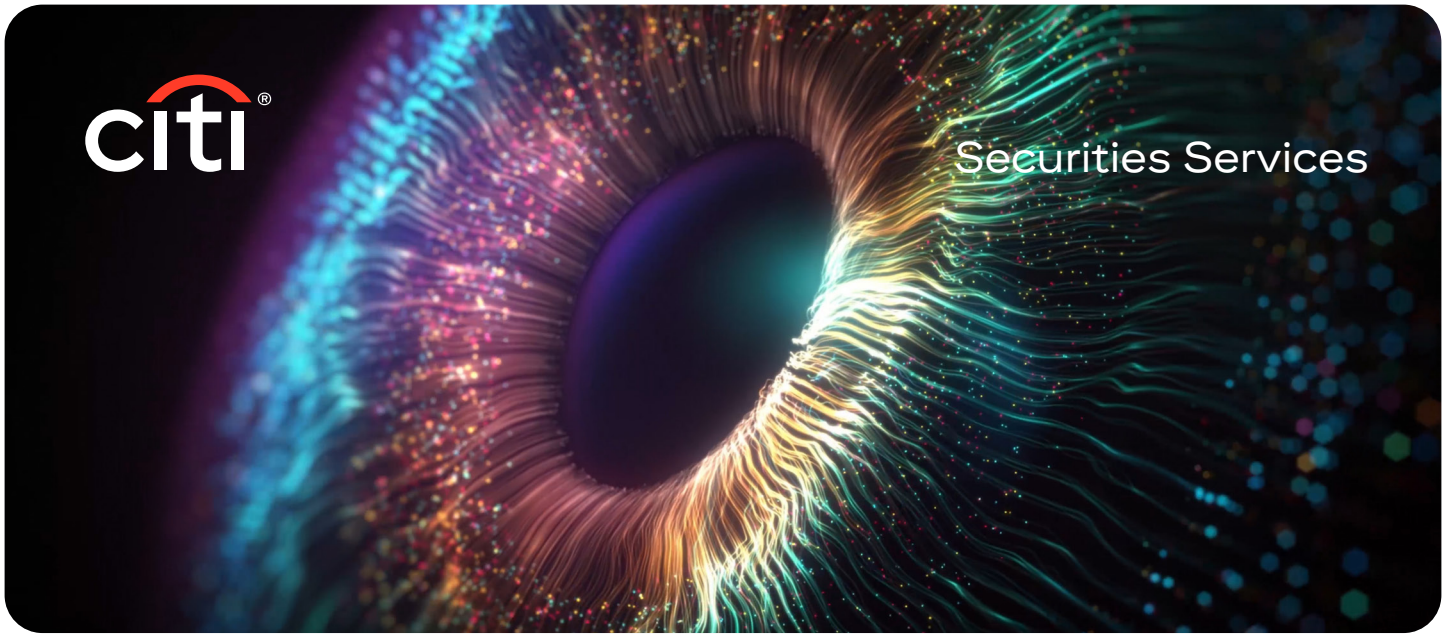


Doing the Right Thing for Business and Society - Generative AI

In recent years, Artificial Intelligence (AI) and Generative AI (Gen AI) have transitioned from the realm of science fiction to reality.

The concept of intelligent machines can be traced back centuries. But true AI development really started with Alan Turing's Test for computer intelligence[1] and neural networks[2] in the 1950s, followed by the development of AI programming languages and microworld programs[3] during the 60s.

The following decade witnessed the establishment of fuzzy logic[4] in the 70s, and the exploration of symbolic[5] and nouvelle[6] approaches in the 80s. The 90s brought the Internet and increasing computation power which enabled significant advances in machine learning and evolutionary algorithms.

Big data, deep learning, autonomous agents, and language models have led to AI's ubiquitous entry into everyday life.

But that is all history, fast forward to 2024 - where we are now familiar with many Gen AI powered tools that are on the market with capabilities to generate text, write code, create audio, and make images and video on-demand.

Large language models (LLMs) and enhancements in computing power have unlocked new capabilities, allowing us to create content, analyse text, and converse with machines like never before. Such models have been credited by some as kick-starting the current AI boom[7], leading to ongoing rapid investment in, and public attention to, the field of AI.

Whilst society appears to be assimilating the use of such tools into everyday life, this is not necessarily the case for financial services firms, as there are certain regulatory and ethical guardrails to be honoured.

What is AI and Gen AI and why does it matter?

Defining AI is often a challenge due to AI's broad scope, quick advancements in the field, and the type of AI in question. In its most basic form, AI is the simulation of human intelligence by machines. More formally, Citi defines Artificial Intelligence and other related concepts to ensure there is consistency in understanding across the organisation.

Citi's definitions

'**Artificial Intelligence**' refers to a quantitative method, system, or approach ('techniques') that emulates human intelligence via computer programs to generate content (textual, audio, video, or images) and/or make estimates, predictions, recommendations, or decisions in manners that go beyond classical statistical, mathematical, econometric, or financial approaches.

'**Generative AI (Gen AI)**', being a subset of AI, which is a type of Artificial Intelligence that creates new content, such as images, text, audio, video, code based on their training data and in response to prompts.

Generative AI systems are designed to generate new and original content that is not explicitly programmed, but rather generated based on patterns and information learned from existing data.

Why it matters

During a recent podcast 'Good Things Happen - Artificial Intelligence: What is it and why it matters' Dr. Prag Sharma, Global Head, AI Centre of Excellence, Citigroup, discussed his stance on the use of Gen AI in business. He talked about the responsible use of Generative AI at Citi, highlighting the potential of this technology along with the risks that the current version of LLMs pose.

Separately, Dr. Sharma provided his thoughts on Citi's use of Gen AI.

