dedicated AI officer. After a failed attempt in 2023, UK lawmakers are again discussing requiring financial services firms to designate an individual to supervise the use of AI in their organisations. As a private member's bill, the idea is not likely to become law but may be picked up elsewhere, either as a formal requirement or recommended best practice.

Data and controls

As described in more detail in Chapter 4, data that is incomplete, inaccurate or mislabelled (or that embeds bias) is likely to generate problematic outputs (for example, poor or biased credit decisions).

Since AI poses challenges to the proper use of data, boards should attach real priority to the governance of data. This will include considering what data should be used, how it should be modelled and tested, and whether the outcomes derived from the data are correct.

Humans to remain accountable

UK regulators and the US Federal Trade Commission have both clarified that the adoption of systems centred on AI or machine learning technologies will not reduce the existing accountability burden on humans. They stand ready to challenge firms' existing approach to allocating accountability where necessary. Regulators question whether responsibility will be shifted both towards the board but potentially also to more junior, technical staff, which in the long run may mean less responsibility for front-office middle management.

This will bring a significant shift to how accountability for regulated firms has worked so far, which has been traditionally applied to senior individuals rather than employees in operational functions. The HKMA, for example, has been clear that boards and senior management will remain accountable for the outcome of AI applications; however, they also recommend clearly defining the roles and responsibilities of the three lines of defence in developing and monitoring the AI, underlining the importance of the role of humans at all levels. Boards are encouraged to continue to focus on the oversight of human incentives and accountabilities within AI and machine learning-centric systems.

Execution risk at board level

As the rate of adoption of AI in financial services accelerates, boards have to deal with the increased potential for execution risk. So far, firms have embraced either a piecemeal approach or a more general firm-wide approach to adoption. Regulators acknowledge the costs of aligning internal processes, systems and controls and underline the need for firms to make sure that there are senior managers with the appropriate skillset to deal with these new technological and legal challenges.

Boards should reflect on the skills and controls that are necessary to oversee the transition. Many of the challenges raised by this transition can only be brought together at, or near, the top of the organisation.

Systems and policies

In addition, regulated firms are obliged to have adequate systems and controls to deal with operational and other risks, as well as clear and documented policies for business continuity and contingency planning.

For example, the People's Bank of China has published a framework for real-time risk monitoring of fintech firms' business systems, application program interfaces, software development toolkits and apps. The framework requires the use of a combination of institutional reporting, information capture, automated testing and investigation, manual verification, information sharing, public monitoring and complaints procedures.

A clear governance policy taking into account all the chain of individuals making decisions in relation to the training and usage of algorithms seems the most prudent approach to current regulatory expectations.

The Senior Managers Regime and decision-making

The UK Senior Managers Regime is intended to enhance individual accountability within the financial services industry. Under the regime, senior managers must take reasonable steps to avoid a breach in the part of the business for which they are responsible. Senior managers will therefore take a particular interest in how AI is deployed within the scope of their responsibility.

Likewise, the Singapore regime under the MAS Guidelines on Individual Accountability and Conduct (IAC) and Hong Kong's SFC Manager in Charge regime, which applies to regulated financial institutions, are intended to promote senior managers' individual accountability, strengthen oversight over material risk personnel and reinforce standards of proper conduct among all employees. In particular, each senior manager's areas of responsibility must be clearly specified to ensure that senior managers are held to account for matters under their purview.

A significant hurdle for senior managers is likely to be transparency in AI systems (as described in more detail below).

Both the UK Senior Managers Regime and Singapore IAC Regime are likely to be used as a tool for ensuring that firms take responsibility for assessing AI-related risks and allocating that responsibility appropriately within the organisation. Firms implementing AI systems need to consider who is ultimately responsible for those systems, both operationally and in terms of their output.

72% of firms using or planning to use AI stated that they allocate accountability for AI use cases and their outputs to executive leadership in 2024. This is followed by developers and data science teams (64%), and business area users (57%)."

Source: Artificial intelligence in UK financial services – 2024 | Bank of England

Transparency

How transparent do you have to be to your customers?

Transparency with customers is an important pillar of responsible AI adoption. Firms should consider their obligations carefully. For example, in the UK, one of the FCA's high-level Principles requires firms to pay due regard to the information needs of their clients. The EU AI Act sets transparency obligations for providers and deployers of some AI systems. The US FTC has also warned that companies could face enforcement action if they mislead consumers about their use of AI (although note the shift in the FTC's enforcement approach discussed in Chapter 2).

More generally, greater transparency can help demonstrate trustworthiness, which is relevant not only to the public acceptance of the underlying technology but also for the firm which is using it.

Firms should also think about how they communicate about their use of AI beyond their customer base. Internally, key stakeholders will include senior management and risk managers, as well as representatives in other control functions. Externally, key stakeholders include auditors and regulators (including data regulators).

The 'explainability problem'

Machine learning is not always amenable to a meaningful explanation, as explanations are not a natural by-product of complex AI algorithms. For example, an AI model used to predict mortgage defaults may consist of hundreds of large decision trees deployed in parallel, making it difficult to summarise how the model works intuitively.

A high proportion of firms using AI report employing explainability methods. These can, for example, explain how much each input variable contributes to a model's predictions.

Even if systems allow for some degree of explainability, there is no consensus on at what level of detail the decision-making process should be explained.

Sector-specific considerations

Different types of business need to manage different types of risk. For example, participants in the data-rich and technology-reliant payments industry must pay particular attention to how they maintain the operational resilience of their core services as they seek to take advantage of AI's opportunities. This includes having robust governance processes in place and putting AI at the heart of their implementation of DORA or similar operational resilience regimes.

Read more: Unlocking the future of payments: The role of AI in transforming transactions (June 2024)

There are also broader legal risks associated with the development of agentic payments and the execution of payment transactions by AI agents.

Read more: Agentic payments: what are they, what are the legal risks and what's next? (August 2025)

Approach to AI in financial services in key jurisdictions

EU approach to AI in financial services

EU guidance

The European Supervisory Authorities have started developing regulatory and supervisory guidance on the use of AI in finance. For example, ESMA has published supervisory expectations on the use of AI in providing investment services to retail clients under MiFID and is preparing to report on how EU firms use AI in the securities markets following a survey in 2025. EIOPA has released an opinion clarifying how insurance legislation applies to the AI governance and risk management. Common themes are explainability and transparency, data quality, governance and human oversight.

Member State national guidance

Several EU Member States including France, Germany, Luxembourg and the Netherlands have issued discussion/consultation papers, reports and guidance on the use of AI in financial services.

Parliament report on application of financial regulation to AI

The Committee on Economic and Monetary Affairs within the European Parliament published a draft report on the impact of AI on the financial sector in May 2025. Read more: European Parliament drafts demand for Commission to spell out how financial regulation applies to AI (May 2025).

More guidance to come

As well as applying financial regulation to AI, EU financial services firms must also apply the AI Act. In February 2025 the Commission is due to issue more guidance on the AI Act including high-risk use cases (see Chapter 2). This will be supported by findings from a 2024 consultation which sought to gather evidence from the financial services sector on the practical impact of AI and AI regulation, in particular for high-risk use cases. Read more: EU institutions explore AI in financial services (July 2024).

UK approach to AI in financial services

Growth agenda

The current government is continuing to take a principles-based, rather than a rules-driven, approach to AI regulation in support of its wider growth agenda. The FCA and PRA have been given objectives to support the growth and international competitiveness of the UK economy. With this in mind, the FCA has identified that AI can *"make a significant contribution to economic growth, capital market efficiencies, [and] improve consumer outcomes"* and has launched an AI Lab to encourage responsible AI innovation. The FCA published a webpage titled "AI and the FCA: our approach", in September 2025, aimed at supporting the safe and responsible adoption of AI in UK financial markets.

Tech-neutral

UK financial services regulators generally take a 'technology-neutral' approach to rule-making, meaning that using AI should not change the rules with which firms must comply. That said, existing requirements warrant careful consideration in light of the novel features of AI.

Financial stability

If the Bank of England changes its technology-agnostic approach, it will likely be in response to financial stability concerns. Its Financial Policy Committee has identified four key channels of concern: the risk of wide-scale model failures when firms rely on common algorithms or data libraries; potential market disruption if AI-driven trading strategies behave unpredictably under stress; heightened operational dependence on third-party AI vendors; and novel cyber security threats arising from adversarial use of AI.

Consumer duty

The FCA has told firms to ensure that AI-driven customer interactions deliver fair, transparent outcomes, by embedding robust governance, monitoring algorithms for bias and drift and providing clear explanations of automated decisions to uphold the cross-cutting rules of good faith, avoiding foreseeable harm and supporting customers' financial objectives. Firms deploying AI in both a consumer-facing context, or in a manufacturing chain where consumers may ultimately be involved, need to align with the FCA's Consumer Duty. This should be factored in at every stage of the product/intervention lifecycle, from design and deployment through to monitoring and evaluation.

Governance

The FCA and PRA have indicated that the Senior Managers & Certification Regime offers the primary accountability mechanism for AI governance. Firms must integrate AI-specific risk management into existing frameworks by certifying staff in high-harm AI roles, maintaining robust challenge and escalation processes, and ensuring that algorithms remain explainable and free from bias or drift. Existing governance fora should ensure that they are integrating AI risk and oversight into their existing conversations about risk management within their remit and treat it as any other emerging risk or business change programme.

Supervision and enforcement

By embedding machine learning and natural language processing into their investigative toolkits, the FCA and PRA are strengthening their ability to detect and pursue misconduct (eg, the FCA's Market Abuse Surveillance TechSprint in 2024 showed how LLMs can filter false positives, anomaly-detection algorithms can flag unusual trading behaviour and pattern-recognition tools can pinpoint subtle price manipulations). The PRA and Bank of England have likewise brought advanced analytics into their frontline supervision. Regulators are more likely than ever to request large and granular datasets and may engage more proactively with firms based on insights from AI-driven market monitoring. Firms can get on the front foot by running AI tools over these datasets themselves.

Mainland China approach to AI in financial services

Guidance for data security of financial institutions

- China's National Financial Regulatory Administration's Banking and Insurance Institutions Data Security Management Measures (2024) require financial institutions to protect the interests of individuals, and to ensure transparency, fairness and impartiality (such as by establishing an assessment mechanism for on-boarding of algorithms and AI-related products, in each case when designing algorithms, carrying out automated decision-making activities, labelling data or training AI models).
- The People's Bank of China (PBC)'s Administrative Measures on Data Security in the People's Bank of China Business Areas (2025) strengthen data security management for relevant financial institutions with requirements such as ensuring that training data is accurate and diverse, and requiring an ethical review of AI model evaluation incentive rules. In-scope entities must also establish robust risk assessment and control systems over business data processing algorithms, addressing risks such as lack of explainability and vulnerability in algorithms.

Fintech specific approach

China takes a balanced approach to fintech regulation, promoting growth while enhancing oversight. The PBC's Fintech Development Plan (2022-2025), and multi-regulator digital finance development action plan aim to align digital finance with the growth of the digital economy. A 'look-through' regulatory approach aims to integrate fintech into the wider financial regulatory regime, increasing focus on network security, data protection, antitrust and consumer protection.

AI and technology ethics in finance

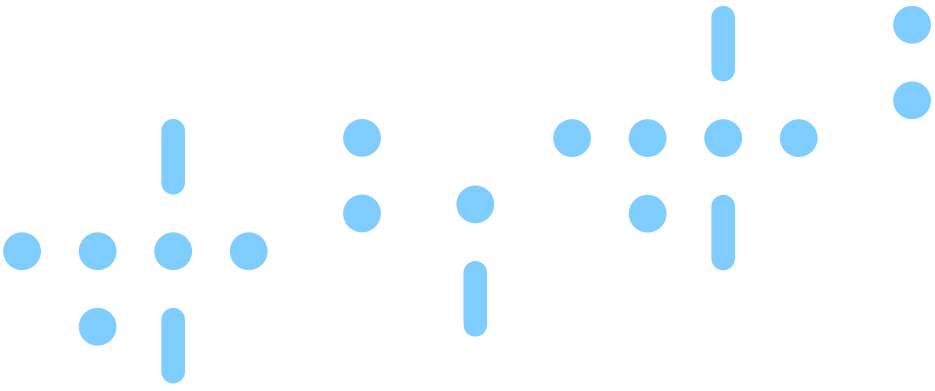
China has promulgated multiple sets of fintech and AI ethics guidance, underlining the general principles, responsibilities of AI developers and scrutiny requirements of AI and other fintech. The PBC's guidelines on technology ethics in the financial industry (2022) set out specific ethical standards for fintech development, including compliance with financial licensing requirements, data security, transparency, risk control, fair competition, green and low-carbon principles.

AI and operational resilience

The PBC's financial standard on AI algorithms (2021) applies to financial institutions, algorithm providers and third-party security assessors. It sets out detailed evaluation criteria on security, interpretability and accuracy of AI algorithms, aiming to eliminate the 'black box' effect and promote AI system continuity. Sector-specific rules also address resilience, for example the "Super Guidance" for asset managers (2018) requires disclosure to investors of AI models' key parameters, the general logic for asset allocation, the AI algorithms' defects and risks, alongside specific regulatory intervention rights in the event of a malfunction undermining financial market stability.

The Personal Financial Information Protection Technical Specification

This PBC-recommended industry standard (2020) sets out requirements for financial institutions to regularly assess the safety of external automated tools (such as algorithm models and software development kits) used in the sharing, transferring or entrusting of personal financial information.



Hong Kong SAR approach to AI in financial services

Focus on financial services

Given the uptick in use of AI by banks which presents new risk management challenges in banking, the Hong Kong banking and payments regulator, the Hong Kong Monetary Authority (HKMA) and the Securities and Futures Commission (SFC) have developed generative AI-specific guidance which firms should take into consideration when implementing both client-facing and internal processes which use generative AI.

