OECD*publishing*

GENERATIVE ARTIFICIAL INTELLIGENCE IN FINANCE

OECD ARTIFICIAL INTELLIGENCE PAPERS

December 2023 **No. 9**

OECD
BETTER POLICIES FOR BETTER LIVES

The rapid acceleration in the pace of Artificial Intelligence (AI) innovation in recent years and the advent of content generating capabilities (Generative AI or GenAI) have increased interest in AI innovation in finance, in part due to the user-friendliness and intuitive interface of GenAI tools. Currently, the use of GenAI in financial markets involving full end-to-end automation without any human intervention remains largely at development phase, but its wider deployment could amplify risks already present in financial markets and give rise to new challenges. This paper presents recent evolutions in GenAI and its slow-paced deployment in finance, analyses the potential risks from a wider use of GenAI tools by financial market participants, and discusses associated policy implications.

Foreword

This report examines use cases of Generative AI in finance, discusses the current state of deployment of such tools in financial markets, analyses the main risks and challenges associated with a wider use of such tools in the future, and considers policy recommendations to mitigate those risks.

The report has been drafted by Iota Kaousar Nassr under the supervision of Fatos Koc and with oversight from Serdar Çelik from the Division of Capital Markets and Financial Institutions Division of the OECD Directorate for Financial and Enterprise Affairs. Nayana Satpathy provided market research assistance and Eva Abbott provided editorial and communication support.

The report supports the work of the OECD Committee on Financial Markets, chaired by Aerdt Houben. and was discussed by the Committee in October 2023. This paper was approved and declassified by the OECD Committee on Financial Markets on 1 December 2023.

The authors gratefully acknowledge valuable input and feedback provided by the following individuals and organisations: Adam Głogowski and Pawel Gasiorowski, National Bank of Poland; Daniela Gariboldi, Commissione Nazionale per le Societa e la Borsa (CONSOB); Giuseppe Grande, Banca d'Italia; Eleftheria Kostika, Bank of Greece; Benjamin Müller, Laura Felber and Lucas Marc Fuhrer, Swiss National Bank; Borut Poljšak, Bank of Slovenia; Paull Randt, Peter Grills and Rohan Paris, US Treasury; Tatu Rasanen, Bank of Finland; María Antonieta Campa Rojas, Banco de México; Mai Santamaria, Jefferson Vieira and Mark Curran, Department of Finance, Ireland; Angelika Schlögel and Kamran Kazemzadeh, Austrian Ministry of Finance and Vanessa Redak, Oesterreichische Nationalbank, Austria; Gibran Watfe and Hans Brits, De Nederlandsche Bank; as well as Antonio Gomes, Karine Perset, Luis Aranda, Rashad Abelson, Barbara Bijelic, Greta Gabbarini, Seohyun Kim and Masayoshi Chida from the OECD.

Table of contents

FIGURES

TABLES

Abbreviations and acronyms

AI	Artificial Intelligence
AGI	Artificial General Intelligence
ESG	Environmental Social and Governance
EU	European Union
FSB	Financial Stability Board
GANs	Generative Adversarial Networks
GenAI	Generative Artificial Intelligence
GPT	Generative Pre-trained Transformer
IoT	Internet of Things
LLMs	Large Language Models
ML	Machine Learning
NLP	Natural Language Processing
OECD	Organisation for Economic Co-operation and Development
P&L	Profit and Loss Statement
R&D	Research and Development
RBC	Responsible Business Conduct

Executive Summary

AI technologies and tools have been around for many decades, and their deployment in finance has been increasing in recent years (OECD, 2021[1]). The rapid acceleration in the pace of AI innovation, coupled with the advent of content generating capabilities, have contributed to renewed interest in this innovation. Most recently, the unveiling of conversational chatbot solutions (e.g. ChatGPT-4, Bard) represented a breakthrough in the evolution of AI. While 'traditional' AI models have been predominantly used for pattern identification, classification and prediction, Generative AI (GenAI) models are able to create 'original' output that is often indistinguishable from human-generated content.

The user-friendliness and intuitive interface of GenAI tools have accelerated the adoption speed at the demand side. Perhaps for the first time in AI-driven models, the average user can interact with complex AI technologies in an intuitive manner that fits human cognition. This has led to a surge in public attention and an increase in the direct usage by the public given easy and cheap (oftentimes free) access to some of these models. **The use of GenAI presents immense opportunities for efficiencies across sectors, including in finance, but comes with important risks that warrants the attention, and possible action, of policy makers.** Such considerations are likely to become increasingly important as the deployment of AI becomes ubiquitous across markets and its pace of development increases.

Despite the hype around GenAI, advanced use cases of AI in financial markets involving full end-to-end automation without any human intervention remain largely at development phase, if any. AI in production and live AI use cases are mostly used for process automation, as a way to enable efficiencies and improve productivity at the back-office (operations) and middle-office (compliance and risk management) of financial service providers and are used in an indirect manner (i.e. without direct interaction with the financial consumer). Large language models (LLMs), a type of GenAI, are being used as information points and as a way to simplify data analysis and reporting on the basis of the firm's data.

Such *slow-paced deployment* of GenAI by financial market participants could be attributed to the fact that finance is a highly regulated space (including model governance and risk management), as well as to the significant risks that GenAI involves in terms of false or deceptive outputs or other adverse consumer impact. In the future, however, GenAI advances are expected to increasingly support production and deployment of additional front-end use cases and applications and to further accelerate the existing use of such tools at the operational level. Financial market participants using such tools tend to deploy versions of foundation (base) models specifically made for their firm and deploy such restricted versions "offline" within their firewalls, or at the private cloud of their firm, with a view to ensure both data sovereignty and security and comply with existing frameworks for privacy, security and model governance.

Wider deployment of AI in finance could amplify risks already present in financial markets while also giving rise to new challenges and risks (OECD, 2021[1]). **GenAI exacerbates AI-related risks given its augmented capabilities and raises a number of additional novel issues.** Compared to other forms of AI, GenAI allows new possibilities for malicious activity including market manipulation, cyber-attacks from terrorists or state-related entities; or pure fraud tailored to the individual level. GenAI exacerbates explainability issues and the difficulty to explain poor performance and/or operational failures is magnified. Difficulties to assign accountability and governance-related issues are exacerbated, while

GenAI use also augments the relevance of third-party dependence and concentration risks involved if GenAI leads to wider use of third parties overall.

AI-driven models may intentionally or inadvertently generate biased or discriminatory results, and GenAI accentuates such risks given that models can potentially train from almost any data available, perpetuating biases and toxicity reflected on the web. The difficulty in understanding why and how AI-based models generate results, described as explainability, makes the detection of biases difficult, and is massively exacerbated in GenAI. GenAI model complexity accentuates the mismatch between non-linear multi-dimensional AI and human-scale interpretation that fits human cognition.

Many of the GenAI-related perils are associated with data management and governance, given the fundamental role of data for the training of models and the feedback loops of models trained also through user input. The quality and appropriateness of data used and their representativeness are associated with risks of misleading model outputs and inaccurate or unreliable models. Data privacy challenges are exacerbated in GenAI given the enormous amounts of unstructured datasets on which they are trained and that can come from any public source. These are likely to contain IP-protected information, possibly without appropriate permission or copyright, raising additional issues of authenticity of outputs.

GenAI accentuates risks related to the quality and reliability of the model output given risks of 'hallucinations' or other form of deception and disinformation/misinformation that can undermine or compromise the credibility of the financial market practitioners and is particularly concerning when it comes to the adverse impact on financial consumers, and retail investors in particular. The risks of limited trustworthiness of such models are aggravated by the possible limited awareness of the model limitations by its users and by the recipients of their outputs. Deepfakes could be used to spread disinformation that is difficult to detect and identify as false and if used by malicious actors at mass scale, could be used for market manipulation at large. Such risk of market manipulation would augment if versions of GenAI models used by financial institutions were to be constantly fed with information from the web. At the systemic level, the deployment of AI in finance could involve potential financial stability risks related to one-way markets, market liquidity and volatility, interconnectedness and market concentration to a few dominant actors.

The role of policy makers is important in supporting innovation in the sector while ensuring that financial market participants are duly protected and the markets around such products and services remain fair, orderly and transparent. Efforts to mitigate emerging risks could help instil trust and confidence and promote the adoption of such innovative techniques. Policy consideration and potential action could be looked at under a contextual, proportional and risk-based framework, and with a view to provide future proof policies that can withstand the test of time given the rapid pace of AI innovation.

Policy makers may need to consider reinforcing policies and strengthening defences and guard rails against risks emerging from, or exacerbated by, the use of GenAI (and other AI classes) in finance, focusing on a number of areas, including: strengthening of data governance including safeguards to overcome risk of bias/ discrimination; efforts to improve levels of explainability; foster transparency and consider disclosure requirements depending on the case; strengthen model governance and promote accountability mechanisms; promote safety, robustness and resilience of models (including for cyber risk) and mitigate risks of deception and market manipulation. Policymakers could encourage a human-centric approach and place emphasis in human primacy in decision making, particularly for higher-value use cases in finance (e.g. lending). Investment in R&D, skills and capacity will be needed to keep pace with advances, raise awareness of the potential perils and create tools to mitigate associated risks. Awareness should be raised as to the limitations of GenAI models, such as LLMs, starting with the fact that these have no reasoning capacity or comprehensive understanding of the world thus far, and currently have a limited ability to perform quantitative tasks. In line with the OECD AI Principles, there is a need to promote international multi-disciplinary and multi-stakeholder cooperation. Industry-led commitments, public education and financial education are additional ways to build awareness of benefits and perils and instil trust and confidence in the safe adoption of this transformative innovation in finance and beyond.

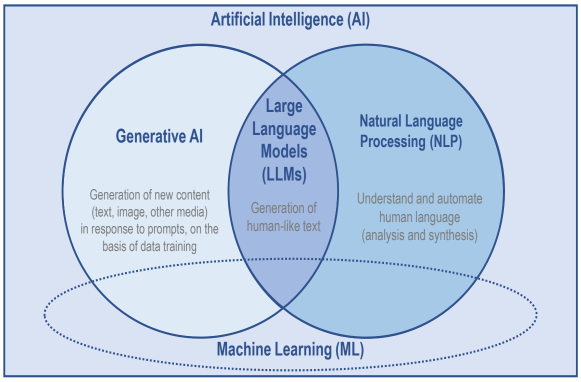
1 Generative AI (GenAI) and Finance

The emergence of models such as ChapGPT (OpenAI), Bard (Google), Bing Chat (Microsoft), Claude (Anthropic), or Ernie Bot (Baidu), and their functionalities have captured public attention and renewed interest of financial markets and financial policymakers to AI. The recent hype has been very much concentrated on GenAI models, and in particular Large Language Models (LLMs) such as the ones mentioned above, although LLMs are just one subset of GenAI. This paper explains GenAI and analyses use cases in finance and associated benefits.

1.1. GenAI: how is it different to other types of AI

Generative AI is a subset of AI comprising models that can create new content in response to prompts, based on their training data (OECD, 2023[2]). LLMs are one type of GenAI, involving the generation of language/text (e.g. ChatGPT), however, GenAI is also involved in generation of visual outputs (e.g. Sunthesia), audio (e.g. Speechify), images (e.g. DALL-E), code (e.g. GitHub Copilot) of other types of content generation. GenAI models are able to process and learn from massive amounts of unstructured datasets on which they are trained, including feedback received by users. They can create instantly new content as output based on various algorithms and mathematical architecture models such as the commonly used Generative Adversarial Networks (GANs)[1] which employ deep neural networks[2].