> Citi has been using AI for a number of years across the bank. The recent emergence of Gen AI has turbocharged these efforts. We are taking Gen AI from the lab to the factory floor – harnessing its power and scaling its use across the firm. We remain on the front foot – empowering our businesses to deploy the technology in areas that will move their businesses forward.
>
> Our work with Gen AI is aimed at amplifying the power of our people and is guided by our commitment to transform and strengthen Citi. As we carefully scale and operationalize Gen AI across the firm, we're also staying vigilant in managing the risks associated with it and ensuring that we align with the evolving regulatory landscape.
>
> Dr. Prag Sharma, Global Head, AI Centre of Excellence, Citigroup

The evolving regulatory lens for AI

In Dr. Sharma's quote he referred to the 'evolving regulatory landscape'. Rapid developments in AI resulted in significant policy attention and action, which according to the Organisation for Economic Co-operation and Development (OECD) are evidenced by more than a thousand AI initiatives in over 70 countries and jurisdictions.[8]

Global regulatory aims and objectives on the use of Gen AI (Figure 1):

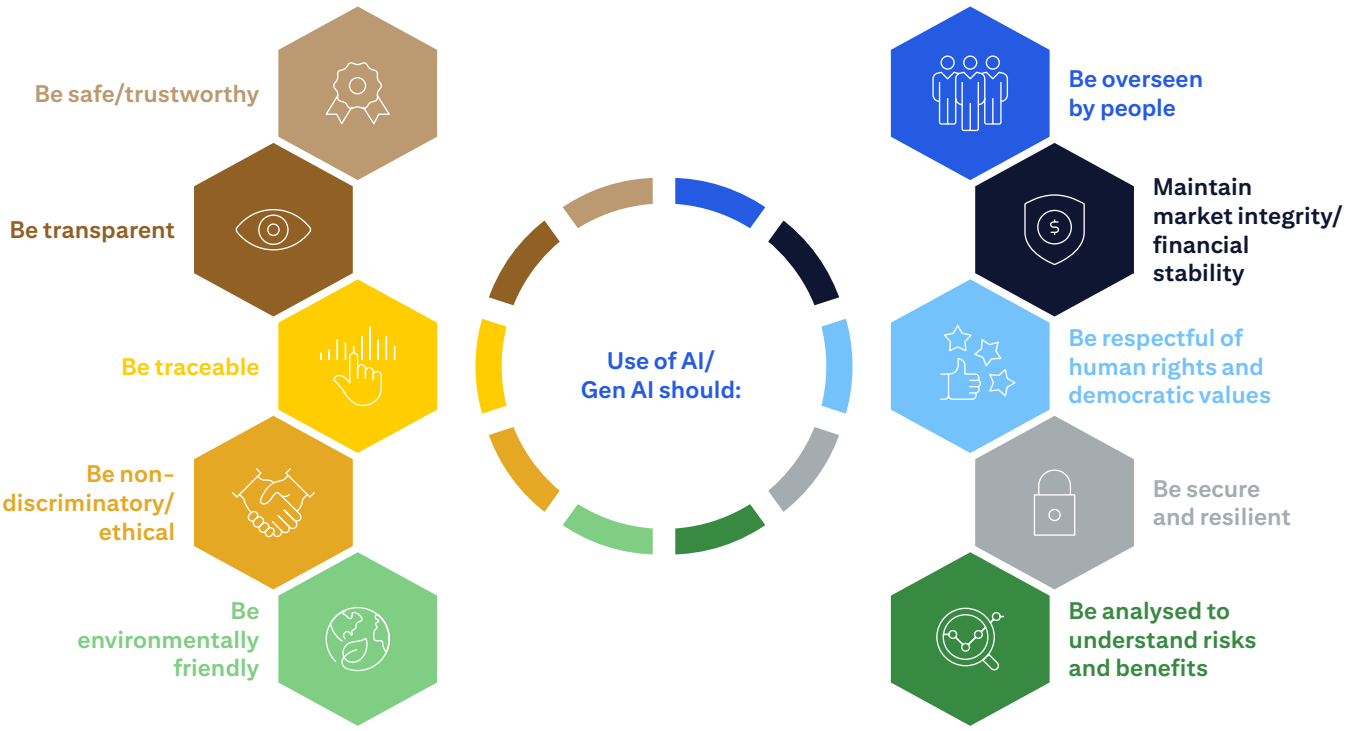


Figure 1 demonstrates consistency in shared themes and expected guardrails on a global level; however, individual countries are taking nuanced approaches to implementation.

To learn more about how different countries are approaching regulatory aims and objectives please refer to the Appendix 'Latest regulatory commentary on AI/Gen AI' which can be found at the end of this article.

The potential of Gen AI

According to Citi's GPS Report 'Disruptive Innovations X - Ten More Things to Stop and Think About' Gen AI has revolutionary potential across financial services. In many respects, finance is the perfect sector for the application of Gen AI.

The sector is data-rich with structured, semi-structured, and unstructured raw materials. For many years, some forms of AI including machine learning, deep learning, and natural language processing have been used in finance on critical use cases such as fraud detection, reconciliation, and risk management.

Most enterprise data (about 80%-90%) is unstructured - this means the data is locked away in emails, transcripts, documents, and reports.

Tod McKenna, Global Head of Data Science and AI for Securities Services, explains that:

> "Gen AI will expand the traditional AI use case portfolio deeper into the realm of semi-structured and unstructured data.
>
> These use cases will cover question-answering, text extraction and analysis, and far more advanced and less annoying chatbot experiences."

McKenna further notes that:

> "A significant amount of data quality problems occur at the point unstructured data is turned into structured data to fit traditional database structures. Transcription errors, transposition errors, formatting problems, data truncation, and other fat-finger mistakes can be mitigated with Gen AI. There's huge potential."

AI/Gen AI - Areas of potential:

Unstructured Data 	• Gen AI has potential to expand the use cases of ML and AI into unstructured data to change customer service, sales and marketing, content generation and coding.
Autonomous Agents 	• AI assistants could reduce headcount in customer support functions 20%-30% by 2026.
Automation 	• Two-thirds of work in banking and insurance could be automated or augmented by AI.

Source: Citi GPS 'Disruptive Innovations X - Ten More Things to Stop and Think About'

Gen AI use-cases

Insurance firms are abundantly rich in terms of data sources, making them a natural fit for Gen AI use. In areas such as actuarial risk modelling, financial reporting, and regulatory compliance.