HKMA guidance

The HKMA circular on generative AI guidance issued in 2024 builds on previous HKMA guidance from 2019 on (1) High level Principles on Artificial Intelligence and (2) Guidance on Consumer Protection in respect of Use of Big Data Analytics and Artificial Intelligence by Authorised Institutions. The HKMA approach to generative AI is to treat it as a subset of AI, and therefore the regulator provides additional generative AI-specific principles to ensure that appropriate safeguards for consumer protection are in place when generative AI is adopted for customer-facing applications.

SFC comments

The SFC's circular on the use of generative AI language models is its first piece of AI-specific guidance, highlighting both the benefits and risks (such as inaccuracies, data privacy threats and over-reliance). It calls for strong governance, robust validation and cyber security controls – and sets out examples of use cases which are high-risk – such as generative AI for investment recommendations, client investment advice or client investment research) which trigger regulatory notification requirements. Read more: Key implications of Hong Kong's new SFC circular on GenAI Language Models (November 2024).

HKIMR report highlights responsible generative AI adoption in financial services

The Hong Kong Institute for Monetary and Financial Research released a report in April 2025,examining generative AI's impact on financial services and regulation. This reveals that 75% of local financial institutions have adopted or are testing generative AI use cases, a figure expected to rise to 87% within five years. Regulatory intervention has contributed to this growth, with the HKMA supporting the use of AI by banks in areas such as AML monitoring and setting up a sandbox to trial generative AI banking solutions.

HKMA GenAI Sandbox

The Hong Kong Monetary Authority Generative AI Sandbox, launched in August 2024, has had strong industry engagement and entered its second cohort in April 2025. This provides a risk-controlled environment for regulated banks to develop and test innovative AI-based solutions.

Singapore approach to AI in financial services

Guidance and initiatives

The Monetary Authority of Singapore (MAS) has published a number of guidance documents regarding the use of AI in financial services, largely focusing on the risks which AI and generative AI present. MAS has also engaged in various industry and cross-jurisdiction initiatives relating to AI. Going forward, MAS intends to continue to work with industry to facilitate AI and generative AI governance and risk management efforts and is considering publishing further supervisory guidance in this space.

AI model risk management information paper

This MAS information paper (2024) sets out good practices for AI and Generative AI model risk management for all types of financial institutions, following its thematic review on this topic. MAS emphasised the need for financial institutions to use AI responsibly, given the risks that AI poses. The information paper focuses on: (1) oversight and governance of AI; (2) key risk management systems and processes; (3) development and deployment of AI; (4) third-party AI; and (5) generative AI.

Existing financial services regulations

MAS has also emphasised in its AI model risk management information paper that financial market participants should continue to meet existing regulatory requirements and supervisory expectations (eg, in relation to technology and cyber risk management, outsourcing risk management, etc.) when developing and deploying AI and data analytics in their financial products/services.

Generative AI cyber risks information paper

Another MAS information paper (2024) aimed at raising awareness among financial institutions of the key cyber threats arising from generative AI, the risk implications and some of the mitigation measures that financial institutions could take to address those risks (eg, its use by threat actors in deepfakes, phishing and malware generation and enhancement).

FEAT principles

MAS co-created the Fairness, Ethics, Accountability and Transparency principles (the FEAT principles) (2018) with the financial industry to promote the responsible use of AI in financial services. To support their implementation, MAS has co-created and continues to work on the Veritas Initiative, which aims to enable financial institutions to evaluate their AI-driven solutions against the FEAT Principles.

Project MindForge

MAS also continues to work on Project MindForge (supported by a consortium of banks) to develop an AI governance for the financial industry. As part of phase one, the consortium developed a comprehensive generative AI risk framework and the next phase will include insurance and asset management industries.

UK and Singapore collaboration

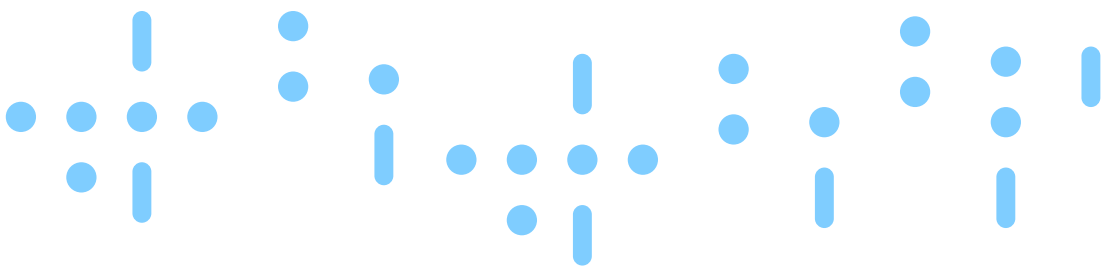
The UK and Singapore held their 10th Financial Dialogue in London in July 2025. The FCA and MAS discussed the use of AI in financial services and their respective approaches to AI regulation, agreeing to develop joint collaboration on AI by sharing innovative AI solutions and discussing cross-border AI developments.

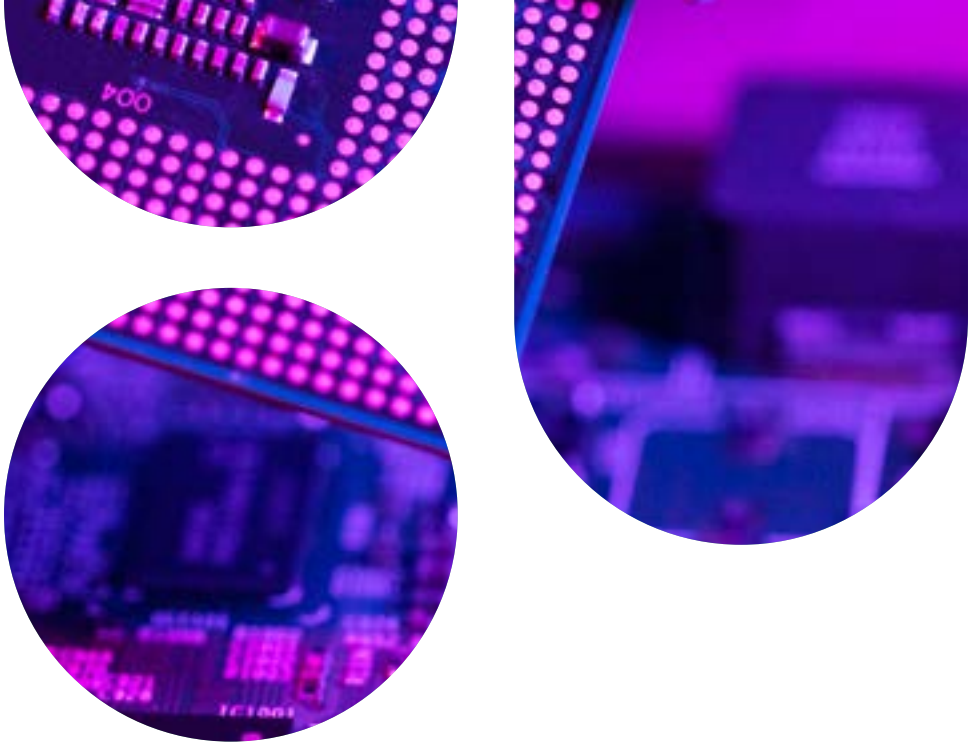
MAS Pathfinder Programme (PathFin.ai)

The MAS launched an initiative with the financial industry in August 2025 to foster knowledge exchange in AI solutions through a curated library of proven AI solutions and implementation best practices among financial institutions.

Addressing cyber risks from deepfakes

MAS highlighted emerging cyber threats in the financial sector linked to deepfake technology in its information paper on Cyber Risks Associated with Deepfakes (September 2025). This guidance builds on MAS's ongoing focus on digital resilience and provides recommendations to FIs for how to combat deepfake threats. The guidance standards are set high, signalling that this is a high priority topic for MAS and important for Singapore financial institutions to apply.





US approach to AI in financial services

Federal approach

The House Bipartisan Task Force Report on AI (2024) recommended a principles-based approach to AI regulation in financial services, focusing on existing regulatory frameworks. The Trump administration's AI Action Plan has gone on to embed a broad deregulatory agenda, discouraging federal regulators from imposing new AI-specific rules. The GAO Report on AI in Financial Services (May 2025) urges a proactive and adaptable regulatory approach, especially as AI adoption accelerates – encouraging regulators to refine guidance to address AI risk.

Accelerating federal use of AI

As part of Executive Order 14179 Removing Barriers to American Leadership in AI (see Chapter 2), federal agencies have been directed to accelerate federal use of use of AI through innovation, governance and public trust. Financial regulators are responding by appointing Chief AI Officers and establishing AI governing boards. The broad focus is on minimum safety standards for high-impact AI, such as deep fakes, and AI washing.

- **Federal Reserve:** The Fed regulates the use of AI primarily through existing frameworks such as model risk management, requiring banks to validate and monitor AI models for accuracy, fairness, and compliance with laws which prohibit discriminatory practices in lending. It has launched its own AI Program led by Chief AI Officer Anderson Monken.
- **Office of the Comptroller of the Currency:** The OCC takes a risk-based, innovation-friendly approach, applying existing model governance principles to AI, promoting financial inclusion through alternative data, and supporting responsible innovation while cautioning against systemic risks. It encourages banks to address challenges such as explainability, bias, and cybersecurity threats while leveraging AI for operational efficiency and fraud detection.
- **Securities and Exchange Commission:** The SEC regulates securities issuers' use of AI through disclosure rules, conflict-of-interest proposals and enforcement sweeps, aiming to ensure transparency, investor protection and responsible AI deployment. The SEC's AI Task Force, led by Chief AI Officer Valerie Szczepanik, aims to coordinate and accelerate the responsible use of AI across the agency. It has hosted public roundtables and issued compliance plans that address the risks and opportunities of AI in financial markets.
- **Commodity Futures Trading Commission:** The CFTC has issued AI regulatory guidance reminding registrants that AI must be used in compliance with existing rules under the Commodity Exchange Act. It has appointed Chief AI Officer Dr Ted Kaouk to lead AI strategy, is engaging with the public through comment requests, is promoting responsible AI principles via its Technology Advisory Committee and is actively engaged in efforts towards international regulatory alignment.
- **Treasury Department:** The Treasury Report on AI in Financial Services (2024) recommended enhanced risk management, transparency and interagency collaboration to address AI risks in financial services. Within the Treasury the Financial Crimes Enforcement Network (FinCEN) aims to accelerate AI use via public-private partnerships, encouraging innovation within financial institutions, and developing FinCEN's advanced analytical tools. It also has issued alerts to combat the criminal misuse of AI and deepfakes.

4 Ensuring AI is compliant with global data protection laws

The interaction between AI and data protection legislation is complex and still not fully resolved. In the EU, regulators are considering fundamental issues such as when personal data taken from the internet can be used to train an AI. They also expect those using AI to apply high governance standards, including completing detailed impact assessments.

The 'law of everything'

The broad scope of the EU's General Data Protection Regulation (GDPR), and its flexible and technology-neutral rules, mean that the GDPR is often described as the 'law of everything'. Prior to the application of the AI Act, the GDPR has taken a central role in the regulation of emerging technology such as AI, with regulators forcefully applying the rules in the GDPR to new AI models and services.

Whether looking at the EU or more widely, data protection is also an area in which black letter compliance with the law is not enough and failure to adequately protect personal data can be reputationally damaging. Firms also need to consider the expectations of customers and employees: if an AI project crosses the 'creepy line', it is unlikely to succeed.

Implications for financial services

Data protection regulators will be interested in the use of AI in financial services just as much as any other sector, particularly where there is the potential for consumer harm. For example, the use of AI to assess eligibility for financial products or the size of insurance premiums, or facial recognition as a means to conduct remote authentication, all fall squarely within the scope of the GDPR.

This is an area in which there is a degree of overlap with the EU AI Act. The sorts of activities identified as high-risk in the EU AI Act are also issues that data protection regulations are likely to focus on.[1] Data protection regulators will expect these activities to be carried out in a transparent manner, using an accurate and non-discriminatory algorithm that reflects the controls on robo-decisions (discussed below).

Financial regulators are likely to take a similar interest in innovative uses of AI. For example, the UK's FCA is undertaking a 'transformation programme' and expects to be a data regulator as much as a financial one by 2026. To achieve this, it is putting more resources into its data and technological capabilities.

Data as the 'new oil'

Underpinning many advances in AI is data, which is necessary to train the relevant AI model. Importantly, that data should be high-quality, well-formatted and properly representative of the real-world situations in which the AI will be used. It should also be checked carefully to ensure that it does not embed biases and discrimination. There must also be a proper 'legal basis' to use that data to train the AI model, something regulators are scrutinising closely in relation to generative AI (see below).

AI also opens the door to greater use of 'non-traditional data', such as social media posts, and there are already examples of financial services companies trying to use that information to help price premiums or profile customers. In China, much of Ant Group's record valuation of $315 billion in 2020 was attributed to its powerful use of data from its super apps and affiliated e-commerce business to assist it in credit scoring small loans applicants.

However, in Western markets at least, the use of this type of data should be approached with caution as such projects can easily slip over the 'creepy line' (eg, Facebook stymies Admiral's plans to use social media data to price insurance premiums).

The use of social media posts and other non-traditional data may be unacceptable to customers and could also have a chilling effect on freedom of speech if speaking out on a controversial subject on social media negatively impacts a customer's financial status. Conversely, it also creates the risk of financial exclusion for customers who do not wish to have an online presence.

1 Strictly speaking, the facial recognition as a means of remote authentication is unlikely to be considered high-risk under the EU AI Act. While the list of high-risk AI systems in Annex III refers to "remote biometric identification systems", it expressly carves out biometric verification solely used for authentication.