Figure 1.1. Generative AI



Note: indicative, non-exhaustive representation of AI domains.
Source: OECD authors' illustration.

[1] Generative Adversarial Networks (GAN) emanate in the category of Machine Learning (ML) frameworks, and use deep neural networks to generate (after training) content that aims to preserve the likeness of the original data (Yang et al., 2020[49]).

[2] Deep neural network architectures are inspired by the brain structure and functionality and are designed for unsupervised Machine Learning in the fields such as computer vision, NLP and recommendation engines.

The main advancement between the tasks that 'traditional' AI models are able to perform and GenAI is that the latter has the ability to produce seemingly original content. 'Traditional' AI models have been predominantly used for pattern identification, classification and prediction, while GenAI is able to create original output that is often indistinguishable from human-generated content.

1.1.1. Underlying drivers

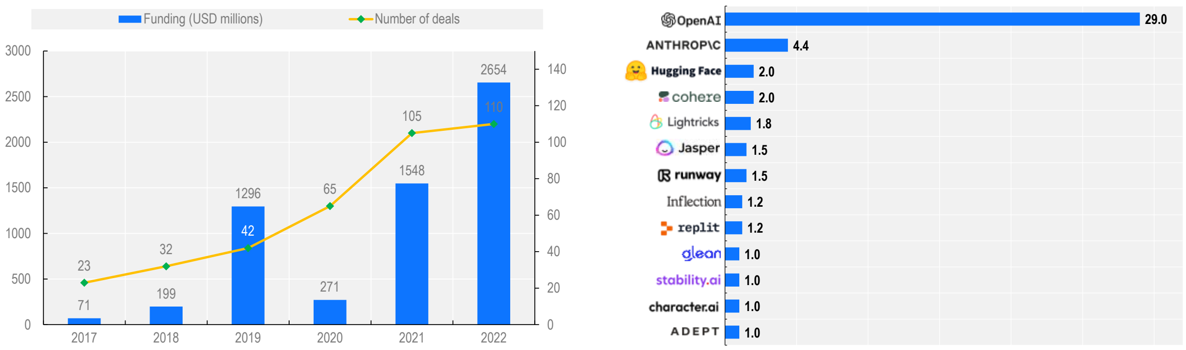
The fast development of AI has been driven by enormous advances in computational power coupled with an exponential growth of data, and of the underlying data processing capacity available at increasingly affordable cost (OECD, 2021[1]). Rapid advances in AI have accelerated over the past years; AI techniques such as deep learning, for example, may date back to the mid-80s, however their true potential has only been realized in the last decade or so (Kotu and Deshpande, 2019[3]). The exponential growth of data availability is also accelerating as big data is enriched with increasingly large amounts of datapoints from social media, transactional data, or data provided by devices connected to the internet of things (IoT) (Ahmed et al., 2017[4]). Progress made in synthetic data generation reduces the possible difficulty associated with the data volume required to train GenAI-based models, further fuelling this trend. Increased private funding is also supporting GenAI's development, with USD 2.6 bn across 110 deals in 2022 (

Figure 1.2).

Figure 1.2. Private funding flowing into GenAI and GenAI unicorns

in USD million (LHS), in absolute figures (RHS)

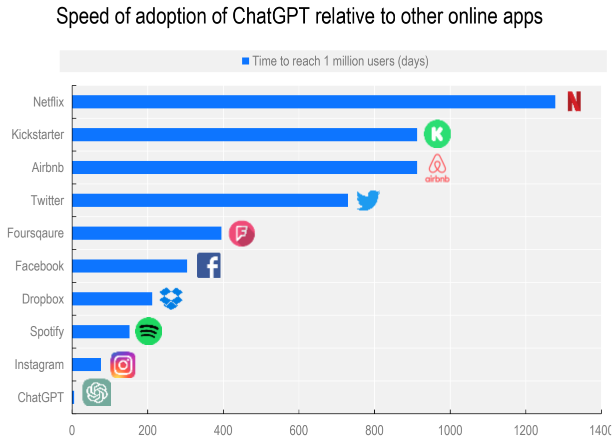
In USD bn, as of 08/05/2023

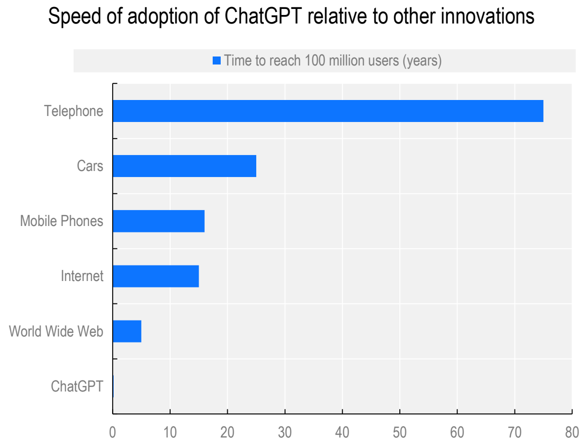


Note: Surge in 2019 is due to USD 1bn funding round of OpenAI.
Source: CB Insights (2023[5]), The state of generative AI in 7 charts, https://www.cbinsights.com/research/generative-ai-funding-top-startups-investors/

The user-friendliness and intuitive interface of GenAI and the conversational element of LLMs that emulate human cognition have accelerated the adoption speed on the demand side and speared public interest. The surge in public attention relates not just to the tremendous capabilities, but also to the fact that ChatGPT-alike models were one of the first AI models easily accessible, interactive and easy to engage with for the average consumer. Unlike other AI classes, such as deep neural or ML models, user-friendly GenAI models interact in a simple way and product outputs that resonate and fit the human cognition, mimicking human interaction, and has driven fast adoption (Figure 1.3). The conversational capability of LLM models such as ChatGPT, as well as the fact that these offered online versions for free, added to the ease of use by the average consumer.

Figure 1.3. Speed of adoption of some GenAI applications





Source: Statista and OECD calculations.

1.1.2. Slow-paced deployment of AI in finance

Despite the hype around AI, including Artificial General Intelligence (AGI)3, advanced use cases of AI in financial markets involving full end-to-end automation without any human intervention remain largely at development phase, if any (i.e. testing and experimentation). AI in production and live AI use cases are mostly found in process automation, as a way to enable efficiencies and improve productivity at the back-office (operations) and middle-office (compliance and risk management) of financial service providers (Figure 1.4). As of today, GenAI and LLMs are being deployed as tools to assist financial service provision (e.g. content generation, summarisation of documents used by financial advisors, human resources processes). Advances in GenAI are expected to support production and deployment of additional front-end use cases and applications and to further accelerate the existing use of such tools at the operational level.

Such slow-paced deployment of AI techniques in finance is only partly explained by the fact that financial market activity is a highly regulated space. Regulatory frameworks in finance safeguard market integrity, consumer protection and ensure financial stability, and require risk management (including of third-party risk), model governance, transparency and other obligations that may not be fully compatible with the most advanced GenAI tools, and outright incompatible with public versions of such models given high risks of security and data breaches (as well as maintenance/update costs). In alignment with their fiduciary duties, financial service providers have the legal responsibility to act in the best interest of the clients, which also means protecting them from risks of deceptive outputs, misinformation or other risks to financial consumers related to GenAI tools in particular (e.g. deceptive model outcomes, deepfakes etc. see Section 2.4).

Additional reasons when it comes to the use of open-source or off-the-shelf re-usable models (e.g. foundation models4) may involve risks of data privacy breaching for financial market participants as well

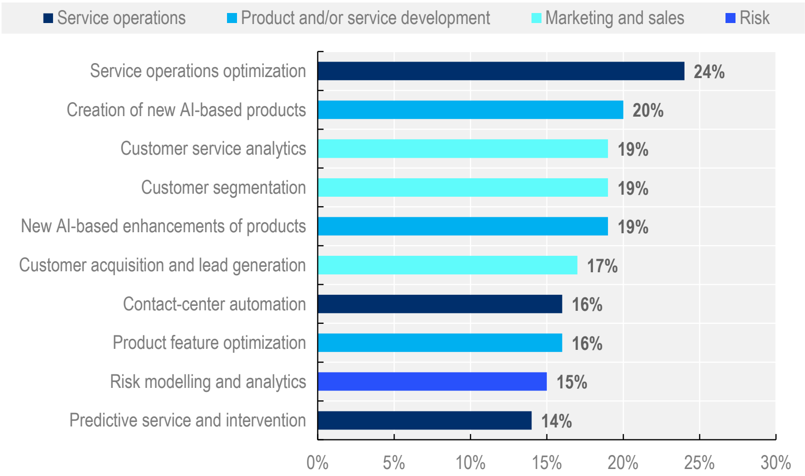
[3] Artificial general intelligence (AGI) displays capabilities that mimic or even surpass generally intelligent systems (e.g. humans) that are capable to learn, reason, and adapt (Goertzel, 2014[46]). Based on industry research, some AI specialists claim that the GPT-4 large language model (LLM) developed by OpenAI is a nascent form of AGI and that the model shows 'sparks' of human-like intelligence despite limitations (factual inaccuracy, hallucinations, inconsistencies) (Bubeck et al., 2023[47]) (OECD, 2023[2]).

[4] Foundation models (e.g. LLM models such as ChatGPT), are models trained in an unsupervised way on a huge amount of unstructured data and which can be adapted to many applications or use cases. The term has been coined

as lack sufficient explainability (see Section 2.3). To that end, financial market participants using such tools deploy version of LLM models specifically made for their firm and deploy such restricted versions “offline” within the firewalls of the firm, or at the private cloud of their firm, with a view to ensure both data sovereignty and security.[5]

Figure 1.4. Most commonly adopted AI use cases by financial institutions, 2022

In % of industry survey respondents



Source: McKinsey survey of financial institutions, (McKinsey, 2022[6]).

Also, adoption of such complex technologies is not straightforward in large institutions with legacy infrastructure and require careful planning before overlaying AI. For example, data management for the amount of unstructured data at the disposal of a bank is a pre-requisite for the deployment of AI models that train on data, however, such data management is not always guaranteed to be in place. Data may be available but not necessarily in a coherent, usable format suitable to the type of AI model used.

While the above limitations exist in supervised ML models in finance, they do not apply to GenAI and LLMs that are fully autonomous, self-supervised learning models. Such models eliminate the need for labelling of training data as they can understand relationships between, and learn from, unstructured data. At the same time, these models are trained primarily with natural language, and their ability to perform quantitative tasks is thus far limited.

Related, talent and skillset availability for the deployment of such technologies contribute to the paced approach of the industry. AI skills need to be present and involved at all levels and functions associated with the service provision using AI, requiring additional organisational capabilities to be employed.

The use of AI in trading is a great example of the slow-paced deployment of AI in finance. Rather than serving the entire chain of action from picking up signal, to devising strategies, and executing them, AI-based algorithms are mostly being used to extract signal from noise in data and convert this information into decision about trades, or to execute trading strategies. It could be envisaged that in the future, AI-

by the Center for Research on Foundation Models (CRFM) of the Stanford Institute for Human-Centered Artificial Intelligence (HAI) (HAI, 2023[48]). A foundation model can be used in infinite downstream AI systems.