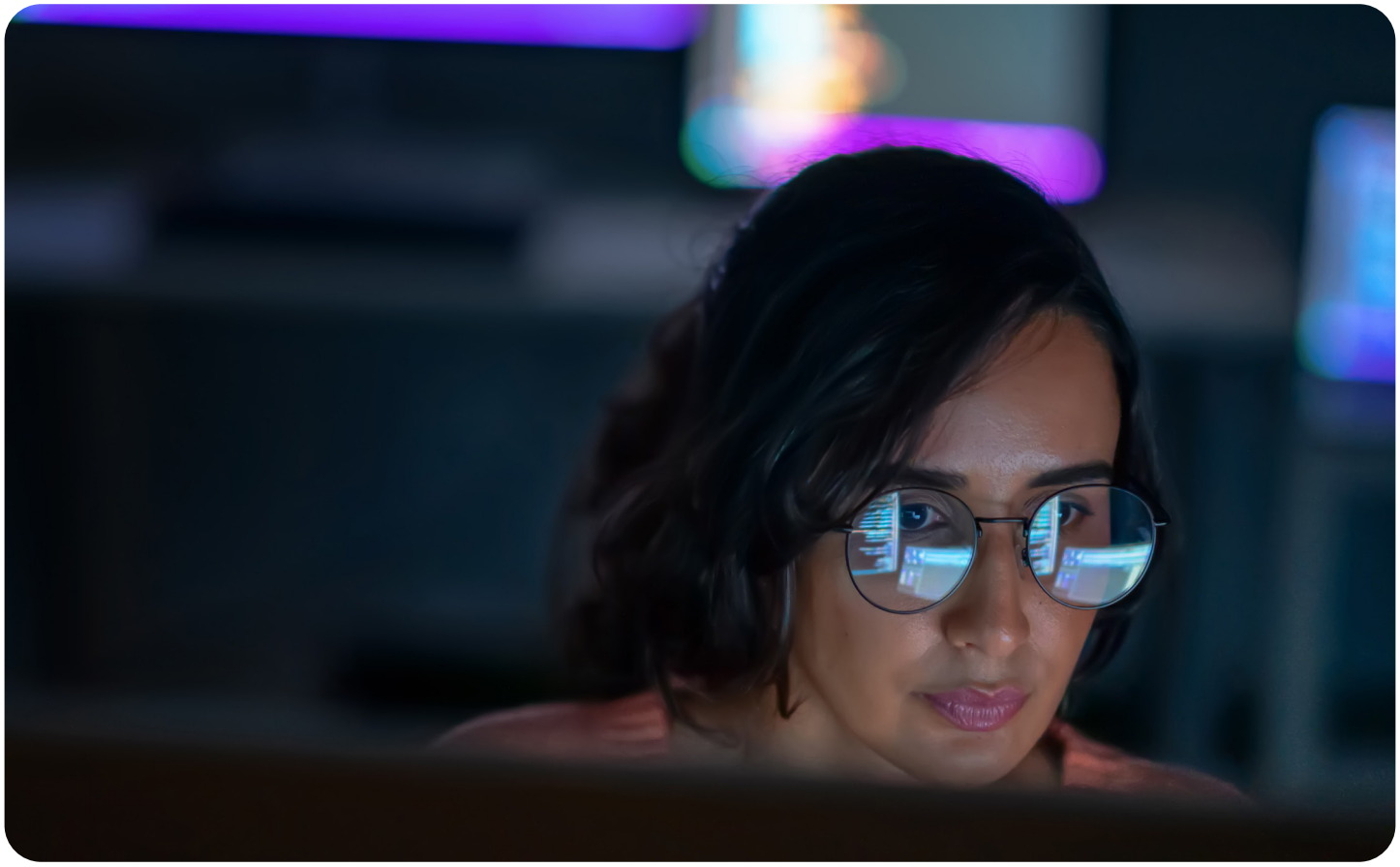
With continuing pressure to cut costs, diversify revenues, and introduce more personalized services, asset managers are also seeing the benefits Gen AI can achieve. For example, improving operational efficiency, reducing key person risks, commoditizing, investment research, optimizing middle-and back-office functions, supporting sales, and improving client services.

In terms of understanding AI use in asset management, in June 2024, Citi's Global Insights | Business Advisory Services (BAS) published a report 'AI in Investment Management - The Pursuit of a Competitive Edge'. The report looks at how Gen AI is being deployed, what use-cases are being considered, and how the future state may look.

BAS conducted over 40 interviews with the C-suites of investment management firms, ranging from hedge funds and sovereign wealth funds to asset managers across different regions of the world, together with industry experts and FinTech's.

The team identified a variety of views, ranging from strong AI enthusiasts to those with a more sceptical point of view. Below is a summary of the high-level findings identified, key messages, and a selection of asset management quotes:

- Creating competitive advantages with AI has become a strategic priority, with Gen AI potentially representing the latest inflection point in the evolution of AI.
- AI adoption in investment management is still in its early stages, with applications being used across the organization under human supervision - 'human in the loop' is still critical.
- AI use cases are in the areas of investment, distribution, and operations with a three-pronged approach adopted: 1) efficiency and productivity gains as the first step; 2) investment research and client engagement; 3) alpha generation, investment co-pilot, automated creation, and management of funds as long-term considerations.
- Current use cases focus primarily on efficiency and productivity gains using both traditional AI and increasingly Gen AI.
- A variety of challenges are hindering AI adoption: culture; resources; performance; governance; infrastructure and timing.
- Different archetypes of asset managers have varying pathways to achieve future state based on their unique goals and value propositions - but change management is a focus for all, with many interviewees highlighting culture change as a key challenge.
- Investment managers generally prefer to build AI solutions in house where they can differentiate by using internal proprietary data and or intellectual property but are also open-minded about collaborating with FinTech providers.
- Whilst using AI can improve productivity and reduce costs, differentiation likely comes from proprietary data and orchestration of internal capabilities and strengths when combined with various AI offerings.
- Data is at the core of the future state of AI in investment management. A future cohesive framework will likely involve both a centralized client insight platform and an AI-augmented data application hub to enable personalization at scale.