Outline of GDPR obligations

Where data relates to information about living individuals, it will be subject to the GDPR. This provides a comprehensive framework covering all stages of the AI development process, including collecting data, using data for training and testing, and final deployment.

The rules in the GDPR are complex, but the underlying aims are simple, and compliance helps to ensure that individuals will trust you with their personal data in the context of AI applications.

The core building blocks with which to build that trust are summarised briefly below, though it is important to note that there are still outstanding questions about the compatibility of generative AI with the GDPR (see box below).

- **Be transparent**: Tell individuals what personal data you hold about them and how you are using it, including that you are using artificial intelligence technology on their personal data. This is backed by detailed disclosure obligations in the GDPR.
- **Don't be creepy**: Ensure that you have a proper justification for using personal data and ensure that data is accurate, not excessive and not kept for longer than necessary.
- **Empower individuals**: Give individuals meaningful choices and control over how their personal data is used. This is backed up by the rights in the GDPR, including those on the prohibition of automated decision-making.
- **Keep personal data secure**: Data security is an important issue for artificial intelligence algorithms, which often use large amounts of personal data.
- **Be fair and avoid discrimination**: Consider the impact of the artificial intelligence model on different groups of individuals and ensure that any difference in their treatment is justified. Do not use sensitive and inherently discriminatory data types, such as information on racial or ethnic origin, unless there is a very good reason to do so.
- **Document your decisions**: Keep records of the decisions you have made. In particular, you are likely to need to complete a data protection impact assessment.

Robo-decisions

One important additional obligation under the GDPR is the prohibition against fully automated decision-making (also known as robo-decisions) where the decision has legal effects on an individual or otherwise significantly affects them. It can only be taken by automated means as explained below:

- **Human involvement**: If a human is involved in the decision-making process, it will not be a decision based solely on automated processing. However, that involvement would have to be meaningful and substantive. It must be more than just rubber-stamping the machine's decision.
- **Explicit consent**: Robo-decisions are permitted where the individual has provided explicit consent. While this sounds like an attractive option, there is a very high threshold for consent.
- **Performance of contract**: Robo-decisions are also permitted where they are necessary for the performance of a contract, or in order to enter into a contract, with an individual. An example might be carrying out credit checks on a new customer.
- **Authorised by law**: Finally, robo-decisions are permitted where authorised by law.

Even where robo-decisions are permitted, suitable safeguards must be put in place to protect the individual's interests. This means notifying the individual and giving them the right to a human evaluation of the decision and to contest the decision.

GDPR in the UK post-Brexit

The data protection framework in the UK is still largely unaffected by Brexit. The UK GDPR forms part of assimilated law in the UK, and applies rules that are substantially the same as those under the EU GDPR.

The UK Government has made a number of amendments to this framework through its Data (Use and Access) Act 2025, although changes are modest and incremental. They include a slight relaxation to the rules on the use of robo-decisions and broader rights to use personal data for research.

Just as important as the black letter of the law is the way in which the flexible principles in the UK GDPR are interpreted and enforced in practice. The Information Commissioner is seen as a pragmatic regulator and, in relation to AI, has taken a more permissive approach than EU regulators, which has allowed greater innovation and faster product development in the UK.

Challenges reconciling generative AI with the GDPR

There are difficult outstanding questions about how to reconcile the training and use of generative AI models with the GDPR. These are not just theoretical risks. At the end of March 2023, the Italian data protection authority (the Garante) banned ChatGPT (before rescinding it a month later) and continues to ban DeepSeek. The Irish Data Protection Commission sought an emergency injunction to prevent X from training its AI on public posts. The EDPB issued guidance on AI models at the end of 2024, but there are a number of outstanding questions likely to be tested by both regulators and privacy activists:

- **Can you use public data to train generative AI models?** Generative AI tools, particularly large language models, are trained on many terabytes of data, much of it likely to have been sourced from the public internet and containing vast amounts of personal data. The only legal basis likely to justify this use is the so-called 'legitimate interests test'.[2] Recognising the many benefits of AI and the implicit approval by the EU legislature of general purpose AI models in the EU AI Act, the EDPB's guidance provides a cautious approval where appropriate steps are taken to mitigate any data protection risks. The questions about the legal basis for training models on data that includes special category personal data remain unanswered. However, at least one court has concluded that this can be accommodated through a flexible application of the rules on special category data (similar to that applied to Google by the CJEU).[3]

2 Any processing of personal data requires a "legal basis" under Art 6(1), GDPR. One such legal basis is the legitimate interests test which requires: (a) there to be a legitimate interest in the processing; (b) the processing to be necessary to achieve that interest; and (c) that the interests of the controller conducting the processing (or other third party) are not outweighed by the interests of the individuals whose personal data is processed.

3 See the decision of the *Oberlandesgericht Köln*, 15 UKl 2/25, 23 May 2025 which applied the CJEU's flexible approach to special category personal data in GC (C-136/17).

- **How does the accuracy principle apply to generative AI?**
  One of the core principles in the GDPR is accuracy. However, this is difficult to reconcile with the 'hallucinations' produced by generative AI models, including those about individuals. The providers of these systems argue that the accuracy principle should be applied contextually, so, given that these 'hallucinations' are well known, the principle does not require strict literal accuracy.[4] However, it is not clear if data protection regulators will accept this approach.
- **What about individual rights, such as the right to be forgotten?**
  The GDPR provides users with a range of rights, including a right to object to the processing of their personal data and the right to erasure. Mapping these rights on to generative AI is potentially challenging. For example, if I exercise my right to erasure, must the provider of the generative AI system trawl their many terabytes of training data to expunge every reference to me? Must they then undertake the computationally expensive process of re-training the whole model without that data? Many tech companies have taken a pragmatic approach to this issue by allowing registered users to opt out of having their data used to train the model initially and then using "output filters" so that other individuals who object can have their personal data filtered out of responses.
- **How is prompt data dealt with?**
  The providers of generative AI models will often want to use the input prompts to that system, and any feedback on the output, to build a better mousetrap. However, the users of these systems will need to consider carefully if any personal data they provide in prompts to the system can be used in this way, particularly if the personal data is confidential or sensitive. In practice, most professional users of generative AI will likely want a sealed instance in which none of that prompt information leaks back to the provider or, at the least, the provider must keep the prompt data confidential and only access it as data processor.

UK ICO Guidance on AI and data protection

The Information Commissioner's Office has taken a keen interest in the potential impact of AI and provides detailed guidance (updated from time to time) on how to comply with data protection law. It addresses a wide range of issues, including:

- **The accountability and governance implications of AI**: This explores how to set a meaningful risk appetite, the need for a data protection impact assessment (DPIA) for most AI projects and whether the processing is undertaken as controller or processor.
- **The need to ensure lawfulness, fairness and transparency**: This explores a range of topics, including identifying the lawful basis for processing. It also provides a useful overview of how to mathematically assess the accuracy of the AI system by reference to both the precision and recall of the system. Finally, this tackles some of the difficult issues of discrimination arising from imbalanced or discriminatory training data.
- **The role of security and data minimisation**: This looks at general data security and also some of the specific attacks on AI systems, including inversion attacks to find 'data ghosts'. Such attacks aim to recover the original (personal) data on which an AI system was trained or on which it relied, raising novel privacy concerns. As part of this, it discusses the differences between black-box and white-box attacks.
- **Providing individual rights**: This discusses how various data protection rights apply with a detailed analysis of the restrictions on robo-decisions.

In addition, the Information Commissioner has issued guidance on Explaining decisions made with AI (updated from time to time) in conjunction with the Alan Turing Institute. This, amongst other things, contains a detailed overview of the various tools currently available to better understand the internal operation of an AI algorithm. While some are likely to be helpful, they do not provide a complete solution and complex multi-dimensional artificial intelligence models are likely to remain largely unknowable.

4 This contextual approach to accuracy has been acknowledged by the English courts, albeit in very different circumstances (see *AB v Chief Constable of British Transport Police* [2022] EWHC 2749).

A note on GDPR and agentic AI

The latest AI trend could also create significant GDPR friction. The 'agentic' concept involves giving an AI agent the authority to act on behalf of another entity. My AI agent might, theoretically, select and purchase a holiday for a university reunion. However, in order to do this, the AI agent will need to know a great deal about me. For example, it might need access to my calendar to find suitable dates, access to my contacts to get details of my university buddies, access to my mail box to message them and access to my bank details to pay for the holiday.

This raises significant questions about who else gets access to this personal data, what controls are available to me and what purposes the personal data is used for. There are technical solutions to help address these issues, such as running the agent locally and giving the AI agent authority over a specific blockchain payment wallet (as a twenty-first century *peculium*) rather than being given free rein over my bank account, but this is a developing area.

Impact of data protection regimes in Asia and the US

The combination of the 'Brussels effect' and the GDPR's formal extra-territorial application means that the GDPR's reach extends beyond the EU borders. It will be interesting to see if this principle also applies to the data protection issues raised by AI or if other jurisdictions will take a more liberal approach to that currently adopted by EU regulators (particularly in relation to generative AI).

We summarise the position in the other key jurisdictions covered in this report below.

Mainland China

- China released a national policy paper in 2022, emphasising data as a strategic national resource while also laying the groundwork for future policies related to data ownership, trading, profit distribution, management and supervision.
- Implementing this policy, China's 17 ministries, including the financial regulators, the People's Bank of China and National Financial Regulatory Administration, jointly issued the "Data Element X" Three-Year Action Plan, seeking to boost use of data in 12 sectors, including the financial service industry, from manufacturing and finance to technological innovation and green and low-carbon development.
- With the implementation of the Cybersecurity Law in 2017 and the Data Security Law and the Personal Information Protection Law (PIPL) in 2021, together with the Network Data Security Management Regulations in 2025 implementing these three-pillar data laws, China has been seeking to ensure state control and protection over valuable personal and non-personal data, including storage of data on Chinese users within the country as a general principle, and the mandatory or recommended use of China's national standards for AI, including over big data, cloud computing and industrial software. The PIPL shares similarities with the GDPR, governing the processing of personal data crucial to AI systems.

  Read more: China's new privacy law passed: the wait is over (August 2021).
- PIPL and several implementation rules and other best practice guidelines apply when data controllers adopt automated decision-making systems that may influence data subjects' interests (such as automated decisions empowered by AI and big data analysis relating to personal credit or loan limits). It requires those data controllers to conduct personal data protection impact assessments of new tools before their deployment and then periodically, and ensure that there are readily accessible channels for customers to object to such automated decision-making, followed by a manual review of the complaints.
- In the United States, federal regulations implementing Executive Order 14117 (Security Requirements for Restricted Transactions) took effect in July 2025. The new rules could constrain bulk flows of personal data from the US to China and other so-called 'countries of concern'. Developers of AI tools who are located in China, subject to Chinese government jurisdiction or majority-owned by Chinese owners, may find it harder to access US personal data to train LLMs as a result of these regulations.

Read more: Data Protected - People's Republic of China | Linklaters

Hong Kong SAR

> Hong Kong's data protection legislation, the Personal Data (Privacy) Ordinance (Cap. 486) (PDPO), does not impose GDPR-like AI-related safeguards on automated decision-making. However, the technology-neutral PDPO governs the collection and handling of individuals' personal data in AI development and use through the six data protection principles:
  1. data must be collected in a lawful, fair and non-excessive manner and adequate notices must be given to individuals;
  2. data must be accurate, up-to-date and be kept for no longer than is necessary;
  3. data must be used for the purposes for which it was originally collected or a directly related purpose, otherwise individuals' consent must be obtained;
  4. data must be protected against unauthorised or accidental access, processing, erasure, loss or use;
  5. entities holding data must be transparent about the kinds of personal data they hold and the main purposes for which personal data is used; and
  6. data subjects must be given rights to access and correct their data.
> The Privacy Commissioner's Guidance on Ethical Development and Use of AI (2021) assists organisations developing AI systems in complying with the PDPO requirements, including adhering to ethical principles such as accountability, transparency, fairness and ensuring human oversight in the development and use of AI.
> The Privacy Commissioner then expanded on the 2021 guidance with its AI Model Personal Data Protection Framework (2024). This is comprehensive best practice guidance for organisations procuring, implementing and using AI solutions (including generative AI and predictive AI) to assist them with how to comply with the data privacy and governance compliance requirements of the PDPO.

  Read more: Hong Kong Privacy Commissioner's new Model AI Framework – What does it mean for businesses adopting AI solutions in Hong Kong? (June 2024).

Read more: Data Protected - Hong Kong | Linklaters

Singapore

Personal Data Protection Act framework

> While the Personal Data Protection Act 2012 (PDPA) of Singapore does not have specific provisions relating to the use of AI, it sets out a data protection framework on the collection, use and disclosure of personal data by private sector organisations in Singapore to protect the personal data of individuals, support public trust in the digital economy and enable innovation in the data space.
> In order to collect, use or disclose an individual's personal data, the PDPA requires an organisation to obtain the individual's prior consent and notify the individual of the purpose of collection and intended use of their personal data, unless an exception under the PDPA applies (eg, if the personal data is publicly available or the collection of personal data is necessary to respond to an emergency that threatens the life, health or safety of the individual or another individual).

Advisory Guidelines on the Use of Personal Data in AI Recommendation and Decision Systems

> Alternative processing grounds to consent have also been introduced, such as the legitimate interests exception or the business improvement exception. In particular, the Personal Data Protection Commission (PDPC) Advisory Guidelines on the Use of Personal Data in AI Recommendation and Decision Systems (March 2024) highlights the business improvement exception and the research exception as potential processing grounds for the development, testing and monitoring of AI systems.
> The PDPC also reiterated that various principles of the PDPA apply in the context of the training and deployment of AI systems. This includes, for example, the requirement for consent and notification (or an available exception) as mentioned above.
> When organisations appoint service providers to provide professional services for the development and deployment of bespoke AI systems, organisations remain responsible for ensuring the compliance of such AI systems with the PDPA. Organisations have a general obligation to be accountable for the processing of personal data, such as developing clear policies and procedures for transparency and developing trust.