[5] Still, in case of use of cloud services, the possibility of signals captured by the foundational model cannot be dismissed at this stage.

based algos will allow for autonomous and dynamic adjustment of their own decision logic while trading, taking the human completely out of the loop. Nevertheless, it should be highlighted that the deployment of AI in trading may increase the risk of prohibited or illegal trading strategies such as spoofing and front-running (OECD, 2021[1]).

Equally, in the case of GenAI, financial market practitioners experiment with offline or private restricted versions of LLM models. At this stage of development of the market, these are reported by the industry as mostly being used as information points and as tools to support internal processes and operations (e.g. summaries, translations) based on publicly available data. In the future, however, given the transformative power of these models, it can be expected that many new use cases will be successfully produced through experimentation and adopted by market participants for everyday use.

1.1.3. Direct vs. indirect scope of use of AI in finance

Irrespective of the type of model, there are different levels of AI interaction with the financial service provider and/or with the end customer, and each involves a different level of associated risk. These different levels may also underpin the slow-paced deployment of AI in finance, with AI today being mostly used to assist, rather than replace, the humans. Future development of these models and of frameworks ensuring their safety and trustworthiness might encourage the transition to more direct interaction of the model with customers.

AI can be used to assist financial service providers to interact with customers in an indirect way, for example preparing portfolio allocation recommendations tailored to the customer's profile. More direct involvement of the customer, for example allowing the model to directly recommend or even automatically execute the suggested recommendations is much riskier, as it takes away the human parameter.

The risks are even greater when the model is used by a non-expert service provider who may rely exclusively on the model's recommendation, without the ability to sense-check and validate the model's output. Similarly, fully automated direct interaction where the model executes its own recommendations without feedback is exposing both the customers and the provider of the service who is using the model to significant risks.

Box 1.1. A brief overview of OpenAI's ChatGPT

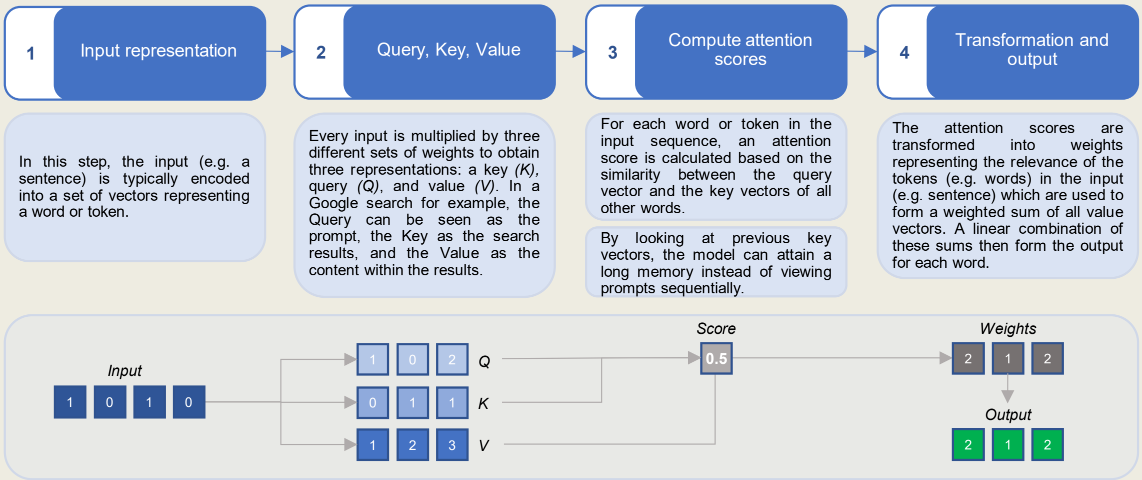
OpenAI's conversational agent ChatGPT has garnered considerable attention since its launch in November 2022 due to its ability to engage in human-like, contextually relevant conversation in response to user prompts. The advanced chatbot is an LLM model based on a Generative Pre-trained Transformer (GPT) architecture powered by neural networks using a self-attention mechanism (see The neural networks are trained on enormous amounts of data and consist of billions of parameters that describe the interrelationships between different nodes in the networks. The continuous computation of attention scores for all input tokens (such as words) during allows the model to pick up on complex relationships between words and identify substance in the conversation.

Figure 1.5). The self-attention mechanism allows the transformer to analyse the entire input prompt (including previous prompts in the conversation) and determine which parts are most relevant for the output it is to generate.

The neural networks are trained on enormous amounts of data and consist of billions of parameters that describe the interrelationships between different nodes in the networks. The continuous

computation of attention scores for all input tokens (such as words) during allows the model to pick up on complex relationships between words and identify substance in the conversation.

Figure 1.5. Self-Attention Mechanism



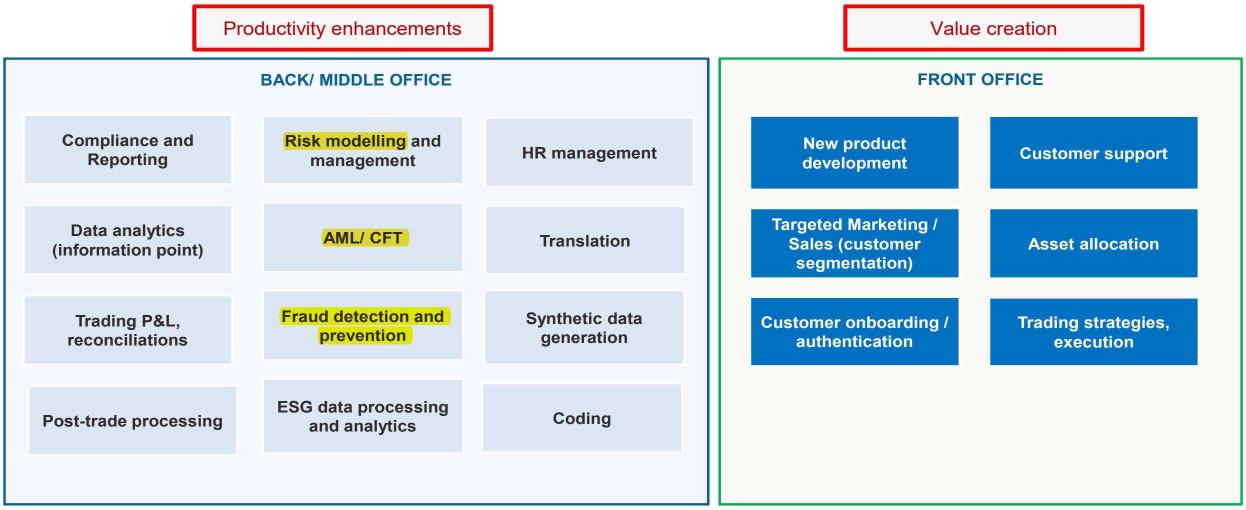
Source: OECD based on Karim (2019[7]) and ChatGPT.

The most advanced version of OpenAI's conversational agent is GPT-4 released in March 2023. GPT-4 is a multimodal LLM (accepting both text and image inputs) that is said to have advanced reasoning skills, capabilities in complex instructions and higher creativity. Compared to its predecessor (i.e. GPT-3.5), GPT-4 accepts a larger amount of input tokens and scores considerably higher on multiple professional and academic benchmarks and tests (e.g. GPT-4 scores in the top 10% of a simulated bar exam while GPT-3.5 scores in the bottom 10%) (OpenAI, 2023[8]). The multimodal model is a larger model compared to GPT-3.5 with a higher number of parameters, however details on the extend of its size have not been disclosed by OpenAI (MIT Technology Review, 2023[9]).

1.2. AI and GenAI use cases in finance and associated benefits

AI tools are being embedded in various parts of the financial markets (securities, asset management, banking, etc.) and at different parts of the value chain of financial products and services (Figure 1.6). Accordingly, AI technology is being deployed across multiple verticals, including asset management (e.g. stock picking; risk management and operations); algorithmic and high-frequency trading (e.g. liquidity management and execution with minimal impact); retail and corporate banking (e.g. onboarding, creditworthiness analysis, customer support) and payment institutions (e.g. AML/CFT, fraud detection) (OECD, 2021[1]). GenAI and large language models' capabilities are expected to further enhance the capabilities of AI-based services by financial institutions, particularly in areas such as sales and marketing, customer support and operations, including coding for data/information management and software development.

Figure 1.6. GenAI use cases in finance



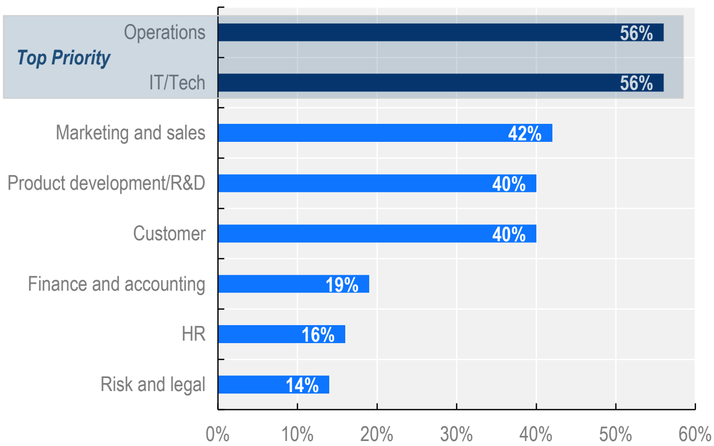
Source: OECD authors' illustration.

AI techniques have been used at the back-office of a wide range of financial market practitioners, given the potential of automation to deliver performance and productivity improvements (Figure 1.7). Operational optimisation is one of the most widely reported use cases of AI in finance today, with the potential to increase both efficiency and accuracy of operational workflows and enhance performance. Manually-intensive and repetitive P&L and other reconciliations can be replaced with faster and cheaper automated ones.

GenAI in particular can be used as an information point (similar to today's search engines) and as a way to 'humanise' and simplify internal and external data analysis and reporting on the basis of the firm's data. The latter can include customer service analytics as well as human resources tasks (e.g. generation of summaries of management reviews), or for translation or summarisation of contracts or other reporting. GenAI models can also be used for individualised communication at the front end, both for product creation and for marketing and sales purposes, as well as for enhanced customer support.

Figure 1.7. GenAI use cases in finance, 2023

In % of industry survey respondents



Source: (KPMG, 2023[10]).

The use of AI-based anomaly detection tools can improve both AML/CFT processes and fraud detection in all types of financial market participants, and in particular in payments. Such models automatically identify outliers in given datasets, that are not conforming to the normal behaviour in a dataset (Kotu and Deshpande, 2019[11]). Indicatively, unusual transaction activity of a consumer that could imply fraudulent activity. Similarly, AI can assist automated onboarding and the performance of KYC checks for banking clients, as well as compliance functions of financial market participants. GenAI can further enhance these uses by producing reporting and other output required for compliance purposes, on the basis of the firm's data.

AI is also used to support and strengthen risk management for asset managers and institutional investors. AI-based risk models can instantly test portfolio performance under multiple market and economic scenarios on the basis of a myriad risk factors monitored on a constant basis. Data and information has always been the cornerstone of many investment strategies, from fundamental analysis to systematic trading and quantitative strategies alike. While structured data was at the core of such 'traditional' strategies, vast amounts of raw or unstructured/semi-structured data are used in AI-based models to provide an informational edge to investors in the implementation of their strategies enhancing sentiment analysis and providing additional insights based on pattern recognition (OECD, 2021[1]).