A selection of asset management quotes provided as part of the BAS interview process

Quote one

> We staged [our AI buildout] across 3 different phases. The 1st phase (next 1-2 years) is the productivity phase, lowering the costs without much computational advantage. The next phase, which is what we see amongst counterparties, is having more domain-specific models - models trained and finetuned across investment cases. It tailors to the lower-level investment needs. The last phase is to be more alpha driven. It really comes with having a more explicit incorporation of models into your investment process.
> - **Sovereign Wealth Fund.**

Quote two

> There are huge gains in every department to be made [from AI] for an Asset Management Company like us with a huge footprint. So, we're doing it, bit by bit; one step at a time. Small wins first and, save the low hanging fruits. First, create the belief that AI actually works and go forward.
> - **Investment Manager AUM <$500 billion.**

Quote three

> We have used AI techniques for sentiment input into our model for longer than 12 months...more recent AI approaches to sentiment classification, analyzing patents and classifying firms into a thematic categorization and the use of NLP on reporting text to classify the firms in different ways.
> - **Investment Manager AUM > $1 trillion.**

Quote four

> I don't think AI is the silver bullet for the industry. [Using AI won't suddenly make a firm] the sole winner of all asset classes because they have some GPT models.
> - **Investment Management AUM > $1 trillion.**

Quote five

> Our biggest concern is being able to have control over the leakage of proprietary information to open models.
> - **Investment Manager AUM > $1 trillion.**

Quote six

> We are not going to build our own LLM. We are going to buy or rent as much of the infrastructures, frameworks, and tooling as we can, with the caveat of having a strict sense of human supervision and a control framework.
> - **Investment Manager AUM > $1 trillion.**

Quote seven

> We have set ourselves targets over the next quarter or two quarters, we want to realize X million dollars of efficiency gain using Gen AI.
> - **Investment Manager AUM ≤ $500 billion.**

If you would like to discuss any of the findings from the AI in Investment Management Report summarised in this section in more detail, please reach out directly to the Business Advisory Services Team business.advisory@citi.com.

Key areas to consider

Having looked at these industry views, what are the key areas that require greatest attention?

The risks

Firms should consider what risks Gen AI can introduce. Some examples include:

- *Data Privacy Risk*: The unauthorized collection, sharing, or use of sensitive information;
- *Security Risk*: Potential vulnerabilities may make Gen AI systems targets for cyberattacks;
- *Intellectual Property Risk*: Uncertainties about ownership and rights to content created by AI;
- *Hallucination Risk*: Generative models may create incorrect, fabricated, or misleading output;
- *Bias Risk*: Amplification of bias in training the data, leading to unfair or discriminatory output;
- *Skill Gap*: Potential lack of sufficient or limited expertise to build responsible and ethical AI; and
- *Regulatory Risk*: Failure to keep up to date with rapidly evolving regulatory frameworks, particularly important given the introduction of the EU AI Act with its extraterritorial impacts.

These, amongst other risks, should be managed through mitigation techniques such as avoidance, controls, or risk transfer.

Upskilling humans

Upskilling humans, alongside developing the use of Gen AI is also important.

In Citi's GPS Report 'What Machines Can't Master - Human Skills to Thrive in the Age of AI' (the Human Skills Report), it looks at which human skills will come to the fore as AI advances.

The Human Skills Report is designed to answer that question and acknowledges that increased awareness and debate is only part of the story. Many experts in Chapter 2 of the Human Skills Report note, "we need action and execution. AI is not waiting for anyone. The race between AI's progress and human's ability to adapt has already started."

Technological developments

Gen AI could also be poised for another boost - quantum computing. As detailed in Citi's GPS Report 'Quantum Computing - Moving Quickly From Theory to Reality' (Quantum Report), quantum computing breaks away from classical computing by using quantum mechanics[9] as its base.

The Quantum Report also explains that computing is at an inflection point. After five decades of increasing computational power in line with Moore's Law (i.e., doubling the number of transistors per integrated circuit every two years), the growth rate of classical computing is reaching its physical limit. At the same time, the capabilities of quantum computers are advancing at a "doubly exponential" rate - in line with Neven's Law.[10]

As quantum computing has the potential to solve practical problems faster or more effectively than classical computers, one of the industries that the Quantum Report explains will likely benefit is AI (which includes Gen AI).

This area is also of interest to regulators, where as recently as 18 July 2024, the Monetary Authority of Singapore (MAS) announced that it will commit an additional SGD 100 million under the Financial Sector Technology and Innovation Grant Scheme (FSTI 3.0) to support financial institutions in building capabilities in quantum and AI technologies, as well as enabling the advancement of quantum and AI related innovation and adoption in financial services.

MAS is also collaborating with banks and technology partners on quantum security,[11] by signing a Memorandum of Understanding (MoU) to embark on quantum security collaboration and study the application of Quantum Key Distribution (QKD)[12] in financial services.

A mixed blessing

Managing the risks of introducing Gen AI into your business is achievable - starting with trusted AI aims and objectives (see Figure 1), guidelines for the ethical development of Gen AI,[13] and a good knowledge of the latest regulatory frameworks. The latest framework being the Council of Europe's convention on AI which was signed on 5 September 2024. The first signatories were the U.S., EU, and the UK, with others expected to follow.

But it's important to remember that whilst Gen AI models can provide many benefits across financial services, the models also need constant supervision and oversight. They should also aim to help humans be better, ideally as a partnership designed to do 'the right thing for business and society'.

If you haven't yet entered the debate, started to consider how Gen AI can be incorporated safely and ethically into your business, or are simply looking to upskill, you can start by reading the series of Citi GPS reports covering various digital innovations, along with the other cited Citi publications referred to in this article - as it's never too early to engage where Gen AI or other types of financial innovation are concerned.