IMDA data protection and AI initiatives

Singapore has introduced three new initiatives focused on data protection and AI adoption:

> The Global AI Assurance Sandbox, as mentioned above in Section 2 for the testing of generative AI applications.
> Following the IMDA's Privacy Enhancing Technologies (PET) Sandbox, which provides opportunities for companies to work with PET digital solution providers to develop use cases and pilot PETs, the IMDA has now also produced a draft adoption guide based on insights from the PET Sandbox, which is focused on helping organisations choose and implement PET solutions.
> In partnership with Enterprise SG and the Singapore Accreditation Council, the IMDA has also made the Data Protection Trust Mark a "Singapore Standard". Businesses can now seek certification under this mark to demonstrate that their data protection practices meet internationally-recognised standards.

Read more: Data Protected - Singapore | Linklaters

US

State privacy law and AI

> In the absence of a national privacy law, US states continue to pass comprehensive privacy laws at record pace. Altogether, 19 states have passed privacy legislation. In the past year, these include: Kentucky, Maryland, Minnesota, Nebraska, New Hampshire, New Jersey, Rhode Island and Tennessee. Many of these state laws address the use of personal information in artificial intelligence systems with respect to automated decision-making technologies (ADMTs) by providing residents the right to opt out of certain automated decision-making, often cabined to that which has a legally or similarly significant impact on the resident.
> See the discussion of artificial intelligence specific legislation enacted in Colorado and introduced in Congress in Chapter 2.

Privacy law, automated decision-making and avoiding bias

> The widespread use of ADMT has the potential to deny consumers access to fundamental services, including financial and lending services, along with necessities such as housing, healthcare, insurance, and employment and educational opportunities. State-wide AI legislation has sought to address concerns with biased and discriminatory uses of this technology, particularly where there is little or no human intervention (for example, Colorado's comprehensive AI Act and California's 18 AI Regulations, including SB-926 and SB-891, aimed at protecting against deepfakes, and AB-2013 requiring transparency in training data).
> At the federal level, President Trump repealed Biden-era Executive Order 14110 on AI and issued multiple AI Executive Orders, alongside the new AI Action Plan, as discussed in Chapter 2.
> Except for Utah and Iowa, every state that has passed comprehensive privacy legislation has afforded its residents the right to opt out of at least certain types of ADMT. Of these, the California Privacy Rights Act (CPRA), which amended the California Consumer Privacy Act (CCPA), grants state residents the most protections, including requiring businesses to provide information about any automated decision-making in response to a resident's data subject access request. These regulations, which will continue to come into effect across the country in the coming months, underscore the importance of transparency with consumers around the use of AI, algorithms, data sets containing personal information, and the use of ADMTs. Algorithmic bias, alongside the potential for algorithms to generate discriminatory or unlawful outputs, presents a real threat.
> Meanwhile, California regulators are using their existing rulemaking authority under the California Consumer Privacy Act (CCPA) to provide consumers with added transparency and opt-out rights relating to the use of their personal data for automated decision-making for "significant decisions" similar to those scoped in the Colorado AI Act.

Privacy law impacting facial recognition

> Some state privacy laws address the use of facial recognition technology. As discussed earlier, Connecticut, Delaware, Florida, Indiana, Iowa, Montana, New Jersey, Oregon, Tennessee, Utah and Virginia have each enacted comprehensive privacy laws that include biometric data and/or facial recognition as a protected category. Other states and local cities or counties have established limits on law enforcement usage of facial recognition technologies (see, eg, Alabama, Boston, Baltimore and Vermont). Colorado's AI Act regulates facial recognition to the extent that it serves as a part of a high-risk AI system that makes or is a substantial factor in making consequential decisions about Colorado residents.
> There are also several biometric privacy laws in the US, including in Illinois, Texas and Washington, which would impact the use of AI for facial recognition and biometric evaluation. For example, Illinois and Texas have implemented strong protections against non-consensual facial recognition and other biometric technologies, through BIPA and CUBI. These laws generally require companies to obtain consent prior to the collection, use and processing of an individual's biometric information (including, but not limited to, records of facial geometry, iris scans and voiceprints).
> Federally, there is no facial recognition law, and there has not been much focus on this at a federal level under the current administration. Concurrently, measures have been introduced in Congress to prohibit the Federal Government from certain uses of facial recognition technology (HR 3782).

Read more: Data Protected - U.S.| Linklaters

5 Regulating AI through competition law – key issues to consider

Antitrust regulators are generally focused on the impact of frontier technologies and developing digital markets on competition. How businesses use emerging technologies, especially AI, is emerging as a key theme for antitrust regulators across the globe. Regulators are not only considering whether competition law is fit for purpose with respect to the impact of AI – for example, with respect to algorithmic collusion, hub and spoke arrangements, tacit collusion and broader harms – but are also balancing these risks with a desire to fuel financial growth and technological innovation.

Impact of competition law in financial services

Competition regulators have historically taken a principled approach to intervention in markets – focusing on maximising the efficiency of the competitive process. Within this framework, AI provides a number of consumer benefits, for example improved innovation, efficiency and potentially lower prices as manual-driven processes are replaced with ML and the associated cost benefits are passed on to consumers.

Within the financial services context specifically, AI could facilitate greater simplicity for consumers and investors alike, more effective and potentially automated identification and control of investment opportunities and risks. However, regulators are increasingly mindful of the potential antitrust risks posed by AI applications.

As explained in Chapter 3, while the underlying rules to promote competition in markets remain largely unchanged, businesses' compliance with those rules will still need to consider the application and effects of AI.

Antitrust regulatory approaches in key jurisdictions

Different approaches are being taken in the key jurisdictions we are focusing on in this report:

EU and UK approach

- **Specific focus on algorithms**: Whilst details have not been made public, the European Commission has disclosed that it is investigating algorithmic pricing at the EU level. Additionally, the European Commission's Joint Research Centre (JRC) announced a new research project into algorithmic-driven risks in online platforms and search engines, inspired by, and supporting, existing work associated with the Digital Services Act (DSA).
- **Scrutiny of digital markets**: Regulators are closely scrutinising developments in digital markets, in particular in payments, to ensure that companies have an incentive to innovate and that dynamic fintech markets are open to competition. Since the development of machine learning, complex algorithms and systems capable of processing vast quantities of data are facilitating the commercialisation of that data. However, they also create the potential for price-fixing and other algorithmic collusion and cartel creation through information-sharing via 'hub and spoke agreements' (discussed further below).
- **Regulatory cooperation**: In the UK, we have historically seen significant cooperation between regulators in the financial services sector, given the FCA's and the CMA's concurrent powers for enforcing competition law. Competition enforcement in financial services firms has historically focused on individual (mis)conduct issues, with a number of large and high-profile enforcement cases being brought by regulators across jurisdictions (for instance, collusion with respect to LIBOR, FX and Supra Sovereign Bonds and derivatives).

- **Work by the Competition and Markets Authority (CMA)**: The CMA has historically taken an active involvement in researching and understanding the potential competitive impacts of AI technology, conducting reviews of AI foundation models, before publishing an AI strategic update in April 2024. The CMA's strategic update suggests that the regulator's main competition concern around AI is the small number of tech firms which are incumbent in the various stages of the AI value chain, and which could "profoundly shape the development of AI-related markets". As of Spring 2024, the CMA indicated its ongoing consideration of potential competition concerns associated with AI innovations, including through its merger control and markets powers. However, having scrutinised five separate AI partnerships between incumbents and specialist providers, the CMA concluded that it either lacked jurisdiction to review the agreements or concluded that the partnerships were not problematic.

China approach

- **Antitrust as a key regulatory tool**: The message from the top leadership in the Chinese mainland which has been consistently maintained since 2020 is to "strengthen anti-monopoly and prevent disorderly expansion of capital". One of the triggering events of a series of internet antitrust enforcement actions concerned capital penetration into retail banking and digital payments by mega internet platform operators. In this respect, antitrust is one of the key tools for regulating the financial industry. Meanwhile, since 2023, authorities have also been making efforts to facilitate economic growth and rebuild confidence through streamlined procedures, soft law tools, and public-private engagement as another aspect of the enforcement approach.
- **Supervisory framework for non-bank payment institutions**: The People's Bank of China, China's central bank, issued regulations in 2023 addressing potential abuse of dominance by non-bank payment institutions – formalising the PBOC's supervisory powers and clarifying the State Administration for Market Regulation (SAMR) enforcement role. If a single payments institution reaches one-third of the market share, or two institutions together reach half, the PBOC may initiate interviews and reviews to assess dominance Institutions with over 50% market share in electronic payments may be deemed monopolistic. The framework now supports full-cycle supervision, including licensing, equity structure oversight, and executive suitability assessment.
- **Use of Big Data to digitalise and facilitate regulation**: The top authority has aimed to digitalise market regulation through the introduction of big data research centres at both the national and provincial level. This is intended to improve market regulation capacity through the introduction of a new set of regulation tools such as digitalised regulatory platforms, facilitation of data and resource sharing and enforcement coordination.

Specifically, SAMR created a Competition Policy and Big Data Centre in 2021 which has since transitioned from a research-focused entity to a strategic enforcement and policy tool within SAMR. It plays a central role in monitoring digital markets, informing merger reviews and abuse of dominance cases and shaping China's evolving antitrust framework in the data economy.

Hong Kong SAR approach

- **Cooperation between regulators**: The Hong Kong Competition Commission (HKCC) signed a Memorandum of Understanding (MoU) with the Hong Kong Securities and Futures Commission in 2020 to enhance cooperation and information exchange between the two agencies. The MoU remains active and continues to serve as a foundational framework for inter-agency cooperation in Hong Kong's financial regulatory landscape. In addition, HKCC also signed a MoU with the Guangdong Administration for Market Regulation (GDAMR) in July 2023 to bolster competition policy and enforcement within the Greater Bay Area (GBA). The collaboration allows HKCC to access to more data and intelligence on financial markets or digital platforms.
- **No formal enforcement**: To date, the HKCC has not formally brought any enforcement actions against financial institutions, even though certain conduct of financial institutions had previously come to the attention of the HKCC (eg, Code of Banking Practice). This Code was revised in December 2023, with enhancements focused on digital banking, consumer protection and transparency. These revisions were led by the Hong Kong Association of Banks (HKAB) and the DTC Association, with endorsement from the Hong Kong Monetary Authority (HKMA), not the HKCC.
- **Future scrutiny is possible**: While HKCC has not yet taken any enforcement action in financial sector, its collaboration with the SFC and interest in financial sector competition suggest that future scrutiny is possible, especially if anti-competitive conduct like collusion, exclusive dealing, or abuse of dominance is identified.

US approach

- **Continued focus on tech and fintech**: The current administration is still actively scrutinising Big Tech and digital markets from an antitrust perspective, continuing major lawsuits and merger challenges in 2025. Particular focus remains on where Big Tech and financial platforms may be seen as gatekeepers. Novel uses of pricing technology also remain a priority.

A word on consumer protection

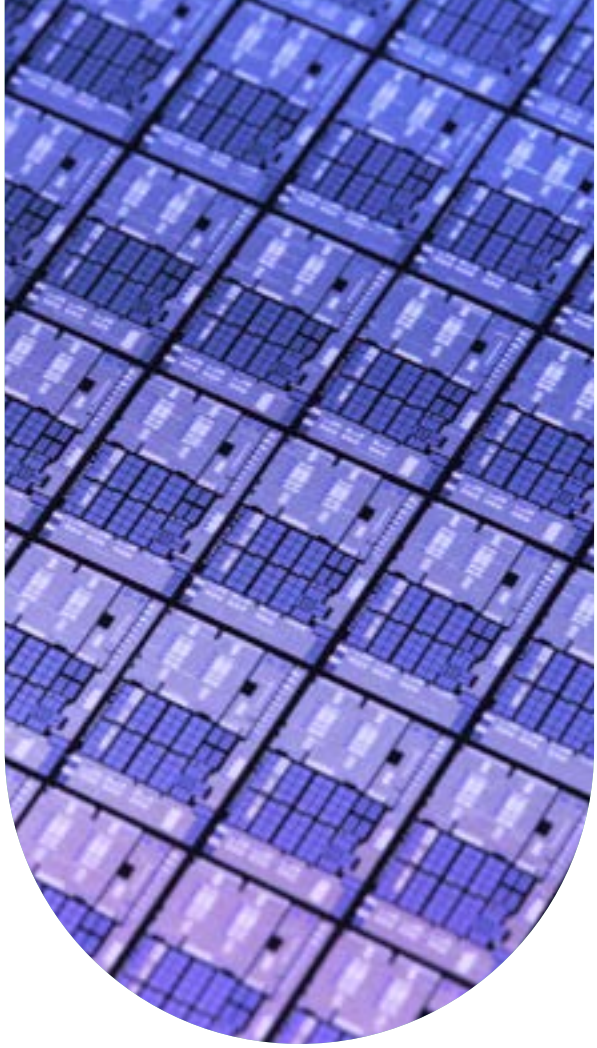
Regulatory interventions have also been shaped by consumer law objectives to reduce unfair contractual terms and unfair commercial practices which may negatively impact consumers. Authorities are paying special attention to practices in the digital markets sector, including the use of online choice architecture, subscription contracts, discounting and urgency claims.

For example, in the UK, the FCA's new Consumer Duty and the CMA's new direct enforcement powers, under the Digital Markets, Competition and Consumer Act (DMCC Act) which came into force in early 2025 increase regulatory focus on consumer protection. We have summarised the two regimes in our recent article. The DMCC Act introduces substantial fines of up to 10% of global turnover for companies which fail to comply with consumer law.

