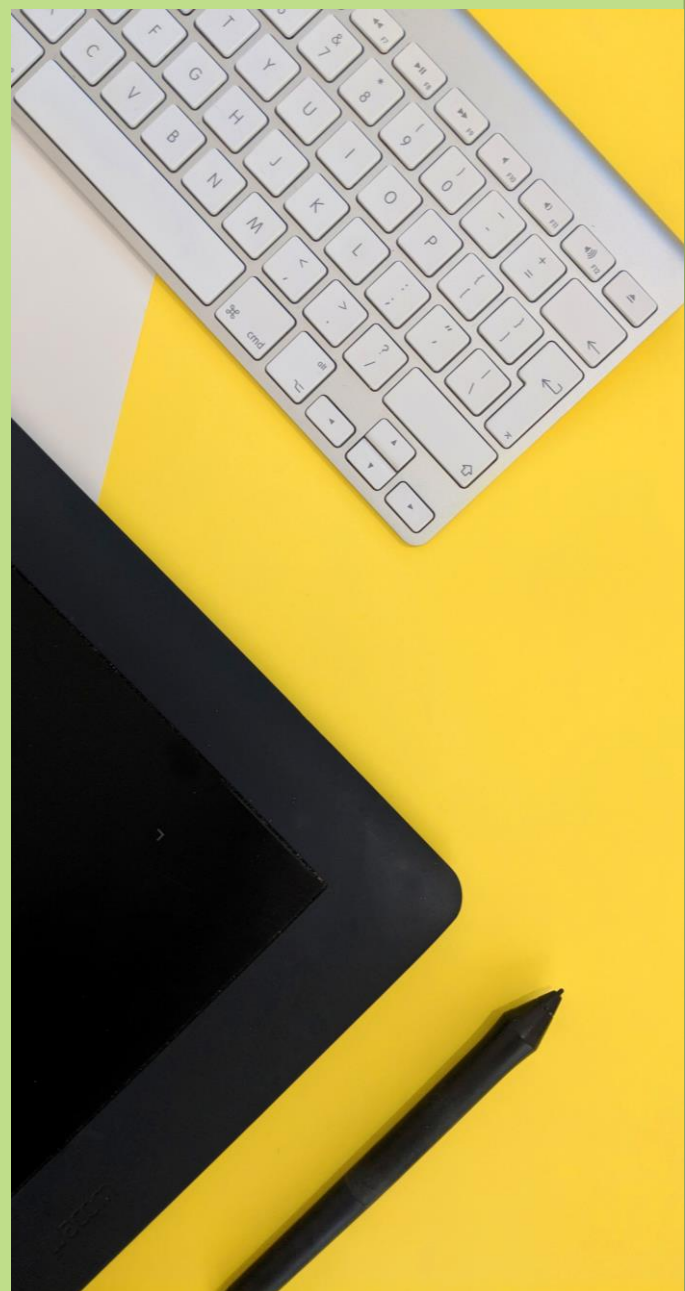


Bird & Bird

AI in FS: risks, opportunities, regulation

Step aside blockchain and the Metaverse. 2023 is the year of artificial intelligence (AI)!



Introduction

After decades of more gradual evolution, AI is having its breakthrough moment. Or should we say “moments” given the sheer number of breakthroughs and innovations being publicised. These breakthroughs have been facilitated by three key factors: greater availability of vast datasets to train AI systems, greater computational power of the underlying computers of these systems including new AI architectures (e.g. transformers) and a wider availability of AI talent and resource. But a key recent development has been the combination of AI learning with natural language referred to as “generative AI” and the creation of large language models (**LLMs**) such as OpenAI’s GPT-4 and Google’s Bard.

AI has the potential to transform the financial services (**FS**) sector by replacing routine tasks requiring the analysis of vast amounts of data and so potentially leading to faster and more accurate outputs. For example, it was recently reported that Morgan Stanley, an international investment bank, had been training an advanced generative AI chatbot based on OpenAI’s latest LLM technology on selected pieces of data vetted by it with the aim of creating a tool that could assist its financial advisors when advising wealth management clients. This tool will provide these advisors with insights at a much greater scale and speed than a human could do, resulting in the potential transformation of the wealth management industry.