ML models in particular are used for pattern recognition, natural language processing (NLP)[6] and computational learning and can provide recommendations that influence decision-making around portfolio allocation and/or stock selection given their predictive capacity (Table 1.1). ML models for predicting returns have received widespread attention in recent years given the ability of non-linear models, such as neural networks, to capture complex relationships between stock characteristics and future returns and to learn from data. This has resulted in the potential for ML-based stock-selection strategies to be included and inform portfolio construction in financial markets. Academic studies examine whether or not such ML-based strategies are 'alpha' generating, with various results (Freyberger et al., 2020[12]; Moritz and Zimmermann, 2016[13]). Interestingly, longer-horizon ML-based strategies select slower signals and load more on traditional asset pricing factors and thereby tend to have a relative worse performance compared to short-term strategies (Blitz et al., 2023[14]).

Table 1.1. Select types of GenAI applications by financial services firms

	Segment	Service	Description
Goldman Sachs	Corporate and Investment Banking	Code generation	ChatGPT-style AI in-house to assist developers with writing code
JPMorgan Chase & Co.	Corporate and Investment Banking	Code generation	Toolkit called Senatus to facilitate the software development process through features such as code recommendations
Deutsche Bank	Corporate and Investment Banking	Financial analyst assistant	Testing Google's generative AI and large language models (LLMs) at scale to provide new insights to financial analysts
Bloomberg	Financial research	Financial assistant	Finance-specific LLM trained on Bloomberg data
Brex	FinTech	Expense management	ChatGPT-style CFO tools. Provides insights on corporate spend and answer critical business questions in real time
الآن ALAAN	FinTech	Expense management	Integration of OpenAI's advanced AI technology onto its user platform to provide insights on corporate spend and answer critical business questions in real-time

[6] Natural language processing (NLP) is an interdisciplinary AI domain aiming at understanding natural languages as well as using them to enable human–computer interaction. It differs from text mining in that it takes into consideration the surrounding information and is concerned with processing the interactions between source data, computers, and human beings (Stephanie Kay Ashenden, 2021[50]).

	Segment	Service	Description
CLEO	FinTech	Financial assistant	App connects to bank accounts and gives clients proactive advice and information on finances, including timely nudges, helping stay on top of their spending
CITADEL	Hedge fund	Code, software development, information management	From helping developers write better code to translating software between languages to analyse various types of information
Morgan Stanley	Wealth management	Financial advisor assistant	Financial Advisors to use GPT-4 capabilities to ask questions and contemplate large amounts of content and data exclusively from MSWM content and with links to the source documents
Morgan Stanley	Wealth management	Sales and marketing	Next Best Action is an internally-built AI-based engine that delivers timely, customized messages to clients and prospects guided by the Financial Advisor
Klarna.	FinTech	Product recommendations	Highly personalized shopping experience through curated product recommendations to users asking for advice
TROVATA	FinTech	Treasury tool	Generative AI Finance & Treasury Tool

Note: Non-exhaustive and based on reported information by financial market participants.
Source: OECD based on web research; (2023[15]), Venture & Growth 2023 Outlook, https://vgb.lazard.com/lazard-vgb-insights-2023-outlook/#FinTech; Dadan and Shetty (2023[16]), Generative AI in Finance and Beyond, https://whitesight.net/generative-ai-in-finance-and-beyond/.

In trading, AI can reduce frictions and create efficiencies, increasing also the speed of trading. AI-based algorithms learn from data inputs and dynamically evolve into computer-programmed algos, able to identify and execute trades without any human intervention. In highly digitised markets, such as equities and FX markets, AI algorithms can enhance liquidity management and execution of large orders with minimal market impact, by optimising size, duration and order size in a dynamic fashion, based on market conditions. Traders can also deploy AI for risk management and order flow management purposes to streamline execution and produce efficiencies (OECD, 2021[1]). Advances in AI make it possible to also anticipate and dynamically respond to trading strategies of other market participants, increasing the ability of traders to predict other traders' behaviour. The latter can translate into increased risk of collusions (OECD, 2021[17]).

AI models in lending could reduce the cost of credit underwriting and facilitate the extension of credit to 'thin file' clients, potentially promoting financial inclusion. The use of AI can create efficiencies in information management and data processing for the assessment of creditworthiness of prospective borrowers, enhance the underwriting decision-making process and improve the lending portfolio management. It can also allow for the provision of credit ratings to 'unscored' clients with limited credit history, supporting the financing of the real economy (e.g. SMEs) and potentially promoting financial inclusion of underbanked populations.

Most of the potential of GenAI is expected to be found at the front-end of financial service provision, given the potential benefits for improved, personalised customer experience. 'Traditional' AI classes were used to power chatbots and automated call centres for customer support. It should be noted that the first chatbot was created as early as 1966, and has since kept evolving with advances of AI classes (Weizenbaum, 1966[18]). GenAI enhances these with a human-like conversational element that often makes it difficult to distinguish whether the interlocutor is a machine or a human.

GenAI's potential is also expected to come in the development of new products, in product-feature optimisation as well as in better targeted sales and marketing (Table 1.2). AI-based enhancements of products can support and enhance both the production of products and services (e.g. investment advice) and its delivery (e.g. GenAI-powered robo-advisors). GenAI can also allow for customer segmentation at the individual level, allowing brokerage firms and other investment advisors enhance their robo-advice with differentiated recommendations produced in a fast, efficient and customised manner, and delivered in a human-like conversational manner tailored to each client.

The combination of financial analysis assistance and communication tool is perhaps the area with the greatest immediate impact in finance, and is becoming even more relevant in light of developments such as platformisation or embedded finance. The information management and predictive capacity of AI can support product recommendations and advice. Recommendation engines are a class of ML techniques that predict a user preference for an item, and are boosted by the implementation of methods such as content-based filtering (Kotu and Deshpande, 2019[19]). These capabilities are complemented by the content generation capacity of GenAI to create sales strategies and marketing campaigns differentiated at the individual customer level. Such strategies could also have a possible benefit of greater financial inclusion, depending on the intended purpose of the customisation.

Coding is the other area with potential immediate impact in finance (and beyond), as GenAI can support software development used across the board of financial services/products. GenAI can be used as an assistant dedicated to coders for the development of software applications or other models. GenAI applications can generate new code, resolve bugs in scripts or provide solutions to coding errors, while they can also perform testing of given code. This use case is rather underdiscussed, despite its enormous practical potential and the relatively limited downside risks.

Related, GenAI systems have the ability to generate synthetic data at scale, and in a customised manner tailored to specific market scenarios. Synthetic data is artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data, with potential for enhanced privacy, lower cost and improved fairness (EDPS, 2021[20]). The most pertinent use case relevant to financial market is the creation of simulated financial market data for scenario analysis, as well as the creation of datasets for testing, validation and calibration of AI-based models in finance (and beyond).

AI could also be used to support sustainable financing and ESG investing, particularly by investment professionals who can use NLP to generate real-time ESG assessments based on firms' communications such as corporate social responsibility reports (ESMA, 2023[21]). Today, the use of AI tools in investment strategies are predominantly made to process ESG-related data due to the unstructured and complex nature of datasets relating to ESG-factors which requires the use of more sophisticated analysis techniques (Papenbrock, GmbH and Ashley, 2022[22]). In light of the growing trend towards ESG investing, asset managers are also becoming involved in the ethical usage of AI by companies they invest in. For example, Norway's (and the world's largest) sovereign wealth fund will be introducing standards on AI use for companies it invests in, with a view to align AI usage to the fund's responsible investment framework and ESG commitments (FT, 2013[23]).

Table 1.2. Select AI companies offering GenAI applications for finance

	Service	Description
ACTIVE.Ai	Customer support	Using AI to provide conversational finance and banking services and help financial companies integrate virtual intelligence assistants into their services
LIVEPERSON	Customer support	Offers meaningful one-to-one support and conversational AI solutions
boost.ai	Customer support	Scalable and user-friendly customer service solution for banks with a focus on customer experience
STREEBO	Customer support	Smart digital omni-channel banking experience powered by Conversational AI
glia	Customer support	Digital customer service technology and interaction platform
FinChat.io	Financial analysis	ChatGPT for financial analysis
TOGGLE	Financial analysis	TOGGLE sifts through billions of real-time data points to provide financial analysis
Portrait Analytics	Financial assistant	Provides personal research assistance that empowers analysts to discover investment ideas at a high page

	Service	Description
Boltzbit	Synthetic data generation and analysis	Offers database linking, portfolio optimization and enhanced prospect profiling through the generation of synthetic financial data.
MOSTLY·AI	Synthetic data generation	AI-generated, privacy compliant synthetic transaction data for banks and financial institutions
FINCRIME DYNAMICS	Synthetic data generation	Generates tailored synthetic data with the latest financial crime simulations to build high performance machine learning crime detection
TONIC THE FAKE DATA COMPANY	Synthetic data generation	Generates synthetic financial data to enhance development tests and compliance requirements for financial services professionals

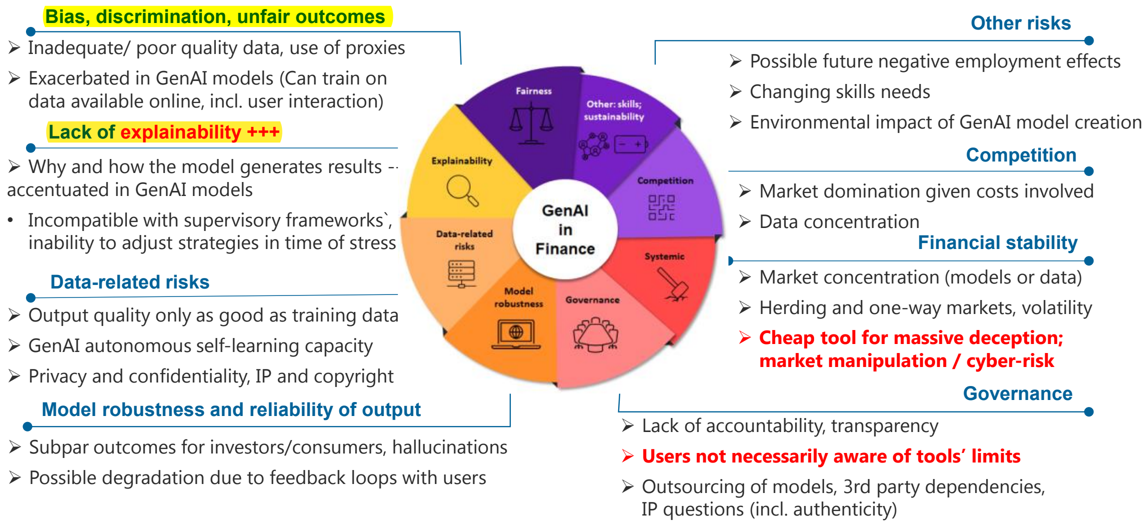
Note: Non exhaustive list.
Source: OECD compilation based on public sources.

2 Risks of GenAI in Finance

The deployment of AI in finance could amplify risks already present in financial markets given their ability to learn and dynamically adjust to evolving conditions in a fully autonomous way, and give rise to new overriding challenges and risks (OECD, 2021[1]). GenAI exacerbates some of these AI-related risks given its augmented capabilities and raises a number of additional novel issues associated with its specificities (e.g. LLMs conversational nature). This section analyses both categories of risks: novel risks related to GenAI or risks present in all classes of AI and accentuated by the use of GenAI in particular (Figure 2.1).