The CMA has indicated its intention to use its new direct enforcement powers in collaboration with other regulators, such as the FCA and Information Commissioner's Office (ICO), recently publishing a joint article with the ICO on how UK regulation supports a range of open and closed-access foundation models. Although elements of financial services are carved out of the DMCC Act (for example, the provisions on subscription contracts), the more general provisions on unfair commercial practices and unfair contractual terms could apply in consumer-facing activities in the market.

Algorithmic-backed dynamic pricing has come under scrutiny by the CMA, for example, in its Ticketmaster investigation. In practice, many businesses do not appear to be using dynamic (or 'surge' pricing in the manner expected by the CMA, however, the tone is clear: this is an area in which the CMA are keen to ensure transparency and fairness and in which the regulator is willing to take action to protect consumers. We understand that the CMA is likely to prioritise enforcement against fake reviews, drip pricing, misleading pricing and urgency claims and greenwashing, in line with its recent enforcement practices and guidance.

Meanwhile, in the US, under the current administration consumer protection in consumer-facing sectors such as consumer banking and payments has become less robust. Key Biden-era rules – such as the cap on overdraft fees and expanded oversight of digital payment platforms – have been repealed, and the Consumer Financial Protection Bureau (CFPB) has significantly reduced its enforcement activity, dropping numerous pending cases and slowing new rulemaking. As a result, state attorneys general have stepped up enforcement efforts, but their authority is more limited than the CFPB's. While broader consumer protection enforcement has continued at the FTC, the agency leadership has clearly moved away from broader rulemaking. While federal bank regulators continue some compliance oversight, the overall regulatory approach is now more deregulatory, with less focus on the consumer than in previous years.



Is competition law fit for purpose?

Competition authorities are increasingly turning their attention to digital markets and the effect of innovative technology on competition. While the development of machine learning, complex algorithms and systems capable of processing vast quantities of data has led to innovative commercial applications for AI, there remains a lack of certainty as to the precise effects such algorithms are likely to have on competition (positive and negative) and, as a result, there remain differing views as to how competition law should deal with these developments.

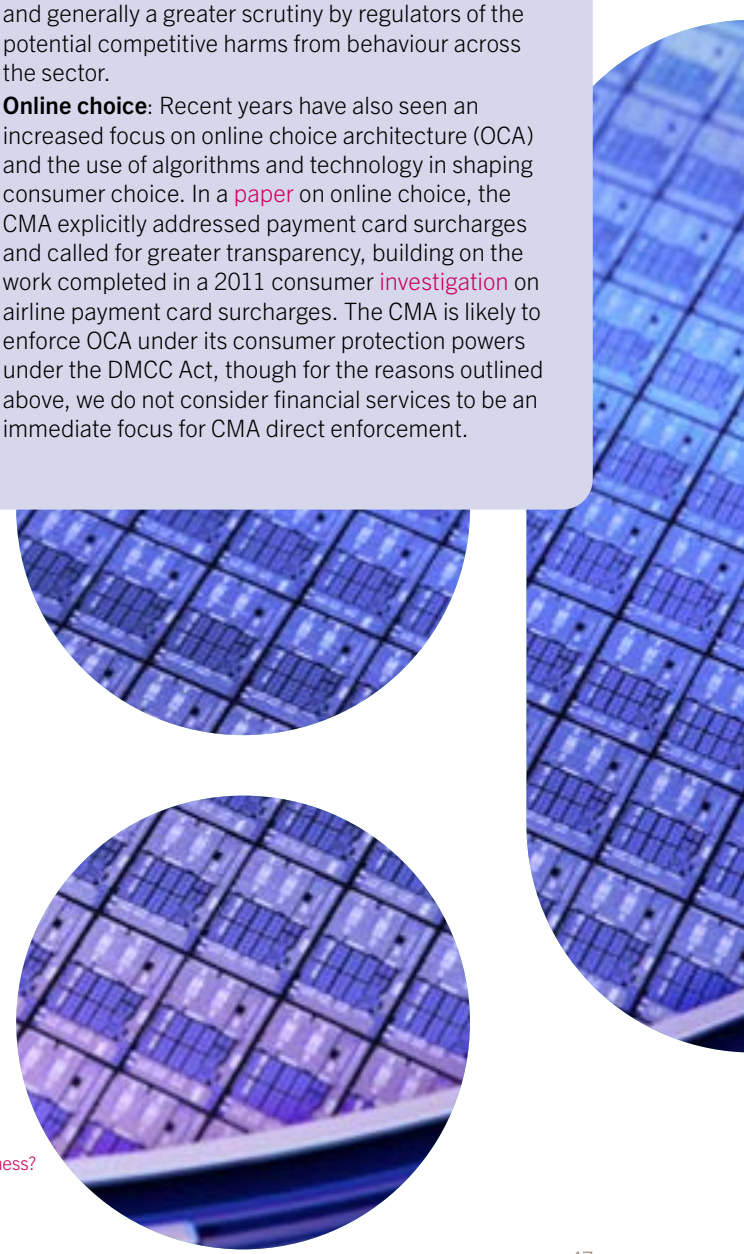
Amid debate around the sufficiency of current regulations in their application to technological developments, some competition authorities have progressed reforms in support of more nuanced tech-specific regulation.

EU approach to tech-specific regulation

- **Digital regulation**: In the EU, the European Commission (EC) has supported sweeping reforms to regulate digital markets, including the new Digital Markets Act and Digital Services Act.
- **Digital Markets Act:** The DMA came into force on 1 November 2022 and on Big Tech, introducing a number of reporting and compliance obligations for so-called gatekeepers, including in relation to interoperability, data-sharing obligations (for the benefit of both end consumers and third-party advertisers who use platform services) and acquisition strategy. The first Gatekeepers were designated by the EC on 6 September 2023 and gatekeepers had until March 2024 to ensure that they followed their obligations under the DMA.
- **Gatekeepers**: Whilst financial services companies themselves are unlikely to be designated gatekeepers (unless considered online intermediaries), digital market leaders who operate in the payments sector, for example through app stores, may see increased scrutiny of transactions, including in the payments space, and be subjected to enhanced reporting and compliance obligations under the DMA, and traditional financial services players could seek to leverage the digital regulations to their advantage in any interactions with so-called 'gatekeepers'.
- **Digital Services Act:** The DSA also came into force in November 2022 and introduces EU-wide rules on liability for digital platforms and the content, products and advertisements those platforms host/distribute. The EC will designate platforms with over 45 million users (10% of the population in Europe) as very large online platforms (VLOPs) or very large online search engines (VLOSEs). These services will have to comply with their obligations under the DSA, including carrying out and providing the EC with their first annual risk assessment, within four months of being designated.
- **EC European Centre for Algorithmic Transparency (ECAT)**: Cooperates with industry representatives, academia and civil society organisations to improve the EC's understanding of how algorithms work, analysing transparency and assessing risks, and will propose new transparent approaches and best practices in line with new measures under the DSA that call for algorithmic accountability and transparency audits.

UK approach to tech-specific regulation

- **Digital Markets Unit**: The UK's CMA has sought to remodel itself to better address the potential harms caused in digital markets, through its Digital Markets Unit (DMU), designed to tackle key competition issues in the regulation of digital markets.
- **Strategic Market Status**: The DMU's primary focus is also the regulation of Big Tech (ie, firms with 'Strategic Market Status'). Although with time this could have knock-on impacts on financial markets (eg, targeting large players in the digital markets which incorporate payments within their business models), for the moment architects of blockchain crypto providers are unlikely to be a focal point for the proposed regulations.
- **Use of algorithms**: In a paper on algorithms and competition, the CMA has also highlighted a particular tension in the use of algorithms. It has called for an increased audit of the algorithms used by companies via several different routes, including opening formal investigations where it thinks a business's algorithms may have infringed competition law, sandbox testing and checking on algorithmic systems that have previously been investigated, a role that will likely be taken up by the DMU. This builds on the work of the FCA[5], with many of the issues raised falling outside the scope of traditional competition law harms, and generally a greater scrutiny by regulators of the potential competitive harms from behaviour across the sector.
- **Online choice**: Recent years have also seen an increased focus on online choice architecture (OCA) and the use of algorithms and technology in shaping consumer choice. In a paper on online choice, the CMA explicitly addressed payment card surcharges and called for greater transparency, building on the work completed in a 2011 consumer investigation on airline payment card surcharges. The CMA is likely to enforce OCA under its consumer protection powers under the DMCC Act, though for the reasons outlined above, we do not consider financial services to be an immediate focus for CMA direct enforcement.



5 FCA, Price discrimination in financial services. How should we deal with questions of fairness? (July 2018) and General Insurance Pricing Practices: Interim Report (October 2019).

Mainland China approach to tech-specific regulation

- **Antitrust Guidelines for Platform Economies (2021)**: These include and acknowledge technologies, data, algorithms and platform rules as means that firms may use to reach or implement monopoly agreements, including hub and spoke arrangements and abuse of dominance practices.
- **Amendment to Anti-Monopoly Law (2022)**: This prohibits monopolistic conduct through data, algorithms, technologies, capital advantage and platform rules, and introduces liabilities for those that organise or facilitate monopoly agreements. This will mean, for example, tech companies, playing intermediary roles by providing data and algorithms, could be liable for facilitating monopoly agreements, even though they may not themselves be engaged in the monopolistic practices.
- **AI regulations with an antitrust element (2022-2023)**: These prohibit monopolistic practices (eg, imposing unreasonable restrictions on users, etc.) by using algorithms, data or platform advantages, for example in the areas of algorithm-based recommendations and generative artificial intelligence.

Read more: China: SAMR joins ranks and sends a strong signal for digital markets (January 2021)

Read more: China Amends 14-Year-Old: Anti-Monopoly Law – what has changed and implications for business (June 2022)

Hong Kong SAR approach to tech-specific regulation

- **No specific guidance**: While the digital economy is one of the HKCC's targeted sectors, the HKCC has not issued any specific guidance on antitrust issues relating to artificial intelligence, technology or digital markets.
- **Scrutiny of digital space**: This lack of guidance, however, has not prevented the HKCC from scrutinising potential anti-competitive conduct in the digital space – the HKCC had previously raised competition concerns on vertical issues involving hotel booking platforms and food delivery platforms, resulting in these platforms changing their market conduct to address the HKCC's concerns.

US approach to tech-specific regulation

- **Existing laws**: No new laws have been implemented to deal with the antitrust concerns around the growing use of AI in the market and the DOJ and FTC have continued to use existing laws to bring claims for the anticompetitive use of AI.
- **Impact of AI on competition**: Overall, the agencies are considering AI just like any other tool to consider how it impacts competition between companies, including potential for signalling, tracking or predicting competitive conduct. One key area is the impact on potential coordination and information exchange. The agencies' draft Merger Guidelines have also focussed on the widespread use of algorithms or AI to track or predict competitor prices or actions, which is in turn seen to promote coordination in concentrated markets by increasing market transparency. The DOJ and FTC are also watching for signs of large tech and non-tech players (eg, big banks) seeking to acquire or undermine disruptive new AI players. The current FTC commissioner has also acknowledged that AI may impact market definitions as AI can be an input and a competitor in several areas.
- **Existing principles**: Algorithms have been targeted in prior enforcement as a means of communicating in broader cartel arrangements in online markets, as a means of collecting sensitive information of competitors and as a means of testing price fixing methods. Back in 2017, the DOJ and FTC concluded in a joint policy paper that existing antitrust principles remain capable of addressing the potential harm to competition presented by the use of new technological tools. In 2024, the FTC and DOJ released a joint statement along with the EU and CMA which outlines three guiding principles concerning the antitrust harms of AI: (1) fair dealing and avoiding exclusionary tactics; (2) interoperability of AI products and services; and (3) choice for businesses and consumers in choosing AI products. The new administration has not disturbed these findings and has used the existing antitrust framework to bring cases against companies who have used AI to harm competition.

Enforcement action to date

Algorithmic collusion

To date, enforcement of competition law regarding algorithms has involved classic collusion, implemented through novel means, for example fixing (and monitoring) prices or facilitating anti-competitive agreements[6]. Regulators on both sides of the channel have recognised that "*where algorithms are designed by humans to [coordinate behaviour], this is merely a new form of the old practice of price-fixing*"[7] and have emphasised that "*[C]ompanies can't escape responsibility for collusion by hiding behind a computer program*"[8]. Algorithmic pricing has historically been considered academically, but not by regulators, though this trend appears to be changing. Whilst details have not been made public, EU Deputy Director-General Linsey McCallum confirmed that there are currently several EU investigations into algorithmic pricing.

In the US, the most prominent US enforcement action has been carried out by the FTC along with several states, who together sued Amazon for using an algorithm called Project Nessie to monitor competitor pricing, and to test if competitors will increase prices in response to Amazon's price increases. The FTC brought their claims under Sec 5 of the FTC Act which penalises unfair methods of competition.

There remain a number of ways in which antitrust laws could be utilised to manage misconduct effected through AI in the financial markets. We are perhaps most likely to see antitrust enforcement 'bite' to sanction collusion between market participants.

As we continue to see technological developments in the financial services space, it should be noted that the nature of decentralised finance (DeFi) presents special practical challenges for enforcement with respect to algorithmic collusion in a blockchain context. This can occur on decentralised exchanges for example, where trading bots and smart contracts can be programmed to monitor and react to competitors' pricing strategies in real time, potentially leading to tacit or even explicit collusion on token prices. Regulators may struggle to identify collusion and the key players involved, as this is inherently (and purposefully) difficult in this context given the lack of transparency in ownership, hash power and decision-making on the chain.

Hub and spoke arrangements

Concerns can arise where several industry players use the same algorithm, which facilitates information-sharing (known as a hub and spoke arrangement). Such algorithms, often provided by a third party, can allow competitors to monitor prices and to thereby determine the 'market price' and/or react swiftly to market developments, all of which can be problematic from an antitrust perspective.