But the opportunities for firms authorised in the FS sector (**Firms**) needs to be balanced against the risks. For example, how do we deal with bias in the training data being used to train AI systems, copyright infringement issues and data protection concerns? How do we deal with the risk that these AI systems’ capabilities outrun their creators’ understanding resulting in systemic risks to the financial system and detrimental outcomes to consumers? How does existing FS regulation apply to this new technology? Is the existing regulatory framework sufficient to address AI risks or is a bespoke framework required?

There is a lot to unpack so we’ve broken down this article into two parts:

- In Part 1 we examine what AI is, how it works and some potential uses cases for the FS sector.
- In Part 2 we examine some of the practical legal, regulatory and commercial risks in-house lawyers working in the FS sector need to think about.

¹ [Machine learning in UK financial services \(bankofengland.co.uk\)](https://www.bankofengland.co.uk/machine-learning-in-uk-financial-services)

PART 1

AI, ML, DL, generative AI – what’s it all mean?

AI



“Artificial Intelligence (AI) is the theory and development of computer systems able to perform tasks which previously required human intelligence.”

There is no consensus on one definition but, generally speaking, AI is a field of computer science relating to the creation of computer systems that perform specific tasks at least as well as most humans (**AI systems**). AI is generally distinguished from artificial general intelligence (**AGI**) which are those systems, which currently do not exist, which can perform all tasks at least as well as most humans.

Machine learning



“Here’s some data, which pattern does it match?”

Machine learning is a subset of AI relating to how AI systems can perform a task without being specifically programmed to perform that task – they do this by creating mathematical models that identify patterns in data and then apply those mathematical models to new data (that it wasn’t trained upon). These systems have been trained to recognise patterns on a massive scale and account for the first major wave of AI systems advancement occurring from around 2013/2014.

Machine learning AI systems can create their own mathematical models for finding patterns in data which are better than those which humans could manually program into them. A well-known example in the context of pattern recognition is how do you tell a computer to spot the difference

between a cat and dog when there are hundreds of species of each, with very different appearances?

Machine learning - how does it work?

In machine learning the AI system turns a problem like image recognition (solving a puzzle such as does this image look more like a dog or a cat) into a statistical problem (is the arrangement of pixels in this image more likely to correspond to patterns of pixels which prior training has shown to be statistically associated with cats or patterns of pixels which are statistically associated with dogs).

Machine learning involves feeding an AI system with vast amounts of “training data” (which may be pre-labelled or unlabelled). Training an AI system on pre-labelled training data is called supervised learning and training it on unlabelled training data is called unsupervised learning. For example, in relation to supervised learning, an AI system is fed with lots of pictures labelled “dog” and “cat”. The AI system ingests this data and then, using an artificial neural network (see below), learns to identify patterns in the training data that, in our example above, distinguishes images of cats from images of dogs (this process is called “training”). It is important to understand that these “learnings” are just numbers (also referred to as the weights – see below) in the neural network that together comprise a mathematical model that is fine-tuned as part of training rather than any easily understandable human rule (for example, “dogs are generally bigger”) – this gives rise to the “black-box” problem of AI which is the difficulty of understanding the inner workings of an AI system (and *why* it decides that a particular image depicts a dog).

Once the AI system has been trained a mathematical model is created and you can then give it a new picture of a cat and the AI system uses this mathematical model to make a statistical prediction as to how likely or not the new picture fits the dog pattern or the cat pattern. The hope, of course, is that the patterns which the system has learned during training will apply to cases which it has never seen before (so that it accurately predicts that an image depicts a dog even though it has never seen that image before).

Artificial Neural Networks

Neural networks are algorithms commonly used in machine learning. They are called “artificial neural networks” because they are loosely inspired by how (we think) neurons in brains work. An artificial neural network contains lots of numbers or ‘weights’ between each artificial neuron (similar to how a brain has lots of neurons and connections or “synapses” between those neurons) that can be adjusted during training to determine how input data gets transformed to outputs (and so to maximise the likelihood of a correct prediction). The weights in artificial neural networks are randomly initialised but then get better as the artificial neuron is trained on a specific task. You can think of the weights as the learnings the artificial neural network (or digital brain) acquires from being trained on the training data. As mentioned above, these weights or learnings create a mathematical model which, when trained properly, has learned to identify certain patterns in data and then can be used to receive new data and provide a statistical prediction as to whether the new data matches the pattern it has been trained to find.