Some of the risks presented by GenAI are noteworthy compared against risks of other AI tools. Advances in GenAI models, such as LLMs, allow new possibilities for malicious activity that could include sophisticated market manipulation strategies, cyber-attacks including by terrorists or state-related entities or pure fraud tailored and personalised to the individual consumer. Explainability difficulties are exacerbated, and it is difficult (if not impossible) to explain poor performance and/or operational failures. Related, assignment of accountability to the human parameter and governance difficulties are magnified; while the use of GenAI also augments the relevance of third-party dependence and concentration risks involved in the use of AI in finance more broadly.

Figure 2.1. GenAI risks and challenges



Note: Non-exhaustive list.
Source: OECD authors' illustration.

2.1. Fairness: risk of bias, discrimination and unfair outcomes

The potential for bias and discrimination in algorithmic decision making is a well- established concept since the earlier days of ML model deployment in finance. AI-driven models may inadvertently generate biased or discriminatory results due to the use of poor quality or inadequate datasets to train the model (e.g. the inclusion of gender-based variables or protected category data such zip-code as input to the model). Such results can also be unintentional, as algorithms have the capacity to combine facially neutral data points and treat them as proxies for immutable characteristics such as race or gender, thereby circumventing existing non-discrimination laws and inferring variables not (allowed to be) included in the dataset. Unfair results can also be intentional, if the datasets made available to the model for its training are manipulated to include such biases.

Particularly in the case of GenAI, given that the model can potentially train from almost any data available online, the model can reinforce and exacerbate biases that are already in the data used to train it, perpetuating historical biases reflected on the web. The latter includes toxic and hate speech that language models may use to train on, encouraging discrimination. The under-representation or exclusion of some data, or equally, the predominance of others, can reduce the accuracy of the model and distort results, as was the case of the Gender Shades project for facial recognition (Buolamwini, 2018[24]). The fact that GenAI models can also learn from user feedback and interaction increases the risk of deepening any prejudices and societal inequalities depending on the profile of users. This risk is also present in the use of such tools for screening by human resources HR in finance (and beyond).

A notable example of risk of discrimination and bias in finance is credit allocation and the potential for discriminatory or unfair lending on the basis of AI-based models (OECD, 2021[1]). The exclusive use of such models for credit allocation decision making can raise risks of disparate impact in credit outcomes, while making discrimination in credit allocation even harder to identify. The use of proxies in the dataset can also result in undetected discrimination or bias. Additionally, outputs of AI-based models are difficult to interpret and communicate to declined prospective borrowers given limits to the explainability or interpretability of such models.

2.2. Lack of explainability

The difficulty in understanding why and how AI-based models generate results, described as explainability, can give rise to important risks and incompatibilities in the area of finance. The explainability-related difficulties and risks that were already an issue in ML models are massively accentuated in the case of GenAI. Advances in AI model capacity, as evidenced by recent GenAI applications, exacerbate the mismatch between the increasing complexity of non-linear multi-dimensional AI models and the demands of human-scale reasoning and interpretation that fit the human cognition. The dynamic nature of AI models which learn from the data and feedback they receive in an automated manner (unsupervised models) further obscure their interpretability[7].

The lack of explainability of AI models used in finance make the adjustment of investment or trading strategies difficult in times of poor performance, as there is no linear relationship to be traced or clear understanding of which parameters drove the outcomes or decisions of the model. Limited explainability of the models makes it harder to detect inappropriate use of data or use of unsuitable data in AI-based applications in finance. In lending, for example, consumers have limited ability to identify and contest unfair credit decisions, and even in case of a fair lending decision, prospective borrowers have limited ability to understand what steps they could take to improve their credit outcomes.

The scale of complexity and difficulty in explaining and reproducing the decision mechanism of GenAI models makes it challenging to mitigate risks stemming from its use. The speed of development of AI seems to outpace progress made in explainability of AI-based models, until now mainly concentrated in

[7] Interpretability refers to the meaning of the model's output in the context of their designed functional purposes, while explainability refers to a representation of the mechanisms underlying the AI system's operation (NIST, 2023[27]).

ML models. Importantly, limited explainability levels can result in low levels of trust in AI-assisted financial service provision by customers and more broadly by market participants.

2.3. Data-related risks: data quality, privacy, concentration of data

Data is fundamental both for finance, and for AI. The quality of the output of the model will be only as good as the data used to train the model, and the risk of 'garbage in, garbage out' exists across all types of models. The choice of datasets that the AI will training on has enormous unseen influence on the output of the model, which also translates into unseen influence of the people making the choice of data. In the case of GenAI, data quality questions also involve the input that the users introduce into the model in their interaction with the model, given the autonomous self-learning capacity of the model and the feedback loops between the user and the model.

Data representativeness and relevance are also associated with risk of misleading model outputs and inaccurate or unreliable models altogether. In terms of representativeness, data used needs to provide an exhaustive representation of the population under study, with balanced representation of all relevant subpopulations in order to avoid risk of bias or discrimination. Data relevance is particularly relevant to GenAI tools and relates to possible inclusion in the data of exogenous false or misleading information. This includes user input and corresponding influence, and the possible reduction in model performance or fairness depending on the users' input and interaction with the model, as the model keeps dynamically evolving after its training.

In the future, the wider introduction of plug-ins to private versions of GenAI models that allow the models to access any online content[8] could increase risk of data privacy and confidentiality breaches of private information used for the model's curation or shared with the model during its deployment. The sheer volume of data used and their continuous flowing into AI systems, and the speed of growth of the size of datasets used for their training accentuate such risks of breaches. Information provided by users in prompts used to query to the model can also reveal private or proprietary information that are collected and can be used by the model in future iterations or training. The more specific the data included in the prompt, the more accurate the result of the model will be in the case of LLMs in particular, and the higher the risk of confidential data leakage.

Cyber security risks, risk of hacking and other security risks related to such models are also present, similar to any other digital financial services, with direct implications on data privacy and confidentiality. Fraudsters can make full use of the capabilities of GenAI to produce tailored individualised fraud attacks at massive scale with less resources required. GenAI can offer to malicious actors improved methods for performing social engineering, email phishing, and text messaging smishing attacks compromising access into firms' systems, emails, databases, and technology services (Federal Reserve Board, 2023[25]). When it comes to security, the use of external models exposes the firm to increased risks of cyber security breaches observed in the use of third-party software provision or – to a much greater extent - open-source systems.

Given such risks, large financial market participants opt for private restricted versions of models deployed offline within the firewalls of the firm or within the firm's private cloud to ensure the safety of the application from an operational/IT perspective, while also protecting client data and proprietary intellectual property of the firm. Such models are also easier to oversee from a compliance perspective for the purposes of model governance by the firm (e.g. in terms of data fed into the model and ensuring privacy). However, to the extent that the initial model gets informed by the outputs of any downstream applications, or is able to receive signals by such private versions of the model, gives rise to similar data-related risks described

[8] Reported anecdotally by the industry to be only done at small scale as of today.

above. It is currently unclear to what extent such feedback loops exist from downstream applications to the initial/main model in case of the use of private versions on a private cloud of third parties.

2.3.1. Authenticity: risks to intellectual property (IP) and copyright

Data authenticity and IP-related risks are closely related to data quality and data privacy-related considerations, and are particularly prominent in GenAI models. GenAI models, such as LLMs, are curated on massive amounts of unstructured data that can come from any public source. These are likely to contain IP-protected data, possibly without appropriate permission or copyright. What is more, the authenticity of the output generated by the model can therefore be doubted given the above.

Data provenance and data location are additional parameters for consideration in GenAI in particular. Where are the data on which the model has been trained located? Who can access the data, with customer consent? The intersection of AI model-related data management and data sharing frameworks, such as open banking or open finance, which allow for third parties to access customer data will be particularly interesting to examine in light of IP and data ownership questions.

Additional considerations around ownership of the data that financial services firm hold and may use to train their private versions of GenAI models are to be examined. When it comes to such restricted versions of models for the purposes of a financial services firm, the intellectual property rights of the model itself and its output (that can be included in future iterations of curation of the model) are additional parameters for possible consideration.

2.4. Model robustness and resilience, reliability of outputs and risk of market manipulation

According to the OECD AI Principles, AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed (OECD, 2019[26]). Poor reliability and accuracy of AI-driven model outputs raise risks of subpar outcomes in any type of financial use case involved (e.g. poor or inadequate investment advice). AI models lacking robustness and resilience do not perform as intended; or perform in ways that give rise to potential harms to people if it is operating in an unexpected setting. In other words, such models cannot withstand unexpected adverse events or unexpected changes in their environment or use, to the detriment (also) of the end users (NIST, 2023[27]).

Issues around the quality and adequacy of data used to train the model (discussed in Section 2.3), model drifts (involving concept or data drifts) or risks of over-fitting are some of the examples of risks related to model accuracy and resilience in ML models used in finance (OECD, 2021[1]). For example, tail and unforeseen events give rise to discontinuity in the datasets that were used to train the model, which in turn creates model drifts that undermine the models' predictive capacity particular in times of market turbulence or stress.

In GenAI models, interaction with the users and the feedback loops of autonomous self-learning models may result in a degradation of the model's accuracy, although it is difficult to understand the drivers of such deterioration in the absence of explainability. Findings of recent empirical analyses show that the behaviour of the same LLM model can change substantially in a relatively short amount of time, highlighting the need for continuous monitoring of LLMs (Chen, Zaharia and Zou, 2023[28]). It is difficult, however, to understand whether this is the result of finetuning and updating of the model, or whether any deterioration in the model's accuracy comes as the result of the model's interaction with users (e.g. poor quality input) that the model learns from in a dynamic manner.

GenAI accentuates risks related to the quality and reliability of the model output, as the advent of these AI classes brought about the risk of 'hallucinations'[9] or other form of deception and disinformation/misinformation[10]. The provision of false information and advice by an AI-driven model deployed for the provision of financial services can undermine or compromise the credibility of the financial market practitioner providing the advice, and is particularly concerning when it comes to the adverse impact on financial consumers, and retail investors in particular. The risks of limited trustworthiness of such models are aggravated by the possible limited awareness of the model limitations by its users and by the recipients of their outputs.

GenAI-model deception can be unintentional, as in the case of AI-driven content generation that does not correspond to any real-word input, or intentional, as in the case of identity theft by bad actors. It can involve simple forms of deception, e.g. convincing financial advisors to apply discretionary pricing of products based on the purchasing power or other profile characteristics of the client in an untransparent manner. Overcoming the difficulty in differentiating between the factually correct information and the inaccurate, false or deceptive information will be critical to overcome such risks, whether intentional or unintentional. Researchers and the industry are testing the use of AI tools to identify algorithmically-generated content and help users to learn how to spot these, although these are still early days of experimentation (Groh et al., 2022[29]).

At mass scale, malicious actors could use GenAI's capabilities to intentionally manipulate financial markets at large. The simplest example of such *data* manipulation would involve the provision of false information about stocks or other investments; or the provision of deceptive advice to prospective investors or other financial consumers. Deepfakes[11] (e.g. voice spoofing or fake images generated by GenAI) could be used to spread disinformation that is difficult to detect and identify as false and misleading given the capabilities of these models (e.g. rumours that could cause market instability or panic).