Although several cases in the US have challenged competitors' use of common recommended pricing algorithms as tacit collusion, the law is still developing on whether this should be deemed a 'per se' hub and spoke conspiracy. For example, lower courts considering the use of common pricing software by multi-family building lessors have split on whether to apply a per se or rule of reason approach where there is sufficient evidence of an agreement to use the software.

In the US, the DOJ, state enforcers and private plaintiffs have targeted the use of firms like RealPage that use a pricing algorithm to collect competitively sensitive information from landlords and recommends rent prices to them. While the DOJ has urged courts to treat these as traditional hub and spoke conspiracies, some courts have cautioned that an effects-based approach is more appropriate when dealing with novel technologies. Similarly, two private class actions against several casinos in Las Vegas and Atlantic City alleging that the defendants were using third-party pricing software to fix hotel room prices in the Las Vegas hotel industry were dismissed for failing to sufficiently allege an agreement between the hotel owners to implement the price recommendations.

By way of comparison, in the EU's *Eturas* case the administrator of a Lithuanian online travel-booking system sent an email to its travel agents, notifying the agents of a new technical restriction of the platform that placed a cap on discount rates, which was ultimately viewed as price-fixing between participants in the platform, even though they had no direct contact with other users. While algorithms and AI are omnipresent in many industries, and regulators in both the EU and UK have launched research into algorithms, we have yet to see how competition law enforcers will deal with this new reality (although there are many calls to increase supervision of AI).

The onus remains on businesses to ensure '*compliance by design*', and to understand both the scope of current antitrust laws and the functionality of algorithms before their establishment, since, even where collusion is tacit or is an unforeseen consequence of algorithmic design, businesses will be held responsible for the way their algorithms behave in markets.

6 Eg, CMA, Investigation of Trod re Online sales of posters and frames (September 2016).
7 Eg, CMA, David Currie on the role of competition in stimulating innovation (February 2017).
8 Eg, EC, Algorithms and Competition (Bundeskartellamt 18th Conference on Competition, Berlin) (November 2019).

Algorithms increasing transparency – competition authorities treading carefully

Tacit collusion

Another potential complication stems from algorithms facilitating tacit collusion (whereby firms unilaterally adapt their strategy in light of competitors' behaviour). At present, pure tacit collusion does not constitute an antitrust offence in and of itself, where there is no evidence of collusion, although in the UK the CMA's paper on algorithms specifically addresses the possibility that algorithms are taught to auto-collude.

In its paper, the CMA has noted that firms are responsible for any effective oversight of their systems, including robust governance and impact assessments, even where an algorithm's behaviour is not perfectly anticipated. Nevertheless, with more and more businesses adopting pricing algorithms and posting their current prices, market transparency has increased.

Regulators (for example in Germany and France) are already considering the implications of this sort of development, but the mechanism for addressing these concerns is far from clear. As it is generally agreed that transparency is in principle pro-competitive, in that it allows consumers to easily compare competing offers, competition authorities may be reluctant to intervene to limit this transparency. Furthermore, it is very difficult for any regulator to reliably predict the 'tipping point' from pro-competitive transparency to potentially problematic tacit collusion[9]. Some academics have suggested that algorithms could consistently learn to charge supra-competitive prices without having to communicate with each other[10].

In the US, the DOJ's and FTC's joint 2017 policy paper confirmed that "the use of algorithms may increase price transparency and help to stabilise prices." In 2024, a joint statement by the FTC, the DOJ, the CMA and the EU also highlighted concerns over the use of algorithms in sharing competitively sensitive information, collusion and price fixing.

Broader harms

In the UK, the CMA has identified a number of broader harms that may arise as a result of algorithmic design, including personalised pricing and unfair algorithmic design practices. The CMA's work builds on the extensive work by the FCA and international regulators in this area, which has concluded that unfair pricing practices can lead to reduced competition and higher pricing, to the detriment of consumers[11].

New regulations such as the DMCC Act draw on sanctions for non-compliance with consumer protection as a key tool in combating unfair commercial practices, which is likely to include unfair algorithmic design practices. This is especially the case where algorithms determine pricing practices and discounting, but algorithms are also being reviewed by the CMA more broadly, for example in platforms' ability to monitor compliance on their sites.

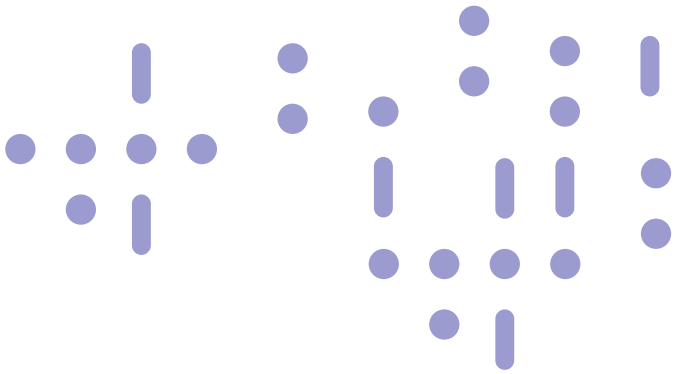
Central to discussions on the ethical design of algorithms are the principles of fairness, transparency and accountability, which have increasingly been emphasised by regulators, including through the Digital Regulatory Cooperation Forum (see Chapter 2).

Avoiding antitrust issues when deploying AI

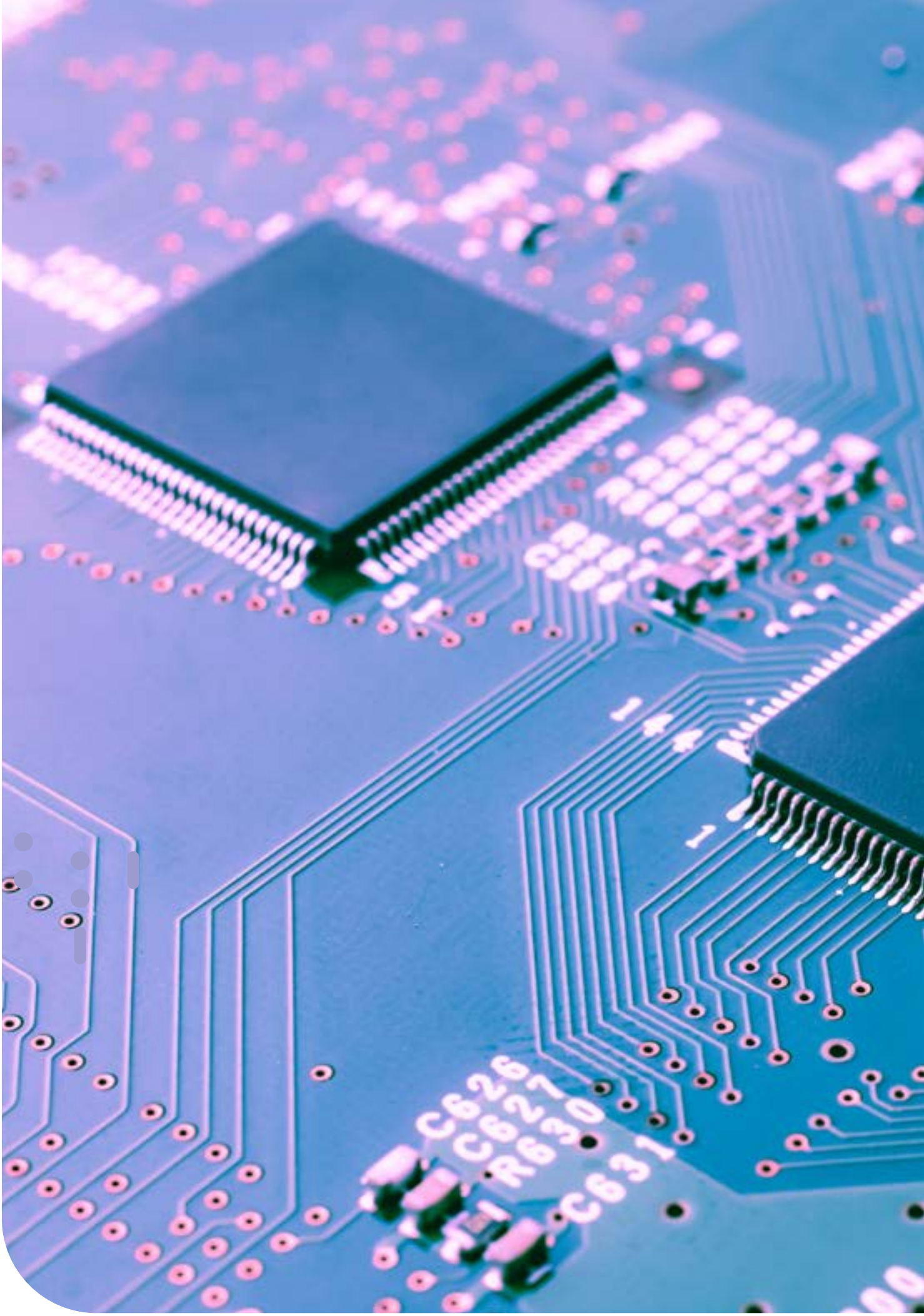
As technology continues to develop and impact both businesses and end consumers, the regulatory landscape applicable to AI is evolving and often unpredictable, meaning that the antitrust implications of using AI are complex and susceptible to change.

Many jurisdictions have not yet launched fully-fledged regimes that regulate the application of technology, leaving many companies bereft of certainty as they consider both the business and the compliance implications of new technological applications.

In light of this, companies should start thinking about these issues and technical ways in which collusion can be prevented when deploying AI solutions in financial services.



9 Eg, CESifo – Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market', Working Paper No. 8521 (August 2020).
10 Eg, CEPR – Artificial intelligence, algorithmic pricing, and collusion (February 2019).
11 FCA, Price discrimination in financial services. How should we deal with questions of fairness? (July 2018) FCA, General Insurance Pricing Practices: Interim Report (October 2019); BEIS, Personalised Pricing and Disclosure (September 2020) and OECD, Personalised Pricing in the Digital Era (November 2018).


C626
C627
R630
C631

6 Practical guidance on managing AI risk

Financial services organisations need to take a holistic, forward-looking approach to anticipating the impact of AI technology on their business. In practical terms, firms need to have a clear understanding of what they want to achieve in deploying any AI tool, and how it will work to achieve that goal, together with a clear plan for identifying and managing any associated technological, reputational and legal risks.

Flexibility is key

It remains to be seen whether governments will coalesce around a particular set of rules for managing AI as the gold standard, in the way that they have done with the GDPR for managing data. Unlike the GDPR, there is a range of different approaches already gaining traction – some with more emphasis on protecting the rights of individuals, while others are more focused on the overall safety of AI for mankind.

As a result, businesses seeking to implement large-scale AI solutions in the near to medium term will need to develop a compliance and risk management strategy that strikes the right balance between local specificity and global consistency. While investments into such strategy do not yield the immediate feel of return that AI trials can produce, they are a pre-requisite to scaling generative AI across the enterprise.

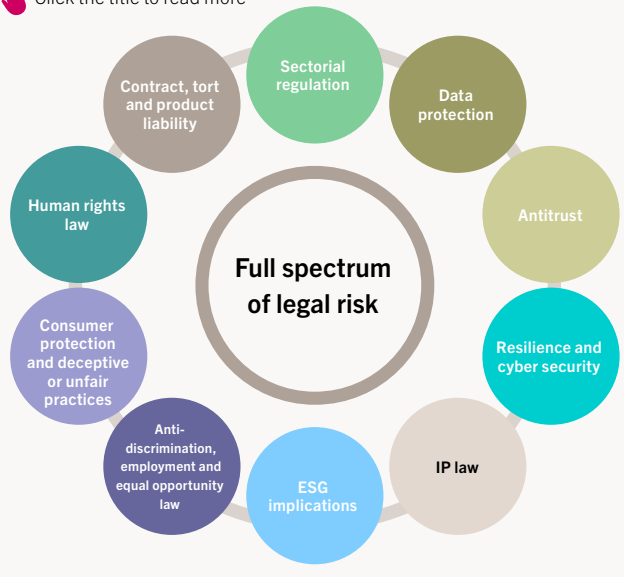
The strategy will also need to be sufficiently flexible to evolve with the changing nature of the technology itself and the varying sets of international rules for AI, as well as the increasing enforcement of existing legal regimes, various of which are being adapted by regulators to focus more effectively on the specific risks of AI.

Mapping AI to a broad spectrum of potential legal risk

In terms of legal risk, firms should appreciate that sectoral regulation, data protection and competition law – as well as developing AI-specific law, regulation and guidance – will have an impact on how they implement AI solutions, as covered by this report. A solid understanding of when those laws and regulations can bite on AI is required to address potential legal and broader risk issues.

A number of these areas are beyond the scope of this report, but they include:

Click the title to read more



Sectorial regulation

In this report we have focused on the various aspects of financial regulation that can apply when deploying AI in financial services (see Chapter 3).

The overlapping nature of AI-specific regimes with those for operational resilience, data protection, outsourcing and third-party risk, financial services conduct and fairness requirements, IP law (particularly as applicable to model training data) and competition law (such as inadvertent data sharing via pretrained models) presents a complex compliance challenge.

Financial institutions need to establish integrated risk assessment flows – bringing together data, legal, compliance, cyber security, operational resilience, model risk management and technology perspectives in a unified governance structure – rather than assessing AI legal obligations in isolation.

Translating ethical principles into business strategies

Given the increasing regulatory scrutiny from all angles on the use of AI – and in particular GenAI – companies adopting AI without appropriate controls risk regulatory enforcement and litigation as well as significant reputational damage. Existing laws, rapidly evolving regulation and increasing public concern create a complex matrix for organisations to navigate with overlapping sets of rules that lack geographic and thematic consistency. However, the broad principles underpinning the various regimes have some common elements which need to be addressed.

Any business using AI should be mindful of the wider ethical implications of its use. This means taking a broad and forward-looking view of the likely implications of AI on the business, its employees, the environment and the customers with whom it interacts.