Deep Learning

Deep learning is a field of machine learning that uses neural networks which have at least three “layers”. Deep learning is a particularly powerful form of machine learning which excels at pattern recognition and classification tasks.

Generative AI



“Take this pattern and make me more new patterns like this.”

Think of generative AI (“generative” because it is generating something) as taking the machine learning model to identify patterns from training data and then reversing it to create new data using those patterns. So, the AI system divines a pattern based on its training data and now you want the AI system to make new data that matches this pattern. The AI system is automating the creation of something entirely new

– contracts, code, images – based on similar data on which it has been trained.

This is an even more advanced form of AI that has seen significant progress since 2021, and hit the mainstream during Q3/Q4 of 2022.

ChatGPT (an AI chatbot interface powered by OpenAI’s GPT-3.5 or GPT-4 LLM) is an example of a generative AI system that can generate new combinations of text in the form of essays, articles, poems or songs based on input prompts and parallel training data available to it (e.g. existing written material, coding data or musical compositions).

Generative AI models such as GPT 4 which powers ChatGPT are called LLMs. LLMs are advanced mathematical models used for language generation which include an artificial neural network.

Generative AI - how does it work?

The key development behind LLMs is the use of “transformer” models. Transformers are a type of artificial neural network that can “learn” context and meaning by tracking relationships in sequential data like words in a sentence. The LLM has an “attention network” that analyses its training data and makes connections or associations between different words which enables it to understand how language is structured and then it tweaks the weights in its neural network in light of this training.

For example, an LLM will receive input data and quiz itself on the text. It does this by taking a chunk of the data and covering up some words at the end and then guessing what might go there. The LLM uncovers the answer and compares it to its guess. The answers are the data itself so it can be trained in a self-supervised manner on massive datasets without human labellers. The model’s goal is to make its guesses as good as possible with minimum errors. The artificial neural network then iteratively updates its weights to ensure it produces answers with fewest errors.

As a result of this training, pairs of words are given a weight indicating how much the model should pay attention to one pair when processing the other – this helps the model form connections between words. Once trained for a specific task (in this example the LLM is providing text responses to text prompts), the LLM will have created a statistical model that helps it predict the most likely sequence of words to select as a response to a prompt. So, if you give the LLM a prompt it will provide an answer based on the

relationships it has identified between the words in its model and the probability of which words should come together based on the model's weightings (which are fine-tuned as part of its training) to form the output answer

LLMs are now being used in combination with other software applications (referred to as AI agents) in what is being referred to as the next evolution of AI. The AI agent sits on top of (interoperates with) the AI system (e.g. ChatGPT). For example, rather than the user having to type in the correct prompts to the AI system to get the right solution it can input its problem into the AI agent and then the AI agent will handle the prompt engineering (i.e. sending the correct prompts to the AI system in order to get the best answer to the requested problem).

FS use cases

It is certain that, in the near future, businesses will make increasing use of AI (perhaps starting with AI-powered customer service chatbots but moving on to much more sophisticated applications). There are, however, many use cases of AI which are specific to the FS sector given the treasure trove of data they can use to help train AI systems. Typical applications for the use of AI will include:

- Insurance –to price premiums, make underwriting decisions and to assess claims.
- Investment management –to make investment recommendations to clients.
- Derivatives –to price derivative contracts.
- Credit –to make underwriting and affordability decisions.
- Fraud detection –to screen for fraudulent transactions.
- Regulatory compliance – to make sure that a firm is dealing with the right people in the right way

Let's take one use case and expand on it. A bank could build a machine-learning AI system to help it identify potential instances of money laundering or terrorist financing. It could provide the AI system with vast amounts of transaction data labelled as "suspicious" (i.e. these transactions that have previously triggered a suspicious activity report or SAR) and vast amounts of transaction data labelled "not suspicious" (i.e. these transactions have not triggered a SAR). The AI system learns from the data and builds up a picture of what hallmarks in the data are needed to fit the "suspicious data pattern" and the "not suspicious pattern". The AI system is then fed new transaction data and identifies which pattern the input data most likely fits (i.e. "suspicious" or not). This could not only help the bank save time and money identifying suspicious behaviour using human resources, but could also potentially identify suspicious activity more quickly and effectively.