The risk of such market manipulation would augment if versions of GenAI models used by financial institutions were to be constantly fed with information from the web (e.g. social media) in real time. Currently, this is not the case as special restricted versions of GenAI models are being mostly used offline or in private cloud environment, without interaction with the web, to comply with existing risk frameworks. In a future scenario where plug-ins allow real time internet data to feed such proprietary versions of the model, the risk of market or investor manipulation is significantly magnified and the possibilities of malicious actors manipulating the markets increase considerably (e.g. spread of false rumours about financial results of a listed company through social media or other media outlets).

---

[9] Artificial hallucination refers to the phenomenon of a machine, such as a chatbot, generating seemingly realistic sensory experiences that do not correspond to any real-world input. This can include visual, auditory, or other types of hallucinations. Artificial hallucination is not common in chatbots, as they are typically designed to respond based on pre-programmed rules and data sets rather than generating new information. However, there have been instances where advanced AI systems, such as generative models, have been found to produce hallucinations, particularly when trained on large amounts of unsupervised data (Ji et al., 2023[52]).

[10] Disinformation in AI LLMs defined as deliberate fabrication of untrue content designed to deceive (e.g. writing untrue texts and articles), while misinformation involves the false or misleading information that does not intend to harm (e.g. creating falsehoods for entertainment) that can damage public trust in democratic institutions (Lesher, Pawelec and Desai, 2022[51]) (OECD, 2023[2]).

[11] Deepfakes refer to fake videos or other fake content generated using AI technology. At the extreme scenario, it may not be possible to differentiate whether the interlocutor is a machine or a human being.

2.5. Governance-related risks: lack of accountability and transparency

Similar to any type of model, the use of AI-based models by financial market participants would need to comply with their existing model governance frameworks and oversight arrangements. This includes clear lines of responsibility for the development and overseeing of AI-based systems throughout their lifecycle, from development to deployment, and explicit designation of accountability for any adverse outcome produced by the model.

Accountability, however, presupposes transparency (NIST, 2023[27]). For transparency to be meaningful in advanced GenAI models, extensive information would need to be included in the disclosure around the model and the data used by the model, including for example data sources, copyrighted data, compute- and performance-related information, limitations of the model, foreseeable risks and action taken to mitigate such risks, including evaluation, as well as information around the environmental impact of such models. The latter is increasingly important for financial market participants wishing to align the use of AI applications with their ESG practices.

There may be challenges to the achievement of high transparency levels for GenAI models, depending on their specific characteristics. For example, it may be difficult to disclose information about the copyright status of training data when such data includes unstructured information curated from the internet (Bommasani R. et al., 2023[30]). Similarly, energy usage and emissions reporting may be challenging given difficulties in accurately measuring such impact (similar to the case of DLTs (OECD, 2022[31])). Additionally, accountability in downstream[12] applications of a model given its influence on downstream use may be challenging to disentangle.

As GenAI solutions become more commoditised, AI-driven tools and applications may become increasingly ubiquitous across financial service providers. This means non-qualified practitioners may be using such tools in the future without even knowing that these are driven by, or incorporate, AI technology. Any lack of awareness by practitioners of the risks associated with the use such AI tools could aggravate the risk profile of such tools, exposing them and their end customers to important risks. As such, educational efforts and suitability requirements may need to be considered as part of the governance frameworks for financial market practitioners wishing to use such tools.

Additional governance-related challenges are associated with outsourcing and third-party provision of services and infrastructure relative to AI models (for example, model providers or cloud service providers). Who is accountable for the adverse outcome of the model when the model was conceived and trained by a third party? This challenge involves also questions around intellectual property, as buying ‘off-the-shelf’ AI models does not grant the financial service provider intellectual property ownership, whilst the financial service provider feeds the model valuable proprietary data and business information that could be accessed by the third-party service provider. This relates to the model provider vs. deployer role differentiation and could add to the complexity of oversight and enforcement.

2.6. Financial stability-related risks: herding, volatility, flash crashes, interconnectedness and concentration

The deployment of AI in finance could involve potential financial stability risks related to one-way markets, market liquidity and volatility, interconnectedness and market concentration. The use of the same AI-driven model by a large number of finance practitioners could potentially prompt herding behaviour and one-way markets, which in turn may raise risks for liquidity and stability of the system, particularly in times of stress (OECD, 2021[1]).

[12] Referred to as the impact of the output of the model on subsequent actions, for foundation models.

AI in trading, for instance, could increase market volatility through large sales or purchases executed simultaneously, giving rise to new sources of vulnerabilities (FSB, 2017[32]). Convergence of trading strategies creates the risk of self-reinforcing feedback loops that can, in turn, trigger sharp price moves and pro-cyclicality. Herding behaviour by investors can also lead to bouts of illiquidity during times of stress and to flash crashes as observed in recent years in algo-high frequency trading.13 Such convergence also increases the risk of cyber-attacks, as it becomes easier for cyber-criminals to influence agents acting in the same way. The abovementioned risks exist in all kinds of algorithmic trading, and are amplified in cases of use of AI models that learn and dynamically adjust to evolving conditions in a fully autonomous way. This is the case of unsupervised learning-based GenAI models that have the capacity to learn, adjust their behaviour and front run based on the earliest of signals.

The deployment of AI in trading may also increase the interconnectedness of financial markets and institutions in unexpected ways, and potentially increase correlations and dependencies of previously unrelated variables (FSB, 2017[32]). It can also amplify network effects, such as unexpected changes in the scale and direction of market moves. The emergence of AI-as-a-Service providers - particularly those providing specific models - also may increase network interconnectedness (Gensler and Bailey, 2020[33]).

This also raises important risks of market dominance by a small number of model providers or user interface providers and the possible risk of concentration of the market, with various systemic implications. Concentration of data adds to these risks (Gensler and Bailey, 2020[33]). Possible operational failures of dominant players could have systemic implications for the markets, depending on the extent of usage of such models by financial market participants. In case of outsourcing, the dependency of financial market participants on third party providers of models would add an extra layer of fragility on top of the infrastructure dependence to the same third party providers (e.g. cloud).

2.7. Competition-related risks

In addition to the systemic implications of a concentration of activity in a small number of providers, there are competition issues related to market dominance in particular when it comes to GenAI models (e.g. refusal of access to models or data, barriers to switching). Given the enormous amounts of compute power and data required to develop and train GenAI models, activity could end up concentrated in a few players, particularly those with a first mover advantage or with the resources available to undertake design, training and maintenance of models. Financial market participants dependent on dominant third parties for their models could face the consequences of lack of competition, which could then be passed on to their customers (e.g. in terms of associated costs).

The cost and capacity requirements associated with GenAI models at the current stage of their development14 could also disadvantage entire nations without the economic resources to build, train and maintain their own models. On the user side, AI models of such scale may be reserved to larger financial market participants who have the capacity and resources to invest in such technologies.

Data concentration is another risk related to dominance of incumbents with cheaper or easier access to datasets (e.g. social platforms). Access to data is crucial for the success of AI models such as LLMs, and data concentration by BigTech or other platforms could exacerbate the risk of dominance of few large companies with excess power and systemic relevance.

[13] Algorithmically driven high frequency trading strategies appear to have contributed to extreme market volatility, reduced liquidity and exacerbated flash crashes that have occurred with growing frequency over the past several years (OECD, 2019[53]). Spoofing and other illegal market manipulation strategies, as well as collusion of ML models are additional risks of AI use in high frequency trading (OECD, 2021[1]).

[14] Presuming that processing power required for GenAI models is decreasing with advances in technology.

Additionally, possible adversarial usage of AI models could also be used to promote and sustain monopolies or oligopolies and suppress competition, undermining competitive market dynamics. For example, depending on the role of the model, it could be used to persuade and influence investor preferences.

2.8. Other risks

2.8.1. Risks related to employment and skills

Although there is significant level of uncertainty around the current and especially future impact of AI in the labour market, there is little evidence of significant negative employment effects due to AI so far (OECD, 2023[34]). This may be because AI adoption is still relatively low and/or because firms so far prefer to rely on voluntary workforce adjustments, therefore any negative employment effects of AI may take time to materialise (OECD, 2023[34]).

The use of AI could free up resources that could be used in higher value-added tasks, but at the same time there is a significant risk of impact in the job market in the future. GenAI in particular has a transformative potential for automation of a great variety of back-office and middle-office functions in finance (see Section 1). As such, the possible future widespread adoption of such models by the financial industry may give rise to some employment challenges.

The impact of AI on tasks and jobs is expected to engender changing skills needs (OECD, 2023[34]). The absence of adequate skills is a potential source of vulnerabilities for both the industry side and the regulatory/supervisory side, and which may give rise to potential employment issues in the financial industry. The deployment of AI in finance requires skillsets and understanding that a relatively small segment of financial practitioners possess at the moment. Inadequate usage or lack of awareness of the risks and unintended consequences of such models, particularly GenAI (given the relative ease of use), could lead to adverse impact for finance market participants and their clients.

2.8.2. Environmental impact

The computational needs of AI systems are growing, raising sustainability concerns (OECD, 2022[35]). GenAI and LLMs require enormous amounts of compute for their training, and are energy-consuming in terms of hosting and run inferences, with associated environmental impact that may need to be further examined. The same applies to data centres given the critical importance of data for the training of these models.

Similar to other innovative technologies used in finance, there is limited availability of credible data on the AI environmental footprint that can be used to inform possible policy discussions (OECD, 2022[31]; OECD, 2022[35]).

3 Policy considerations

The use of GenAI in finance has the potential to deliver important benefits to financial consumers and market participants, by producing efficiencies and improving customer welfare, but comes with great risks and challenges. Rapid developments in the area of AI and GenAI and its increasing relevance to financial markets calls for policy discussion and potential action to ensure the safe and responsible use of such tools. Financial regulators and supervisors have a role in ensuring that any deployment of GenAI in finance is consistent with the policy objectives of securing financial stability, protecting financial consumers, promoting market integrity [and fair competition].

3.1. OECD Principles on AI and select ongoing international efforts

The OECD Principles on AI, adopted in 2019, constitutes the first international standard agreed by governments for the responsible stewardship of trustworthy AI, remain highly relevant for the application of AI and GenAI tools in finance (OECD, 2019[26]). The AI Principles recognise the potential risks AI systems pose to human rights, privacy, fairness, and equality; robustness and safety; and the need to address these, such as by building transparency, accountability, and security into AI systems and enabling continuous monitoring and improvement (OECD, 2023[36]).

A number of national or regional initiatives have also been launched with the aim of providing guidance or guard rails for the safe and trustworthy development of AI across sectors. For example, in the EU, the EU Artificial Intelligence Act is the first proposal for a comprehensive legislative framework for AI across sectors, proposed in 2021 and voted by the European Parliament in June 2023 (European Commission, 2021[37]; European Parliament, 2023[38]). In the US, the White House Office of Science and Technology Policy has identified principles that should guide the design, use, and deployment of such automated systems in the Blueprint for an AI Bill of Rights (The White House, 2022[39]) and most recently an Executive Order has been issued on Safe, Secure, and Trustworthy Artificial Intelligence (The White House, 2023[40]).

At the G20 level, the financial stability implications of artificial intelligence and machine learning in financial services have been discussed by the Financial Stability Board in 2017 (FSB, 2017[32]), while the G7 in 2020 has analysed the cyber risks posed by artificial intelligence in the financial sector. Most recently, the G7 Leaders welcomed the Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems and the Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems (G7, 2023[41]; G7, 2023[42]; G7, 2023[43]).