Contributors

Citi GPS
Citi Business Advisory Services Team
Matthew Cherrill
Jennifer English
Amanda Hale
Tod McKenna
Andrew Newson
Dr. Prag Sharma

This article is accurate as at the time of writing: 9 October 2024

Appendix – Latest Regulatory Commentary on AI/Gen AI

International

G7

The International Guiding Principles on Artificial Intelligence were developed under the "Hiroshima Process" which was established at the G7 Summit in May 2023, as part of coordinated efforts by a number of international bodies (including the OECD and the EU-U.S. Trade and Technology Council) to establish global guardrails and standards for AI.

The accompanying International Code of Conduct for Advanced AI Systems contains a set of rules that AI developers are encouraged to follow, on a voluntary basis, to mitigate risks throughout the AI lifecycle.

IOSCO

In September 2021 the International Organization of Securities Commissions (IOSCO) published its final report 'The use of artificial intelligence and machine learning by market intermediaries and asset managers' (final report).

In the final report IOSCO identified that the use of technologies such as AI, could create efficiencies and benefits for firms, but also amplify certain risks with the potential to cause consumer harm.

Rolling forward a few years to 2024, IOSCO published an updated Workplan. In terms of new topics in the Workplan, IOSCO said that it would turn its attention to key new technological developments in financial markets such as the use of AI.

IOSCO's AI workstream will consist of a two-year policy initiative, aimed at ensuring the development of a shared understanding among IOSCO members on the issues, risks and challenges presented by emerging AI technology through the lens of market integrity, financial stability, and investor protection. It also aims to assist IOSCO members in their policy responses - so is intended to be directional for regulators globally.

Following publication of the Workplan, IOSCO have now launched a new AI survey (Investment Association membership required to access) to assess the use and role of AI in the financial services industry. The survey will support IOSCO's new AI workstream and consists of close-ended questions focused on two key areas:

(1) Where firms/organizations are using AI in relation to certain products and services, and

(2) The material risks associated with AI, with responses ranging from Strongly Disagree to Strongly Agree.

The Survey closed on 13 September 2024.

FSB

On 11 July 2024, the Chair of the Financial Stability Board (the FSB) delivered a speech at the IMF/World Bank Constituency meeting 'The AI adventure: how artificial intelligence may shape the economy and the financial system'.

During this speech, the Chair covered Gen AI from different angles, starting with the impact that Gen AI may have in the future, plus how it could shake up labour markets.

In terms of regulation, the Chair noted that there are open questions concerning AI. Further saying it can be a technology that can be used for benefit if the right policies and regulations are implemented.

This year, the Chair advised that the FSB will be updating its paper on the financial stability implications of artificial intelligence, originally published in 2017. Early emerging consensus is that the risks identified in the earlier report are still there, with the most important ones being concentration risk; third-party risks; increases in herding behaviour; and model risk - including challenges with regards to explainability.

Potential new forms of interconnectedness in the financial system could emerge. For example, the Chair explained that autonomous trading agents may interact to create new dynamics in financial markets, with some studies finding that AI-powered algorithms consistently learn to charge higher prices through collusive strategies, even if there is no direct communication between them.

Such interdependencies may be pronounced if the market for data and model providers is highly concentrated. This is the case for generative AI models, because although there are lots of applications out there, in practice they are based on only a handful of models, just three or four.

The Chair concluded by saying that at this stage, the FSB's work is purely analytical. The FSB is not currently developing policy options or coordinating across standard setting bodies or international organisations. But the FSB is ready to do what's required, to monitor these risks and implement effective regulatory frameworks.

OECD

The OECD AI Principles initially adopted in 2019, were amended in 2023 to update the definition of an "AI System", also clarifying that the definition includes Gen AI.

They were updated again in May 2024, this time in consideration of new technological and policy developments to address AI-associated challenges involving privacy, intellectual property rights, safety, and information integrity.

United States

On 13 September 2023, the U.S. Senate held an inaugural bipartisan 'AI Insight Forum' with an aim to familiarize senators with the nature of AI and its associated risks and to discuss required safeguards and legislation.[14] These forums are ongoing.

Then on 30 October 2023, U.S. President Biden signed the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.

In relation to supervisory measures, it was in March 2024 (in a press release), that the Securities and Exchange Commission (SEC) announced it was bringing charges against two investment advisers for making false and misleading statements about their use of AI ("AI Washing"). This followed a joint Investor Alert issued on 25 January, 2024 by the SEC, the North American Securities Administrators Association (NASAA), and the Financial Industry Regulatory Authority (FINRA) intended to make investors aware of the increase of investment frauds involving the purported

use of AI and other emerging technologies.[15] SEC Enforcement Director Gurbir Grewal has characterized the use of AI and AI Washing as areas of potential risk and enforcement activity in a 15 April 2024 speech and said that "proactive compliance" is a firm's best defence.[16]

In a notice to member firms dated 27 June 2024, FINRA outlined regulatory obligations when using generative AI tools and LLMs. Including the adoption of policies and procedures that address technology governance, including model risk management, data privacy and integrity, reliability, and accuracy of the AI model, among other things.[17]

Then more recently, during August 2024, a bipartisan group of lawmakers introduced a bill in both the House and Senate that would allow financial federal agencies to experiment with AI. The Bill 'The Unleasing AI Innovation in Financial Services Act' is designed to foster innovation by introducing a safe environment where new financial products and services making use of AI can be tested, without expectation of enforcement action.

At the state level, Colorado signed the Colorado Artificial Intelligence Act (SB 205) into law on 17 May 2024, making Colorado the first state to enact an AI law. The bill establishes standards and requirements for those who develop or deploy AI systems to protect against algorithmic discrimination, beginning 1 February 2026. This includes risk management strategies, impact assessments, and annual reviews.[18]

Draft legislation in California would require developers of AI models trained using a specified level of computing power ("covered model") to meet certain safety requirements before training and releasing the model and establish other oversight and protective mechanisms.[19]

Finally, on 23 September 2024, the U.S. Department of Justice (DOJ) updated its Evaluation of Corporate Compliance Programs to reflect DOJ's evolving expectations with respect to corporate compliance programs, including how those programs appropriately address the compliance risks of new technology such as AI.