Taking a holistic, ethical approach to AI governance is key to being on the front foot for compliance with AI regulation, as the rules are currently mainly set at the level of over-arching principles rather than detailed technical standards. This is an ongoing responsibility to be considered at every stage in the process of technology adoption. At the design stage, building in compliance by design (a concept borrowed from the GDPR) should be a key objective. And once the tool is adopted, good ongoing monitoring procedures and clear communications with management, risk/compliance and customers are essential.

Adopting this approach should foster the trust that firms need from their end users, to ensure that the use of AI is a success. As discussed in Chapter 3, a variety of frameworks are developing which are built around key principles.

Taking the UK as an example, these are:

1. **Deliver fairness**: Firms need to deliver fairness when employing AI solutions. To do this, they need to make sure that AI-made decisions are justifiable with no bias and no discrimination. Demonstrating fairness will be key to avoiding liability.
2. **Demonstrate transparency**: explain how AI is being used: Firms need to demonstrate transparency, both about their use of AI and about how the decisions are being made. Considering how to deliver explainability, and what level of explainability you need to deliver to satisfy your management, customers and regulators, will be a key issue to address.
3. **Provide accountability and governance**: Firms need to consider how to allocate responsibility for maintaining appropriate oversight of the AI and keeping a 'human in the loop'. Doing that exercise of attributing responsibility up-front also really helps to ensure that fairness and transparency requirements are met.
4. **Anticipate contestability and redress**: Regulators consider that anyone suffering a harmful decision or outcome generated by AI should have the right to contest that decision and to seek appropriate redress – firms will need to establish processes to manage such claims.
5. **Safety, security, and resilience**: AI systems should be safe to use, robust and resilient in operation and secure in terms of cyber security. Technical standards may be needed, and regulations may require regular testing. Firms should validate initial findings and monitor products as they evolve.

For UK firms, emphasis is placed on maintaining an effective governance framework and ensuring technical skills development among employees, with regulators focused on the systemic risks that AI can introduce. The Alan Turing Institute has also been influential in setting the UK policy agenda.

Read more in Chapter 9 of our AI Toolkit.

IBM case studies

Case study 1

Fidelity Investments

Reaping ROI through repurposed use case implementation

Responsible AI initiatives are embedded into each phase of AI use cases at Fidelity Investments, beginning with robust data management practices and feeding into a dynamic review process driven by the company's AI Center of Excellence. The financial services firm invested heavily in these features to take full advantage of its foundational pillars – rather than as compliance box-checking exercises.

Each business line at Fidelity has a dedicated team for AI use case development and vendor management. This work is guided by the expertise of senior officers, who include representatives from the data governance and risk offices, who receive specialised AI training.

The AI Center of Excellence is involved in each stage of this process, from solution selection to model evaluation. Its leaders in Fidelity's data function and includes specialists from across the business in risk offices, information security, and compliance. In addition to legal and even internal communications, this process also allows Fidelity to consider adverse clients' incoming and outgoing information requests with heightened attention.

Resistance to responsible AI initiatives is alleviated when the rationale for controls is explained. More information relating to AI than to other controls are important, just to "remind them compliance policy, but there are implications to the firm if we get this wrong," said the senior business executive. Fidelity has also repurposed the provenance architecture previously adopted by the privacy office, which is integral to the success of AI projects and by streamlining the overall governance process.

> You have to explain that the reason controls are so important is not just some random compliance policy, but that there are implications to the firm if we get this wrong."
>
> A senior vice president with responsibility for data policy at Fidelity Investments.

Case study 2

Global financial services firm

Justifying positive returns with a lowered reputational risk

For one senior leader responsible for AI governance at a global financial services firm, AI development and ethics starts with education. He advocates hosting workshops that discuss ethical principles and values – empowering leadership to understand trade-offs. "We need to talk about AI in a way that interests leadership, not just in processes and procedures," he observes.

In discussing how to measure the return on investments in AI ethics, the senior leader offers the "creepy line" metaphor. Often, organisations find themselves in situations in which they are doing something perfectly legal that is highly profitable, yet still feel uncertain about the ethicality of their actions – a sense of crossing the "creepy line."

In such situations, he says that organisations must examine the activity through the lens of both current and future generations, in conjunction with the component of ethical considerations. As long as these considerations are covered satisfactorily, the organisation should feel reassured that the "creepy line" is not breached.

He also notes, "Reputational risk is a key factor in justifying positive returns. We aim to decrease reputational risk while achieving new AI and ethics principles."

For example, his team conducted an ethical fairness review of loan pricing involving a credit scoring algorithm. In conducting this review, the team analysed all 165 features of the model, asking if there were any potential causal mechanisms for why that particular data feature may correlate with an individual's ability to pay back a loan.

Ultimately, these data features were removed because a causal link did not exist, thus avoiding the lack of fairness in using this AI technology.

> We need to talk about AI in a way that interests leadership, not just in processes and procedures."
>
> Senior leader responsible for AI governance at a global financial services firm.

Source: Assessing ROI of AI ethics and governance initiatives | IBM

Tackling AI with risk management and governance

Agentic AI demands not merely technological deployment but a deep consideration of how work is structured, how decisions are made, and how humans and machines collaborate. Organisations must develop new governance frameworks that balance autonomy with accountability, speed with safety, innovation with reliability.

With the wave of AI regulation, we expect that regulators, policymakers, shareholders and the public will hold companies to account for failures to address AI risks *much faster* than they have for other risks with a 'technical' component because of the nature of the risks. An AI governance framework will help to: (i) promote responsible use of AI; (ii) help organisations identify and proactively manage risks; and (iii) ensure that AI is used in a way which is fit for purpose and proportionate to the risks.

In less than two years, generative AI, has moved from novelty to business necessity. With gen AI strategies shifting rapidly from exploring to focusing to expanding – 77% of business leaders in a recent IBM IBV study say they are convinced that gen AI is not only market-ready, but that quick adoption is necessary to maintain competitiveness.

What does good governance look like?

Today's AI landscape demands that organisations do more than simply adopt powerful tools. They must build robust governance and training structures, maintain thorough oversight, and engage in continuous risk management. Making sure that the organisation can explain how it uses AI and mitigates its risk is a legal necessity. Put simply, to develop trust in AI, effective governance is essential.

Trust factor 1

Accountability: Who is in charge of AI governance?

Effective AI governance must be a funded mandate from senior leadership.

Organisation-wide adoption requires reliable governance frameworks to mitigate risks and achieve business goals.

Trust factor 2

Transparency: How do you assess sources of data and what is shared about them?

Diverse and multidisciplinary teams should be deployed to assess data used to build models, matching the broad range of needs and expectations of AI users. This will answer questions about how models are audited and how they perform compared to humans.

Trust factor 3

Explainability: How do you explain the output of AI systems and models?

Deep collaboration between people and AI systems is built on transparency and explainability, adding the human touch to AI-informed decision-making. Sharing the provenance of models is fundamental to trust.

Clear AI governance practices and policies are at the core of addressing these trust factors. Without governance, the adoption of trustworthy and ethical AI systems can be inhibited. At the same time, open AI itself can help improve governance – all across the enterprise.

Source: The enterprise guide to AI governance | IBM

There is a further commercial imperative to implement a governance system that will scale with the increasing use and change of models. As LLM providers continue to release new and improved models, users need a means of measuring the changes in the performance, accuracy and tone of those models against a grounded set of approved outcomes.

Continuous collection of feedback on model outputs is key. The method behind the collection of validated model outputs by those humans in the loop is an important factor in good governance. Building validation methods into existing workflows can remove frictions to validation, which, in turn, allows those responsible for AI governance to check model 'accuracy' and identify and correct anomalies when they occur.

Guidance for Boards

1. **Don't trust it – you still need to check**: AI's accuracy has improved, but it still produces some plausible, yet strikingly incorrect answers, making it risky to rely on without human oversight.

Read more: UK – The LinksAI English law benchmark (Version 2).

2. **Data usage risks remain but can be mitigated if you build your own system**: To prevent the risk of disclosing confidential information, organisations should either develop internal AI systems or, when using externally hosted AI systems, implement thorough due diligence procedures and obtain contractual assurances about data usage.
3. **Regulation and reputation mean that organisations must take a degree of responsibility for outputs**: Organisations will need to take more responsibility for AI outputs, manage their AI processes, comply with emerging regulations and disclose when the output is AI-generated, while ensuring that failures are monitored and addressed.
4. **Be alert that some intellectual property issues are unresolved and data enforcement is increasing**: The use of copyrighted materials to train AI has led to significant litigation and uncertainties persist about the application of IP rights laws to AI. With data enforcement increasing, it is crucial that lawyers, technology and marketing teams collaborate to protect IP rights and avoid infringements.
5. **Be careful about bias and discrimination**: Ensuring data integrity is key to successful deployment of AI: As AI's use becomes more pervasive, so do the risks of bias and discrimination, which require close cooperation between technology, procurement and legal teams to address. Data provenance and integrity, along with proper training and testing, are essential for the successful adoption of AI.
6. **Consider outsourcing rules, supply chain and staff**: Focus is shifting from the establishment of terms and rules to ensure that the use of AI complies with procurement and outsourcing rules, to the longer-term impacts of increasing AI adoption on staffing requirements, and the balance between AI as assistant and AI as replacement for human resources. The near-term focus remains on training, enabling the organisation to fully benefit from the technology and ensuring that staff can effectively adapt to this new form of hybrid working.
7. **2025: Focus on compliance and managing regulatory risks**: Regulation now in place is compelling organisations at each level of the AI supply chain to adopt appropriate pro-active controls. AI use cases can materialise across the business and central functions, and specialists from impacted areas must be involved early on to assess risks for their area, using tools from legal, compliance and risk teams to avoid management logjams.

Additionally, regulations impose AI literacy on organisations, training staff on AI impacts, management being aware about AI risks and informed decisions being made. Boards must be informed about strategic AI issues and the associated legal, operational and ethical risks and mitigants.

Read more: Issues for Boards: Managing AI risk in 2025 - updated guidance.

NIST recommendations

As discussed in Chapter 2, the US NIST agency has produced the most comprehensive and holistic AI risk management framework to date. This builds on previous frameworks, for example in cyber, in providing practical guidance and can be scaled for the relevant organisation, use case and the level of risk associated with the AI system in question.

The NIST rules are summarised as follows: "*AI risk management is a key component of responsible development and use of AI systems. Responsible AI practices can help align the decisions about AI system design, development, and uses with intended aim and values. Core concepts in responsible AI emphasise human centricity, social responsibility, and sustainability*".

There are four core elements:

- **Govern**: Organisations need to cultivate and implement a risk management culture – prioritised by senior leadership – around the AI full product lifecycle (see below).
- **Map**: Organisations should understand the intended purpose and benefits of what the AI system is trying to achieve and map risks and impacts.
- **Measure**: Organisations should use quantitative and qualitative techniques to analyse and assess the risks of the system and how trustworthy it is.
- **Manage**: Identified risks should be managed on an ongoing basis with **priority given to higher-risk AI systems**. See suggestions in our AI Toolkit: periodically update incident response plans; ad hoc legal and compliance advice, incident postmortems and remediation.

Internal governance and controls: 10 steps to risk management

Regulation now in place is compelling organisations at each level of the AI supply chain to adopt appropriate pro-active controls.

Additionally, regulations impose AI literacy requirements on organisations, including training staff on AI impacts, ensuring that management understand AI risks and that informed decisions are made. Boards must be informed about strategic AI issues and the associated legal, operational and ethical risks and mitigants.

Large multi-national organisations need to also take into account various different guidance that has been issued by a number of key jurisdictions (as covered throughout this report). Having reviewed a range of international approaches to AI risk management, we have identified 10 suggested steps as a basis for establishing an effective framework for financial firms to consider when adopting any AI programme:

10 steps to risk management

1. Interaction of AI and ESG – Consider how AI fits into the firm's wider business values and the approach to corporate social responsibility and ESG disclosure requirements. Note that AI enables much deeper analysis – and scrutiny – of ESG performance.

2. Holistic approach to technology risk management – Consider your risk tolerance versus risk profile and update your technology risk and model risk management frameworks to document and govern AI usage and risk. Consider providing ethical guidelines for developers.

3. Specific governance structures – Consider whether an ethics board or other AI-specific board committee is needed to address AI risks if they are significant in the context of the business. Various roles to consider are ethics management, risk assessment, regulatory compliance management and policy analysts. Many organisations signal the importance to their organisations of AI adoption and governance through the appointment of Chief AI Officers, and similar positions, to the board.

4. AI and data governance policies – Consider whether the firm's data governance framework needs to be updated to address the potential risks of AI: this may involve embedding ethical guidance, revising data use and hygiene policies, introducing data deletion standards and implementing model behaviour constraints and ensuring some appropriate level of human oversight and review. Create AI guiding principles and set AI risk tolerance. Is the data used owned by the firm, or are necessary consents to use it in place? And are appropriate safeguards in place to protect data in a world of AI model integrations?

5. Training and development of employees – Deliver AI risk management training and ensure technical skill development by training employees how to use, manage and audit AI-based systems (depending on their role) and be aware of AI ethics.

6. Output labelling – Be open and transparent (internally and externally) about your use of AI, where possible – for example, establishing management updates and clearly labelling outputs as AI-generated.

7. AI incidents and resilience planning – Ensure there is a documented AI incident response plan for unintended consequences of the use of AI, taking into account any reporting/compliance obligations (eg, with respect to a data breach and operational/cyber resilience). We suggest conducting a tabletop exercise on AI incidents.

8. AI supply chain risk management – Address procurement and oversight of provision of AI tools and services throughout the lifecycle of any supplier engagement. Conduct thorough vendor due diligence and minimise gaps in compliance and risk between vendor and supplier by using appropriate contractual mechanisms.