Companies in the AI value chain play a critical role in identifying risks and addressing risks, but also using their leverage over business relationships to take action on AI risks and impacts. The OECD Guidelines for Multinational Enterprises on Responsible Business Conduct (MNE Guidelines) are a set of government-backed voluntary recommendations for business to proactively address potential harms they may cause, contribute to, or are directly linked to through business relationships (OECD, 2023[44]). The MNE Guidelines specifically recommend that companies carry out risk-based due diligence to identify and address any adverse impacts associated with their operations, their value chains or other business relationships. This approach to maximise the positive potential of business by first minimising the negative impacts forms the foundation for responsible business conduct (RBC). The OECD is currently working with governments and a multi-stakeholder expert group to develop tailored, concrete RBC guidance for actors in the AI value chain (OECD, 2023[45]).

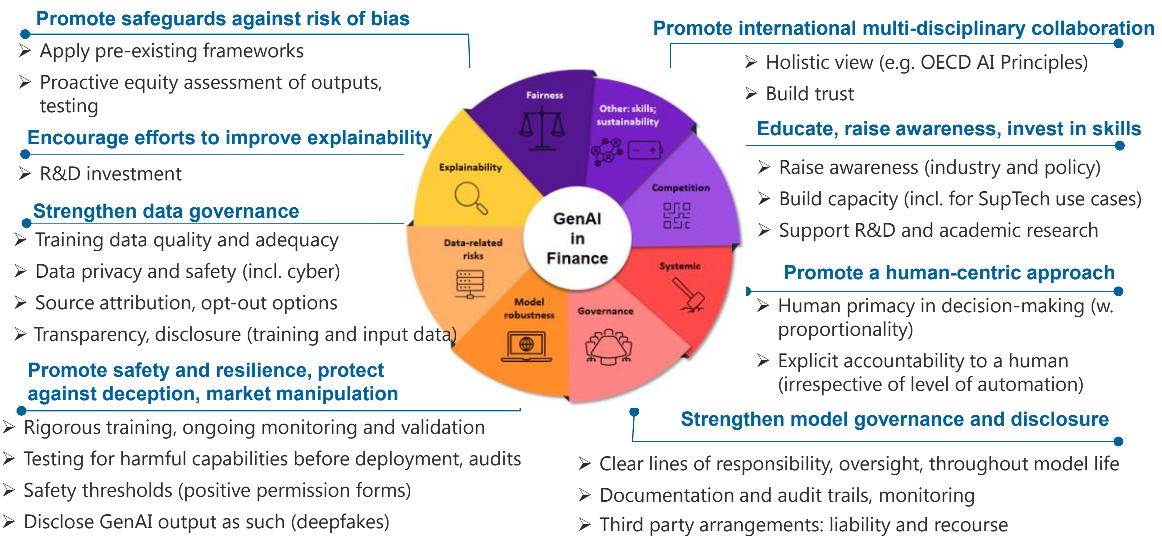
3.2. Policy consideration and potential action

Risks from the deployment of GenAI tools in finance will need to be identified and mitigated to support and promote the use of responsible and safe AI, without stifling innovation. The use of GenAI in finance exacerbates some of the 'generic' AI-related risks given its augmented capabilities, while it also raises a number of additional novel challenges associated with its specificities (e.g. deepfakes).

The application of existing guard rails applicable in AI models may need to be clarified and potentially adjusted to effectively address some of the novel challenges of advanced AI tools, if and where needed. Any perceived incompatibilities of existing arrangements with developments in AI may also need to be considered, such as the case of explainability in GenAI models.

Policy consideration and potential action could be looked at under a contextual and proportional framework, using on a risk-based approach depending on the criticality of the application (i.e. depending on the specific use-case) and the potential impact on the consumer involved (OECD, 2021[1]). Any guidance or policy will also need to be future proof to withstand the test of time given the rapid pace of development of innovation in the AI field.

Figure 3.1. Potential policy considerations to address GenAI risks in finance



Source: OECD authors' illustration.

Policy makers may need to consider reinforcing policies and strengthening defences and guard rails against risks emerging from, or exacerbated by, the use of GenAI (and other AI classes) in finance, focusing on a number of overarching areas (Figure 3.1). In particular:

Strengthen data governance practices by model developers and deployers: The importance of data is undisputed when it comes to the training of GenAI models and their usage by financial market participants. Best practices for data management and governance practices may be considered to ensure data quality, data adequacy depending on the intended use, data privacy when financial consumer data is fed into the model, and data authenticity and appropriate source attribution/copyrighting when applicable. This could include increased transparency, model documentation and reporting about the data used to train the model and any other data introduced into the model, including their location, origin and source attribution for copyrighted data used.[15] Depending on the model, the feasibility of data deletion options or obligations from models after a certain period of time could also be considered (similar to the 'right to be forgotten' of GDPR). This would need to include any data inputted in the model through prompts or otherwise, and the output of the model itself, given feedback loops for its self-training.

When private data are being used, consumers should have the right to opt-out from the use of their data for the training of GenAI models. This becomes particularly important in case the model can scrape data off the internet if it has web browsing capabilities or it if can link to the web in any way. The same considerations around data governance apply on databases purchased by third party providers, and on synthetic data generation based on public and private data.

Safeguards should be in place to overcome risk of bias and discrimination: Firms deploying such models should ensure that pre-existing fairness frameworks in financial services continue to apply. This could also involve proactive equity assessments of the models, impact assessment of model outputs, their sense checking against baseline datasets and other tests to ensure that protected classes cannot be inferred from other attributes in the data. The validation of the appropriateness of variables used by the model and of datasets used for training in terms of their representativeness are additional possible tools to reduce sources of potential biases. The latter applies in particular to LLMs, given potential current under-representation of minority languages in the training of language models (OECD, 2023[2]). Risks diagnosed should be followed by mitigating action and reporting of all the above could be conducive to strengthening user trust.

Encourage efforts to improve levels of explainability: Limited or outright lack of explainability, as could be the case in GenAI models, is a significant source of risks associated with the use of such models in finance (e.g. inability to adjust strategies in times of market stress). It may even be incompatible with existing laws and regulations, as for example the requirement to explain the basis for denial of credit extension to a prospective borrower in some jurisdictions. Progress made in the area of explainability of relatively simpler ML models will need to also be pursued in GenAI models with even greater complexity and lack of explainability. Improved explainability levels will be fundamental to build trust around the deployment of such tools in finance (and beyond).

Foster transparency and consider disclosure requirements depending on the case: Financial consumers should be informed about the use of AI techniques in the delivery of a product, when these have an impact on the customer outcome, as well as about machine-generated content and any potential interaction with an AI system instead of a human being. Financial consumers should also be informed about any collection and/or processing of their data for the purposes of the model and informed consent could be sought to that end. Customers should be offered the option to engage with a human if they so prefer. Active disclosure by financial market participants deploying such tools could be considered to ensure maximum awareness of the customer.

Disclosure requirements could include clear information, in plain language, around the AI system's functionalities and performance, including capabilities and limitations, as well as mitigating action taken to address such limitations. Description of the datasets used to train the model, including any copyrights, could help address data governance risks. Description of the results of any internal testing and independent external evaluation of the model and any impact assessment made (e.g. for disparity testing) could be considered as part of reporting to users. Manuals could be provided for downstream uses of models. Datapoints on the energy requirements for the model (for its training or use) could also be considered, in light of the absence of data around their environmental footprint. The governance framework of the model's development and deployment could also be part of such reporting. Transparency and disclosure will be even more critical in cases of GenAI models as a way to partly compensate for the lack

[15] Data quality and representativeness is to a large extent related to transparency and model documentation.

of explainability. Open benchmarking against academic reference models could provide a good starting point to promote transparency beyond explainability and model documentation.

Strengthen model governance and promote accountability mechanisms: Currently applicable frameworks for model governance in finance may need to be enhanced or adjusted to address incremental risks emerging from advances in AI. Solid governance arrangements and clear accountability mechanisms are fundamental in AI models deployed in high-value use-cases (e.g. in determining access to credit or investment advice). Parties involved in the development and deployment and such models should be held accountable for their proper functioning (OECD, 2019[26]). Explicit governance frameworks could include clear lines of responsibility and oversight throughout their lifecycle[16] and minimum standards or best practice guidelines to be followed. Documentation and audit trails for oversight and supervision should not be limited to the development process, and model behaviour and outcomes need to be monitored and tested throughout the model's lifetime.

Governance arrangements may need to include explicit attribution of accountability to a human irrespective of the level of automation of the model, with a view to also help build trust in AI-driven systems. In other words, this involves explicit accountability of the actor deploying the model for any harm caused by the model they are deploying. Contingency and security planning may also need to be considered to ensure business continuity. This could include the introduction of kill switches or other automatic control mechanisms, and back-up plans, models and processes in place to ensure business continuity in case the models fails or acts in unexpected ways (OECD, 2021[1]). Additional guard rails could be considered for the accountability of third- party providers of (foundation) models that are being adapted for downstream use cases or in other cases of outsourcing. Questions around recourse and legal liability of developers of such models could be also examined.

Promote safety, robustness and resilience of GenAI models (including for cyber risk) and mitigate risks of deception and market manipulation: Frameworks for appropriate training, retraining and rigorous testing of AI models, and their ongoing rigorous monitoring and validation could be the most effective ways to improve model resilience, prevent and address drifts, and ensure the model performs as intended. Monitoring and validation could include independent reviews and external audits both at the development and during deployment, and documentation of each such processes could facilitate supervision. Ongoing monitoring is particularly important for GenAI models that are based on autonomous unsupervised learning and where false or inaccurate information introduced in the model post deployment feeds the model and continues to be part of the model in future loops (e.g. user prompts). Also, datasets used for training, especially when synthetic, need to be large enough to capture non-linear relationships and tail events in the data to cover for unprecedented crises events, and stress testing for such scenarios could be performed.

Testing for dangerous or harmful capabilities of a model before its deployment could be used to understand the ability of the model to act in adversarial ways (e.g. proliferation of misinformation) and to adjust the models' behaviour to account for the results of such tests. Depending on the capabilities of the model, and the results of such impact assessments prior to deployment, content filtering and other restrictions could be introduced upfront to the model based on safety thresholds (e.g. refusal of harmful requests by design). Alternative options to be considered could include positive permission forms of design (i.e. do not do unless it's permitted). Hackathlons or 'red teaming'[17] are other initiatives that could be useful in improving the robustness of production models by explicitly targeting their weaknesses and comparing the results with desired outcomes. GenAI-generated output needs to be explicitly disclosed as such, in order to limit the risk of deepfakes and promote the truthfulness of the model's output. In case of

[16] Design, development, deployment of the model.

[17] Attack simulations or other methods designed to measure how well the system performs in a variety of adversarial scenarios.

large models above a certain level of capabilities that could be considered systemically important, adherence to commonly agreed sets of safety requirements could be envisaged.

Encourage a human-centric approach and place emphasis in human primacy in decision making, particularly for higher-value use cases (e.g. lending): An appropriate degree of human involvement in AI and GenAI-assisted financial market activity may need to be ensured to minimise the risk of harm to individual customers, depending on the criticality of the use case. End customers need to be informed about the involvement of AI in the provision of their service and could have the right to object to its use, opt out of AI-assisted products or services and of the AI model's reach (e.g. for data usage). Customers may need to be given the right to request a human intervention, or challenge the outcome of the model and seek redress. In addition to a mandatory human alternative option for the end customer, humans would also need to be ready to act as a human safety net in case of model disruption to ensure business continuity, avoiding over-reliance of firms in AI-based systems. Keeping the 'human in the loop' can also help build confidence and trust in the use of AI in finance. What is more, due consideration should be given to the impact of the use of GenAI in finance on social and environmental well-being.