With the DOJ's most recent update to this document, this tool now reflects DOJ's focus on disruptive technology risks.

Europe - the EU AI Act

In April 2021, the European Commission proposed the first EU regulatory framework for AI (referred to as the EU AI Act), so it warrants a callout in this article.

The new rules establish obligations for providers and users depending on the level of risk from AI.[20] These risks are further broken down into either unacceptable risk or high-risk. As regards associated transparency requirements, Generative AI, like ChatGPT/OpenAI, will not be classified as high-risk, but will have to comply with transparency requirements and EU copyright law which requires:

- Disclosing content generated by AI.
- Designing the model to prevent it from generating illegal content.
- Publishing summaries of copyrighted data used for training.

High-impact general purpose AI models that might pose systemic risk, such as the more advanced AI model GPT-4, would have to undergo thorough evaluations and any serious incidents would require reporting to the European Commission.

Whilst content that is either generated or modified with the help of AI such as images, audio, or video files (for example deepfakes[21]) - need to be labelled 'clearly' as AI generated so that users are aware when they come across this type of content.

With the introduction of the EU AI Act, a regulatory definition has also been introduced by the European Commission. The AI Act defines an AI system 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.'

The Regulation also provides a definition of 'general-purpose AI system' meaning an AI system which is based on a general-purpose AI model, and which has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems.'

The EU AI Act came into force on 1 August 2024, but it will be phased in. Bans on AI systems with unacceptable risk will take effect after six months, rules for general purpose AI models will apply after 12 months, and high-risk AI systems embedded in regulated products will have a 36-month compliance period.

Also in the EU, on 30 May 2024, the European Securities and Markets Authority (ESMA) issued a Statement providing initial guidance to firms using AI when they provide investment services to retail clients.

When using AI, ESMA expects firms to comply with relevant MiFID II requirements, particularly when it comes to organisational aspects, conduct of business, and their regulatory obligation to act in the best interest of the client.

ESMA states that potential uses of AI by an investment firm which would be covered by requirements under MiFID II include customer support, fraud detection, risk management, compliance, and support to firms in the provision of investment advice and portfolio management.

ESMA and the National Competent Authorities confirmed that they will keep monitoring the use of AI in investment services and the relevant EU legal framework to determine if further action is needed in this area.

On 18 June 2024, the European Commission sought input from industry on the use of AI in finance. The initiatives cover use cases, benefits, barriers, risks, and stakeholder needs.

The consultation and associated workshops target all financial stakeholders and any input received will enable the European Commission to provide guidance to the financial sector for the implementation of the EU AI Act.

Finally, on 25 September 2024, the European Commission announced that over 100 companies have made up the first signatories of the EU AI Pact and its voluntary pledges.

Turkey

On 26 June 2024, the Artificial Intelligence Bill No. 2/2234 was submitted to the Parliamentary Committee, where it is currently in commission. The bill specifically aims to ensure the safe, ethical, and fair use of AI technologies, protect personal data, prevent privacy rights violations, and establish a regulatory framework for the development and use of AI systems.

The bill seeks to maximize the benefits AI systems can provide to society. By ensuring that AI systems operate securely, transparently, and accountably, it aims to increase public trust in these technologies and accelerate the adoption of AI applications.

Progress on the bill can be tracked here.

Asia Pacific

Regulation of AI technology in the Asia Pacific region is developing quickly.

Singapore

The Singapore government has developed various frameworks and tools to guide AI deployment and promote the responsible use of AI. In 2022 for example, Singapore launched AI Verify which is an open source AI governance testing framework and software toolkit that validates the performance of AI systems against a set of eleven internationally recognised AI ethics and governance principles through standardised tests.

In November 2023, the Monetary Authority of Singapore (MAS) announced that it would partner industry to develop a Generative AI Risk Framework for the Financial Sector. Whilst the Gen AI Framework applies to banks, the press release indicates that the scope is likely to be expanded to cover insurance and asset management firms and that an AI governance handbook for the financial services industry will be developed.

Then in December 2023, Singapore unveiled version 2.0 of its (2019) national AI strategy with goals aimed at developing peaks of excellence in AI, to advance the field, and maximise value creation.

However, more recently in January 2024, Singapore's Infocomm and Media Development Authority published a Model AI Governance Framework for Generative AI (Gen AI Framework), with a more comprehensive approach aimed at addressing new risks from the use of generative AI.

The Gen AI Framework provides a governance framework to guide organisations in their development or deployment of Gen AI. It makes recommendations which may require AI developers and companies to shoulder more responsibility around Gen AI, proposing goals around safety, accountability, transparency, and security.

Hong Kong

In November 2019, the Hong Kong Monetary Authority (HKMA) published a circular on 'High-level Principles on Artificial Intelligence' covering three aspects of AI technologies - governance, application design and development, and ongoing monitoring and maintenance. So as not to stifle innovation in the development of AI-related technologies, the principles are formulated to be high-level in nature and banks are only expected to apply them in a manner proportionate to the nature of their AI applications and the level of associated risks.

In the same month, the HKMA also issued a set of guiding principles on consumer protection aspects in the use of AI applications (2019 BDAI Guiding Principles to the use of AI). The principles are centred around four key areas: governance and accountability, fairness, transparency and disclosure, and data privacy and protection. The HKMA reminds authorized institutions to adopt a risk-based approach commensurate with the risks involved in their use of AI applications when employing these principles.

Whilst the Securities and Futures Commission (SFC) is aware of the development of AI and employs AI-assisted technology in its operations, it has yet to issue any specific guidance on AI - this could be because of the existence of guidelines on subjects such as cyber security and risk management.

The SFC's CEO, Ms. Julia Leung, commented on AI in her speech at the HKIFA 16th Annual Conference on 5 June 2023, which can shine some light on what might be expected of firms using AI applications. For example, she said that AI should be used responsibly to augment, rather than replace asset managers in strategic decision making.