9. Third-party AI audit and quality assurance – Consider engaging the services of a third-party AI or algorithmic audit and certification service, to vet compliance with internal policies and governance. Both internal and external audits and activity logs can help with ongoing monitoring by keeping a continuous log of Agentic activity. Read more: AI governance and quality assurance: Lessons from Linklaters and the audit sector (July 2025).

10. Appropriate Tooling – As models, assistants and agents proliferate across organisations, with the need for version control, retesting and continuous monitoring there is a need for robust tooling to manage these multi-model, multi-vendor environments. Organisations should be considering investing in governance platforms such as IBM's watsonx.governance to manage this complexity.

Business readiness – what financial services firms must do

The most pressing challenge is establishing robust internal governance and implementing infrastructure and risk management policies from top down, keeping in time and ideally ahead of the regulatory wave, while recognising that existing regulatory obligations already apply to these firms and will govern how they deploy AI. Many institutions face a disconnect between AI development teams and senior management. Proper AI governance requires buy-in at the highest level: where the use of AI is mandated, so too should be the means by which it will be used.

To ensure regulatory readiness and strategic alignment: C-suites must be briefed on AI risks and opportunities; multi-disciplinary governance forums (including tech, legal, risk) should oversee AI deployment and assurance and institutions should prepare for increased scrutiny and expectations around accountability. The effectiveness of internal governance will ultimately determine how well institutions navigate the evolving regulatory terrain.

Build robust AI governance frameworks under an executive mandate

01

Empower a senior-level executive to lead AI and data governance initiatives

Send the message that governance is a senior management priority. Championing AI and data governance from an enterprise's highest levels minimises the risk of failure due to fragmented ownership and fuzzy accountability.

02

Prioritise and build on responsible AI development and deployment.

Give leaders who will be accountable for AI governance the authority to do the work and provide their teams with the necessary resources to support this mandate. Teams can also build on and evolve governance frameworks already in place.

03

Ensure that senior leadership aligns principles with practices.

Align values related to the development and procurement of AI. Organisations achieve the outcomes they measure, and aligning principles with practices supports measurement of progress towards responsible AI adoption.

04

Develop a cultural foundation for governance structures.

Without a strong cultural foundation, AI governance structures cannot gain traction. Healthy cultures have success measurements, incentives, messaging and communications, diversity and inclusiveness, psychological safety, proactive employee training, and a holistic approach to AI literacy.

05

Foster collaboration with stakeholders and ecosystem partners

Include stakeholders across the entire organisation – all working toward the same goals. Collaboration with governments, trade and industry associations, and other groups helps establish AI governance guidelines, best practices, and regulations for responsible AI use. Ensure third-party software vendors and partners with embedded AI are subject to audits and other governance processes.

Compliance strategy – looking ahead

In financial services, given that many regulators take a technology-neutral approach to enforcing their rulebooks, firms need to continue to map their AI projects against existing law and regulation. And with the AI-specific regulation now in force in major markets like the EU, Mainland China and the US, there will be limited tolerance for firms that do not manage AI effectively.

This means that scanning for changes on the horizon has moved from being a nice-to-have to an essential. Getting the legal and regulatory structuring right up-front can make all the difference to avoiding not only legal consequences, but the potential for serious reputational and financial damage.

We are seeing the regulatory process accelerate in response to the rapid development of generative AI, but in a very fragmented away across the world. What will be interesting to see is how developments at national, regional and international levels play out, and what level of harmonisation across sectors, jurisdictions and regions can be achieved, when competing geopolitical objectives are at play.

Why Linklaters?

Linklaters has one of the largest global teams of AI lawyers. We are advising at the forefront of the rapidly evolving world of AI, advising the world's biggest tech companies, AI innovators, investors, and corporates across the globe. Our multi-disciplinary team excels in providing innovative and practical solutions for clients looking to harness the power of AI, from navigating a shifting legal landscape to enabling complex transactions and advising on contentious matters.

Our team assists in developing and executing robust AI strategies that maximise value while managing potential risks. Our clients benefit from best practice insights from our extensive experience around the world and deep relationships with regulators, particularly in the financial services sector. Our team has been advising on AI and machine learning-related matters since their inception – and our global tech practice is over 30 years old. Our client work and achievements, including a Financial Times Innovation award, demonstrate our market-leading expertise in AI.

We are embracing AI ourselves through a variety of tools, which bring a deeper understanding of AI and enhance our client service. With senior leadership championing our AI strategy, we are delving into exciting developments like agentic AI functions.

We take it upon ourselves to thoroughly test and explore different AI models so that we can share insights with clients based on a deep understanding of the technology (eg, through benchmarking of the capabilities of some of the top legal tech LLMs on the market (UK – The LinksAI English law benchmark (Version 2)).

As thought leaders we will continue to scan the horizon for AI legal and regulatory developments in the financial services, data and competition spheres, both at national and international levels. We bring our cross-disciplinary expertise to bear in advising clients on what steps they need to take as they navigate this rapidly evolving landscape.

Why IBM?

With over 100 years of experience in business and the technology industry with operations across 170 countries, IBM has established itself as a global leader and trusted partner for businesses in every sector. Its history is marked by continuous innovation, with numerous groundbreaking technologies and solutions that have revolutionised industries worldwide.

IBM has a history of technological firsts and continues to lead in investment and research into areas such as Hybrid Cloud, AI and quantum computing that will be critical to the future of all businesses, especially those in financial services. Global financial institutions trust IBM to design and run their operations.

IBM Consulting is the only global consultancy within a major tech leader, helping IBM to drive high-impact outcomes using advanced AI and a science-based approach to tackle the most critical business challenges. Partnering directly with clients, IBM advises, designs, builds and operates business innovation on IBM and partner technologies that matters and delivers results that last.

Please reach out to your usual Linklaters contacts or any of the contacts in this report to learn more.

Contacts

Linklaters

Hong Kong SAR

Kishore Bhindi
Financial Regulation Partner,
Hong Kong SAR
Tel: +852 2901 5066
kishore.bhindi@linklaters.com

Marcus Pollard
Antitrust and Foreign Investment
Partner, Hong Kong SAR
Tel: +852 2901 5121
marcus.pollard@linklaters.com

Yang Fan
TMT Managing Associate,
Hong Kong SAR
Tel: +852 2901 5006
yang.fan@linklaters.com

Mainland China

Arthur Peng
Antitrust & Foreign Investment
Partner, Shanghai
(Zhao Sheng Law Firm)
Tel: +86 10 6535 0651
arthur.peng@linklaterszs.com

Alex Roberts
TMT Partner, Shanghai
Tel: +86 21 2891 1842
alex.roberts@linklaters.com

Singapore

Peiying Chua
Financial Regulation Partner,
Singapore
Tel: +65 6692 5869
peiying.chua@linklaters.com

Adrian Fisher
TMT Partner, Singapore
Tel: +65 6692 5856
adrian.fisher@linklaters.com

Jia-Yi Tay
TMT Counsel, Singapore
Tel: +65 6321 5205
jiayi.tay@linklaters.com

Evan Lam
Financial Regulation Partner,
Singapore
Tel: +65 6321 5289
evan.lam@linklaters.com

EU

Sonia Cissé
TMT/IP Partner, Paris
Tel: +33 1 56 43 57 29
sonia.cisse@linklaters.com

Guillaume Couneson
TMT/IP Partner, Brussels
Tel: +32 2 501 93 05
guillaume.couneson@linklaters.com

Ceyhun Pehlivan
TMT Counsel, Madrid
Tel: +34 91 399 6182
ceyhun.pehlivan@linklaters.com

Marius Raetz
Germany Head of Fintech, Financial
Regulation Counsel, Frankfurt
Tel: +49 69 71003 777
marius.raetz@linklaters.com

UK

Jennifer Calver (Editor)
Global Fintech and Tech Sector Counsel, London
Tel: +44 20 7456 2417
jennifer.calver@linklaters.com

Peter Church
TMT/IP Counsel, London
Tel: +44 20 7456 5495
peter.church@linklaters.com

Julian Cunningham-Day
Global Co-head of Fintech, TMT/IP Partner, London
Tel: +44 20 7456 4048
julian.cunningham-day@linklaters.com

Jonathan Ford
Antitrust and Foreign Investment Partner, London
Tel: +44 20 7456 5295
jonathan.ford@linklaters.com

Richard Hay
Global Co-head of Fintech, Capital Markets Partner, London
Tel: +44 20 7456 2684
richard.hay@linklaters.com

Raza Naeem
Financial Regulation Partner, London
Tel: +44 20 7456 5272
raza.naeem@linklaters.com

Greg Palmer
TMT Partner, London
Tel: +44 20 7456 2925
greg.palmer@linklaters.com

Simon Treacy
Fintech and Financial Regulation Counsel, London
Tel: +44 20 7456 2451
simon.treacy@linklaters.com

US

Brad Caswell
Financial Regulation Partner, New York
Tel: +1 212 903 9362
brad.caswell@linklaters.com

John Eichlin
Antitrust and Foreign Investment Partner, New York
Tel: +1 212 903 9231
john.eichlin@linklaters.com

Kris Ekdhal
TMT Counsel, New York
Tel: +1 212 903 9415
kris.ekdahl@linklaters.com

Ieuan Jolly
TMT Partner, New York
Tel: +1 212 903 9574
ieuan.jolly@linklaters.com

Joshua Ashley Klayman
U.S. Head of Fintech and Head of Blockchain and Digital Assets, Capital Markets Senior Counsel, New York
Tel: +1 212 903 9047
joshua.klayman@linklaters.com

Caitlin Potratz Metcalf
TMT Senior Associate, Washington, D.C.
Tel: +1 202 654 9240
caitlin.metcalf@linklaters.com

Don Waack
Financial Regulation Partner, Washington
Tel: +1 202 654 9245
don.waack@linklaters.com

IBM Consulting

Jeremy Bassinder
Generative AI Consulting Leader

Rachel Lum
Partner and Capital Markets Industry Leader

Contributors

Linklaters

Hong Kong SAR

Kathleen Gooi
Antitrust & Foreign Investment
Managing Associate, Hong Kong SAR
Tel: +852 2901 5375
kathleen.gooi@linklaters.com

Clara Hackney
Financial Regulation Senior Associate,
Hong Kong SAR
Tel: +852 2901 5611
clara.hackney@linklaters.com

Jackie Lau
Antitrust and Foreign investment
Associate, Hong Kong SAR
Tel: +852 2901 5198
jackie.lau@linklaters.com

Jasmine Yung
TMT Associate, Hong Kong SAR
Tel: +852 2901 5293
jasmine.yung@linklaters.com

Mainland China

Tiantian Ke
TMT Associate, Shanghai
(Zhao Sheng Law Firm)
Tel: +86 21 2891 1881
tiantian.ke@linklaterszs.com

Ruoyi Lu
TMT Associate, Shanghai
Tel: +86 21 2891 1989
ruoyi.lu@linklaters.com

Ningying Bao
Corporate Managing Associate,
Shanghai (Zhao Sheng Law Firm)
Tel: +86 21 2891 1824
ningying.bao@linklaterszs.com

Jennifer Wu
Corporate Associate, Shanghai
(Zhao Sheng Law Firm)
Tel: +86 21 2891 1849
jennifer.wu@linklaters.com

Singapore

Elizabeth Webb
Financial Regulation Associate, Singapore
Tel: +65 6692 5718
elizabeth.webb@linklaters.com

Ee Ning Low
TMT Associate, Singapore
Tel: +65 6692 5754
eening.low@linklaters.com

US

Bhavishya Barbhaya
Antitrust & Foreign Investment Associate, Washington, D.C.
Tel: +1 212 903 9015
bhavishya.barbhaya@linklaters.com

Toby Irenshtain
TMT Associate, New York
Tel: +1 212 903 9017
toby.irenshtain@linklaters.com

UK

Kim Rust
Antitrust & Foreign Investment Associate, London
Tel: +44 20 7456 4721
kim.rust@linklaters.com

Elton Qemali
Fintech Paralegal, London
Tel: +44 20 7456 3888
elton.qemali@linklaters.com

Abu Dhabi | Amsterdam | Antwerp | Bangkok | Beijing | Berlin | Brisbane* | Brussels | Cape Town*** | Dubai | Dublin
Düsseldorf | Frankfurt | Hamburg | Hanoi* | Ho Chi Minh City* | Hong Kong SAR | Jakarta** | Johannesburg***
Lisbon | London | Luxembourg | Madrid | Melbourne* | Milan | Munich | New York | Paris | Perth* | Port Moresby*
Riyadh | Rome | São Paulo | Seoul | Shanghai$^{\Delta}$ | Singapore | Stockholm | Sydney* | Tokyo | Warsaw | Washington, D.C.

** Office of integrated alliance partner Allens*
*** Office of formally associated firm Widyawan & Partners*
**** Office of collaborative alliance partner Webber Wentzel*
$^{\Delta}$ Linklaters Shanghai and Linklaters Zhao Sheng (joint operation office with Zhao Sheng Law Firm)

linklaters.com

This content is intended merely to highlight issues and not to be comprehensive, nor to provide legal advice. Should you have any questions on issues reported here, please get in touch. © 2026 Linklaters.

Linklaters LLP is a limited liability partnership registered in England and Wales with registered number OC326345. It is a law firm authorised and regulated by the Solicitors Regulation Authority. The term partner in relation to Linklaters LLP is used to refer to a member of the LLP or an employee or consultant of Linklaters LLP or any of its affiliated firms or entities with equivalent standing and qualifications. A list of the names of the members of Linklaters LLP and of the non-members who are designated as partners and their professional qualifications is open to inspection at its registered office, 20 Ropemaker Street, London, EC2Y 9AR, England or on www.linklaters.com and such persons are either solicitors or registered foreign lawyers. Please refer to www.linklaters.com/regulatory for important information on our regulatory position.

CD701_F/09.26