Invest in R&D, skills and capacity to keep pace with advances in AI, raise awareness of the perils of GenAI and create tools to mitigate some of the associated emerging risks (e.g. hallucinations). Both the public and the private sector will need to deploy resources to invest in research, build skills and raise the awareness of financial market participants and policy makers around the risks of advanced AI models such as GenAI and LLMs. R&D investment could provide solutions and tools to mitigate issues of explainability and mitigate risks of GenAI models (e.g. identify and prevent deceptive outputs). Research is important to ensure safety of future scenarios of fully autonomous models (e.g. AGI). Investment in education and skills in the industry could enable effective AI model governance, while also guiding practitioners and consumers towards safer deployment of such models. Policy makers would also need to keep pace with advancements in AI technology in order to be technically able and prepared to oversee such activity in finance and/or intervene as required. Importantly, the upskilling of policy makers will also allow them to benefit from RegTech/SupTech solutions for effective and efficient supervision of financial market activity more broadly.

In line with the OECD AI Principles, there is a need to promote international multi-disciplinary and multi-stakeholder cooperation (OECD, 2019[26]). Given the transformative effect of AI across sectors of economic activity and across society, there is a need for a holistic view and objective for trustworthy and safe AI deployment. This, in turn, warrants multi-disciplinary and multi-stakeholder dialogue and cooperation, such as the one underpinning the OECD AI Principles. Industry-led commitments, public education and financial education, as well as communication efforts are additional ways to build awareness of benefits and perils and instil trust and confidence in the safe adoption of this transformative innovation in finance and beyond. Support for policy coherence and interoperability among burgeoning AI risk management and accountability frameworks will also be critical given the global nature of financial products and services.

References

Ahmed, E. et al. (2017), "The role of big data analytics in Internet of Things", *Computer Networks*, https://doi.org/10.1016/j.comnet.2017.06.013. [4]

Blitz, D. et al. (2023), "The Term Structure of Machine Learning Alpha", *SSRN Electronic Journal*, https://doi.org/10.2139/SSRN.4474637. [14]

Bommasani R. et al. (2023), *Do Foundation Model Providers Comply with the Draft EU AI Act?*, https://crfm.stanford.edu/2023/06/15/eu-ai-act.html (accessed on 2 August 2023). [30]

Bubeck, S. et al. (2023), "Sparks of Artificial General Intelligence: Early experiments with GPT-4", https://arxiv.org/abs/2303.12712v5 (accessed on 27 July 2023). [47]

Buolamwini, J. (2018), "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification *", *Proceedings of Machine Learning Research*, Vol. 81, pp. 1-15. [24]

CBInsights (2023), *The generative AI landscape: Top startups, venture capital firms, and more*, https://www.cbinsights.com/research/generative-ai-funding-top-startups-investors/ (accessed on 28 July 2023). [5]

Chen, L., M. Zaharia and J. Zou (2023), "How Is ChatGPT's Behavior Changing over Time?". [28]

EDPS (2021), *Synthetic Data | European Data Protection Supervisor*, https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en (accessed on 3 August 2023). [20]

ESMA (2023), "ESMA TRV Risk Analysis Artificial intelligence in EU securities markets", https://doi.org/10.2856/851487. [21]

European Commission (2021), *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legistative Acts*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (accessed on 3 August 2022). [37]

European Parliament (2023), *Artificial Intelligence Act | Legislative Train Schedule*, https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence (accessed on 3 August 2023). [38]

Federal Reserve Board (2023), "Cybersecurity and Financial System Resilience Report - Board of Governors of the Federal Reserve System", http://www.federalreserve.gov/aboutthefed.htm. (accessed on 7 August 2023). [25]

Freyberger, J. et al. (2020), “Dissecting Characteristics Nonparametrically”, *Review of Financial Studies*, Vol. 33/5, pp. 2326-2377, https://doi.org/10.1093/RFS/HHZ123. [12]

FSB (2017), *Artificial intelligence and machine learning in financial services Market developments and financial stability implications*, http://www.fsb.org/emailalert (accessed on 1 December 2020). [32]

FT (2013), *Norway’s $1.4tn wealth fund calls for state regulation of AI | Financial Times*, https://www.ft.com/content/594a4f52-eb98-4da2-beca-4addcf9777c4 (accessed on 26 July 2023). [23]

G7 (2023), *G7 Leaders’ Statement on the Hiroshima AI Process*, https://www.mofa.go.jp/files/100573466.pdf (accessed on 7 November 2023). [41]

G7 (2023), “Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems On the basis of the Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI systems, the Hiroshima Process International Code of Conduct for Organizations”. [43]

G7 (2023), “Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System The Hiroshima Process International Guiding Principles for Organizations Developing Advanced”. [42]

Gensler, G. and L. Bailey (2020), “Deep Learning and Financial Stability”, *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3723132. [33]

Goertzel, B. (2014), “Artificial General Intelligence: Concept, State of the Art, and Future Prospects”, *Journal of Artificial General Intelligence*, Vol. 5/1, pp. 1-48, https://doi.org/10.2478/JAGI-2014-0001. [46]

Groh, M. et al. (2022), “Deepfake detection by human crowds, machines, and machine-informed crowds”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 119/1, https://doi.org/10.1073/PNAS.2110013119. [29]

HAI (2023), *Stanford CRFM*, https://crfm.stanford.edu/ (accessed on 27 July 2023). [48]

Ji, Z. et al. (2023), “Survey of Hallucination in Natural Language Generation”, *ACM Computing Surveys*, Vol. 55/12, https://doi.org/10.1145/3571730. [52]

Karim, R. (2019), *Illustrated: Self-Attention*, Towards Data Science, https://towardsdatascience.com/illustrated-self-attention-2d627e33b20a (accessed on 26 May 2023). [7]

Kotu, V. and B. Deshpande (2019), “Anomaly Detection”, *Data Science*, pp. 447-465, https://doi.org/10.1016/B978-0-12-814761-0.00013-7. [11]

Kotu, V. and B. Deshpande (2019), “Deep Learning”, *Data Science*, pp. 307-342, https://doi.org/10.1016/B978-0-12-814761-0.00010-1. [3]

Kotu, V. and B. Deshpande (2019), “Recommendation Engines”, *Data Science*, pp. 343-394, https://doi.org/10.1016/B978-0-12-814761-0.00011-3. [19]

KPMG (2023), “Generative AI: From buzz to business value”. [10]

Lazard (2023), *SUV Report - Lazard VGB Insights - 2023 Outlook*, https://vgb.lazard.com/lazard-vgb-insights-2023-outlook/#FinTech (accessed on 28 July 2023). [15]

Lesher, M., H. Pawelec and A. Desai (2022), *Disentangling untruths online : Creators, spreaders and how to stop them | OECD Going Digital Toolkit Notes | OECD iLibrary*, https://www.oecd-ilibrary.org/science-and-technology/disentangling-untruths-online_84b62df1-en (accessed on 2 August 2023). [51]

McKinsey (2022), *The state of AI in 2022—and a half decade in review | McKinsey*, https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review (accessed on 28 July 2023). [6]

MIT Technology Review (2023), *GPT-4 is bigger and better than ChatGPT—but OpenAI won't say why | MIT Technology Review*, https://www.technologyreview.com/2023/03/14/1069823/gpt-4-is-bigger-and-better-chatgpt-openai/ (accessed on 7 August 2023). [9]

Moritz, B. and T. Zimmermann (2016), "Tree-Based Conditional Portfolio Sorts: The Relation between Past and Future Stock Returns", *SSRN Electronic Journal*, https://doi.org/10.2139/SSRN.2740751. [13]

NIST (2023), "Artificial Intelligence Risk Management Framework (AI RMF 1.0)", https://doi.org/10.6028/NIST.AI.100-1. [27]

OECD (2023), *Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI | en | OECD*, https://www.oecd.org/sti/advancing-accountability-in-ai-2448f04b-en.htm (accessed on 1 December 2023). [36]

OECD (2023), *AI language models: Technological, socio-economic and policy considerations | en | OECD*, https://www.oecd.org/publications/ai-language-models-13d38f92-en.htm (accessed on 27 July 2023). [2]

OECD (2023), *Common guideposts to promote interoperability in AI risk management | OECD Artificial Intelligence Papers | OECD iLibrary*, https://www.oecd-ilibrary.org/science-and-technology/common-guideposts-to-promote-interoperability-in-ai-risk-management_ba602d18-en (accessed on 1 December 2023). [45]

OECD (2023), *OECD Employment Outlook 2023 : Artificial Intelligence and the Labour Market | OECD iLibrary*, https://www.oecd-ilibrary.org/sites/08785bba-en/index.html?itemId=/content/publication/08785bba-en (accessed on 3 August 2023). [34]

OECD (2023), *OECD Guidelines for Multinational Enterprises on Responsible Business Conduct*, OECD Publishing, Paris, https://doi.org/10.1787/81f92357-en. [44]

OECD (2022), *Environmental impact of digital assets: Crypto-asset mining and distributed ledger technology consensus mechanisms | en | OECD*, https://www.oecd.org/publications/environmental-impact-of-digital-assets-8d834684-en.htm (accessed on 14 March 2023). [31]

OECD (2022), *Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint | en | OECD*, https://www.oecd.org/publications/measuring-the-environmental-impacts-of-artificial-intelligence-compute-and-applications-7babf571-en.htm (accessed on 3 August 2023). [35]

OECD (2021), “Algorithms and Collusion: Competition Policy in the Digital Age”. [17]

OECD (2021), *Artificial Intelligence, Machine Learning and Big Data in Finance Opportunities, Challenges and Implications for Policy Makers*, https://www.oecd.org/finance/artificial-intelligence-machine-learningbig-data-in-finance.htm. (accessed on 22 August 2022). [1]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en. [26]

OECD (2019), “Strenghtening Trust in Business OECD Business and Finance Outlook 2019”. [53]

OpenAI (2023), “GPT-4 Technical Report”. [8]

Papenbrock, J., N. GmbH and J. Ashley (2022), “Accelerated Data Science, AI and GeoAI for Sustainable Finance in Central Banking and Supervision 1”. [22]

Stephanie Kay Ashenden (2021), “The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry”, *The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry*, https://doi.org/10.1016/C2019-0-01262-9. [50]

The White House (2023), *FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence | The White House*, https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/ (accessed on 7 November 2023). [40]

The White House (2022), *Blueprint for an AI Bill of Rights | OSTP | The White House*, https://www.whitehouse.gov/ostp/ai-bill-of-rights/ (accessed on 3 August 2023). [39]

Weizenbaum, J. (1966), “ELIZA a computer program for the study of natural language communication between man and machine”, *Communications of the ACM*, Vol. 9/1, pp. 36-45, https://doi.org/10.1145/365153.365168. [18]

Whitesight (2023), *Generative AI in Finance and Beyond - WhiteSight*, https://whitesight.net/generative-ai-in-finance-and-beyond/ (accessed on 28 July 2023). [16]

Yang, D. et al. (2020), “Segmentation using adversarial image-to-image networks”, pp. 165-182, https://doi.org/10.1016/B978-0-12-816176-0.00012-0. [49]