The speech highlighted that firms must be aware of potential risks, stay alert and make sure clients are treated fairly. Ms Julia Leung went on the explain that the onus will remain with the licensed corporation using the AI. So, they should thoroughly test AI to address any potential issues before deployment and keep a close watch on the quality of data used by the AI.

The speech also stated that firms should have qualified staff managing their AI tools, as well as proper senior management oversight and a robust governance framework for AI applications.

Then on 13 August 2024, HKMA and Cyberport launched a Gen A.I. Sandbox to bolster A.I. adoption in the financial sector. Then on the same day made opening remarks at FiNETech2 on "Building Secure A.I. Ecosystem: A Blueprint for Strategic Adoption in Financial Industry". The opening remarks referred to "a paradigm shift, where AI is reshaping the financial services industry."

On 19 August 2024, HKMA issued a circular to provide authorized institutions with a set of guiding principles in respect of the use of Gen AI in customer-facing applications from the consumer protection perspective.

The circular highlights that, as a subfield of Big Data Analytics and AI (BDAI), Gen AI shares a set of similar risk dimensions. As such, regarding consumer protection in customer-facing applications, the HKMA expects all authorized institutions to apply and extend the 2019 BDAI Guiding Principles to the use of Gen AI and continue to adopt a risk-based approach commensurate with the risks involved. However, the circular warns that since Gen AI uses complex models its potential risks could have an even more significant impact on customers.

Most recently, on 27 September 2024, the HKMA wrote to authorized institutions advising them of the publication of a research paper on Gen AI in the financial services sector. This particularly looks at Gen AI through the lens of operational efficiency, risk management, and customer engagement.

Australia
In Australia, a voluntary set of AI Ethics Principles has existed since 2019.

In terms of more recent developments, in January 2024, the Australian government published an interim response relating to its safe and responsible AI consultation. The government indicated that it wants to use a risk-based framework to support the safe use of AI and to prevent harm occurring from the use of AI, including:

- Considering and consulting on new mandatory guardrails for organisations developing and deploying AI systems in high-risk settings.
- The development of a practical, voluntary, best-practice toolkit to help entities ensure that AI systems being developed or deployed are safe and secure.
- Considering the merits of voluntary labelling and watermarking of AI-generated material in high-risk settings.
- Considering further opportunities to strengthen existing laws to address risks and harms from the use of AI.

Most recently, on 5 September 2024 the Australian Federal Government released two key documents as part of its broader agenda to promote safe and responsible use of AI in Australia:

- A proposal paper for introducing mandatory guardrails for AI in high-risk settings, which closed for comments on 4 October; and
- The Voluntary AI Safety Standards.

Republic of Korea
On 14 February 2023, the Ministry of Science and ICT announced a Strategy to realise trustworthy AI, which presented three strategies and ten action plans. These will be implemented on an ongoing basis up to 2025.

The Act on Promotion of the AI Industry and Framework for Establishing Trustworthy AI has now passed the final stage of voting and is under review by the National Assembly.[22]

Taiwan
Following the release of the "Core Principles and Related Promotion Policies for the Use of AI by the Financial Industry" in October 2023, the Financial Supervisory Commission ("FSC") issued "Guidelines for the Use of Artificial Intelligence (AI) in the Financial Industry" ("AI Guidelines") on 20 June 2024. Its public consultation closed for comments on 13 September 2024.

In terms of general provisions these cover:

- Relevant definitions of AI-related terms, description of the AI system life cycle, and risk assessment factors, in circumstances such as:

  a) When using AI systems, financial institutions should implement the core principles based on risk, and assess the risk levels of using AI systems after considering various risk assessment factors; and

  b) Supervision of third-party vendors where financial institutions should have in place supervisory measures when engaging third parties for introduction of AI systems, including establishing appropriate data or system migration mechanisms in the event of termination.

There are also specific chapters in the Guidelines covering: 1) Establishing Governance and Accountability Mechanisms; 2) Emphasising Fairness and Human-Centric Values; 3) Protecting Privacy and Customer Rights; 4) Ensuring System Robustness and Security; 5) Implementing Transparency and Explainability; and 6) Promoting Sustainable Development.

India
On 9 March 2023, Mr. Rajeev Chandrasekhar, the Minister of State for the Ministry of Electronics and Information Technology (MEITY), held a consultation to officially announce that the Information Technology Act of 2000 (IT Act) will be replaced with a new Digital India Act (DIA). The DIA is expected to include AI regulation as part of it.

The Securities and Exchange Board of India has also proposed that investment advisers and research analysts employing AI tools should provide complete disclosure on the extent of their usage of AI to clients.

Brazil

The Brazilian Senate announced, on 12 May 2023, that it will analyse Bill No. 2338 of 2023 to regulate AI systems in Brazil. In particular, the Senate highlighted that the bill creates rules for making intelligence systems available in Brazil, establishes rights for people affected by their operation, and provides for penalties for violations, as well as information regarding the supervising body.

More specifically, the bill specifies requirements for the operation of AI systems in Brazil, including a requirement that such systems undergo a preliminary assessment carried out by the suppliers themselves, to determine whether they can be classified as being of high or excessive risk.

United Kingdom

The UK wants to strike a balance between allowing innovation in AI to develop whilst also safeguarding public interest in AI ethics and governance.

Nikhil Rathi, Chief Executive FCA, delivered a speech at The Economist, in London during July 2023. This raised some questions around the regulation of AI in financial services but reiterated that the FCA does not want to stop innovation. The FCA's view is that its existing rulebook can accommodate AI use cases. This includes the FCA's high-level Principles, the Senior Managers Regime, the Consumer Duty, and operational resilience rules.

The UK Government has also published multiple documents covering the topic of AI, including a National AI Strategy and understanding AI ethics and safety. The most recent being a White Paper 'A pro-innovation approach to AI regulation', published in August 2023. The White Paper encourages regulators to align their AI strategies with five core principles based on fairness, transparency, and accountability.

In November 2023, the first global AI Safety Summit was held in the UK to discuss the risks of AI and the possibility of mandatory and voluntary regulatory frameworks. 28 countries, including the United States, China, and the European Union issued a declaration at the start of the summit, calling for international cooperation to manage the challenges and risks of artificial intelligence.

Then, in February 2024, the UK government published its response to its earlier consultation on its AI White Paper, reaffirming the flexible principles-based approach, and called on regulators to outline their strategic approach to AI. Work in multiple areas will be ongoing from 2024 onwards.

On 27 August 2024, the FCA updated its website on the Digital Regulation Cooperation Forum (DRCF) with information about the new DRCF AI and digital hub. The hub is a new, informal advice service set up by the members of the DRCF - the FCA, the Information Commissioner's Office, the Competition and Markets Authority and Ofcom. It is designed to support innovators with their cross-regulatory queries. The hub will run for a one-year pilot period, ending in March 2025.

Most recently, on 5 September 2024, the UK Government announced it had signed the first legally-binding treaty governing the safe use of AI.

Once the treaty is ratified and brought into effect in the UK, existing laws and measures will be enhanced.

[1] A test of a machine's ability to exhibit intelligent behaviour equivalent to, or indistinguishable from, that of a human.

[2] Neural Networks are computational models that mimic the complex functions of the human brain. The neural networks consist of interconnected nodes or neurons that process and learn from data, enabling tasks such as pattern recognition and decision making in machine learning.

[3] Programs capable of intelligence behaviour in simpler artificial environments - known as microworlds.

[4] Fuzzy Logic is based on the idea that in many cases, the concept of true or false is too restrictive, and that there are many shades of gray in between. It allows for partial truths, where a statement can be partially true or false, rather than fully true or false.

[5] Symbolic Artificial Intelligence (AI) is a subfield of AI that focuses on the processing and manipulation of symbols or concepts, rather than numerical data.

[6] Nouvelle AI is characterized by its ability to learn from minimal data, adapt to new environments, and make decisions based on abstract reasoning. Unlike conventional AI, which relies heavily on large datasets and specific programming, Nouvelle AI aims to mimic the human brain's flexibility and adaptability.

[7] An ongoing period of rapid progress in the field of AI that started in the late 2010s before gaining global prominence in the early 2020s.

[8] See https://www.oecd.org/en/about/news/press-releases/2024/05/oecd-updates-ai-principles-to-stay-abreast-of-rapid-technological-developments.html

[9] See What Is Quantum Computing? | IBM

[10] The field of quantum computing contains a range of disciplines, including quantum hardware and quantum algorithms. While still in development, quantum technology will be able to solve complex problems that supercomputers can't solve, or can't solve fast enough.

By taking advantage of quantum physics, fully realized quantum computers would be able to process massively complicated problems at orders of magnitude faster than modern machines. For a quantum computer, challenges that might take a classical computer thousands of years to complete might be reduced to a matter of minutes.

The study of subatomic particles, also known as quantum mechanics, reveals unique and fundamental natural principles. Quantum computers harness these fundamental phenomena to compute probabilistically and quantum mechanically.

Nevan's Law is named after Hartmut Neven, the director of the Quantum Artificial Intelligence lab. In late 2018, scientists at Google AI ran a calculation on Google's best quantum processor. They were able to reproduce the computation using a regular laptop. Then in January, they ran the same test on an improved version of the quantum chip. This time they had to use a powerful desktop computer to simulate the result. That rapid improvement has led to what's being called "Neven's law," a new kind of rule to describe how quickly quantum computers are gaining on classical ones.

[11] See Monetary Authority of Singapore (mas.gov.sg)

[12] QKD can help financial institutions (FIs) protect the exchange of cryptographic keys to address the cybersecurity threats posed by quantum computing.

[13] Harvard Business Review 'Managing the Risks of Generative AI' published 6 June 2023: See Managing the Risks of Generative AI (hbr.org)

[14] The Artificial Intelligence Insight forums, also known as the A.I. Insight forums, are a series of forums to build consensus on how the United States Congress should craft A.I. legislation. See U.S. Senate AI 'Insight Forum' Tracker | TechPolicy.Press

[15] Artificial Intelligence (AI) and Investment Fraud: Investor Alert | Investor.gov

[16] https://www.sec.gov/newsroom/speeches-statements/gurbir-remarks-pcce-041524#_ftnref13

[17] Regulatory Notice 24-09 | FINRA.org

[18] 2024a_205_signed.pdf (colorado.gov). California has a similar bill, AB2930, in legislature. Bill Text - AB-2930 Automated decision tools. (ca.gov)

[19] https://digitaldemocracy.calmatters.org/bills/ca_202320240sb1047

[20] EU AI Act contains four levels of risk in total: 'unacceptable risk', 'high risk', 'limited risk' and 'minimal risk'.

[21] 4 ways to future-proof against deepfakes in 2024 and beyond | World Economic Forum (weforum.org). See also EU AI Act definition of 'deepfake' meaning 'AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities, or events and would falsely appear to a person to be authentic or truthful.' (Article 3 (60).

[22] See assembly.go.kr

Please contact for further details:

David Morrison
Global Head of Trustee and Fiduciary Services
david.m.morrison@citi.com
+44 (0) 20 7500 8021

Amanda Hale
Head of Regulatory Services
amanda.jayne.hale@citi.com
+44 (0)20 7508 0178

Kelli O'Brien
United States
Head of Fund Administration
Product
kelli.a.obrien@citi.com
+1 617 859 3468

Ramesh Selva
North & South Asia (ex-Korea)
Head of Trustee & Fiduciary Services
ramesh.selva@citi.com
+65 6657 4142

Sung-Wook Han
Korea
Head of Trustee & Fiduciary Services
sungwook.han@citi.com
+82 22004 2162

Shane Baily
EMEA Head of Fiduciary Services
UK and Europe
shane.baily@citi.com
+353 1 622 6297

Jan-Olov Nord
EMEA Head of Fiduciary Services
Netherlands and New Markets
janolov.nord@citi.com
+31 20 651 4313

www.citibank.com/mss

cbs38605 10/24