Bloomberg has been developing a generative AI model for the finance industry based on a freely available, off the shelf AI model that it has then trained on its own bank of proprietary financial data. Bloomberg plans to offer the generative AI system (called BloombergGPT) as a potential add-on to its existing Bloomberg Terminal systems although no clear plans have been articulated yet. We can definitely see the benefit of an AI chatbot trained on specific financial data being able to be questioned and provide answers to almost any financial question. This will no doubt speed up decision making processes and increase efficiencies in many businesses.

These AI systems, once trained on an FS company's data, will help cut costs, save time and bring new insights to FS companies. But, what are the legal, commercial and regulatory issues FS companies need to consider before utilising these solutions? We explore these issues further in Part 2 below.

PART 2

There are a number of practical legal, commercial and regulatory issues that need to be considered by firms seeking to implement AI-based solutions in the FS sector.

Confidential and sensitive information

AI systems ingest inputs or prompts provided to them and generate outputs based on these inputs. Care should be taken to ensure users do not upload confidential information into the AI system as, depending on the terms and conditions of the relevant AI system, it could retain such data to fine-tune the AI system (this is a process of taking the additional data provided in the input and using it to improve the AI model by adding it to the original training data). This could result in such input information being replayed back to new users (who could be competitors) in the form of outputs. This could also enable an improved AI system to be deployed by a competitor organisation that is using the same tool.

OpenAI in its terms of use dated March 2023 refers to such inputs and outputs as “Content”. The terms of use do helpfully differentiate between use of the AI system via API and non-API (e.g. browser). In relation to API access to the system, the terms make clear that OpenAI does not use the Content provided to develop or improve the services. However, for non-API use then the AI system will use such Content to help develop and improve the services and so greater vigilance is required as this may mean the Content is reviewed and incorporated into the AI systems training data meaning it could be replayed back to new users in the form of outputs.

If you are a user, consider if you can include protections in your terms. For example, you may agree that your inputs cannot be used other than to provide the services to you and must be returned or deleted on termination or expiry of the agreement but you may agree that any learnings (including machine learnings) acquired by the AI system in using such inputs may be retained by the AI system subject to compliance with confidentiality provisions. This means the AI system can retain any improvement to the AI system (e.g. updates to weights in the artificial neural network) following ingestion of the data but not the data itself.

Data protection

The use of AI to process personal data (including in the designing, training and testing phase as

well as in the deployment of an AI system) requires careful consideration of various data protection and privacy issues. The use of AI to process personal data often triggers the need for a data protection impact assessment.

Data protection regulators are increasing their focus on this area. It is a clear strategic priority for the UK Information Commissioner (**ICO**). The ICO’s ICO25 strategic plan includes actions to tackle AI-driven discrimination on the basis that this is an issue that can have damaging consequences for people’s lives. The ICO has published guidance on AI and data protection as well as an AI and data protection risk toolkit. The French data protection authority (*Commission Nationale de l’Informatique et des Libertés*) recently announced that in 2023 it will expand its existing work to generative AIs, LLMs and derived applications (especially chatbots). In relation to ChatGPT, we have seen recent action by the Italian data protection authority (*Garante per la protezione dei dati personali*) and an announcement that the Spanish data protection authority (Agencia Española de Protección de Datos) intends to carry out a preliminary investigation.

The technical complexities of AI systems and the unique risks involved can make data protection issues particularly challenging to deal with and raises a whole host of questions. From an EU and UK perspective for example:

How do you ensure fairness in the use of AI? How do you address the risk of bias and discrimination? Where you use an AI system to make inferences about people, do you ensure that the system is sufficiently statistically accurate for your purposes?

- How do you embed data protection by design and default from design to deployment of an AI system?
- What lawful basis applies to each activity involving personal data (including any relevant research and development, training, testing and design as well as deployment)?
- Is any proposed automated decision taking (i.e. automated decisions without human involvement, including profiling) lawful taking

into account the particular restrictions under the law?

- How will you explain the processing under an AI system to data subjects in a clear and transparent way? How will you provide meaningful information about the logic involved in automated decision taking?

Copyright

Developers of generative AI systems will want to consider the extent to which developing their system could give rise to copyright infringement, or infringement of other third party rights such as confidentiality, contractual restrictions or database rights. These third party rights may attach to training data and using that data without permission could give rise to liability. In some jurisdictions exceptions to copyright may apply, such as fair use in the US. However the rules vary between jurisdictions and depend on the factual circumstances in which the AI training is taking place.

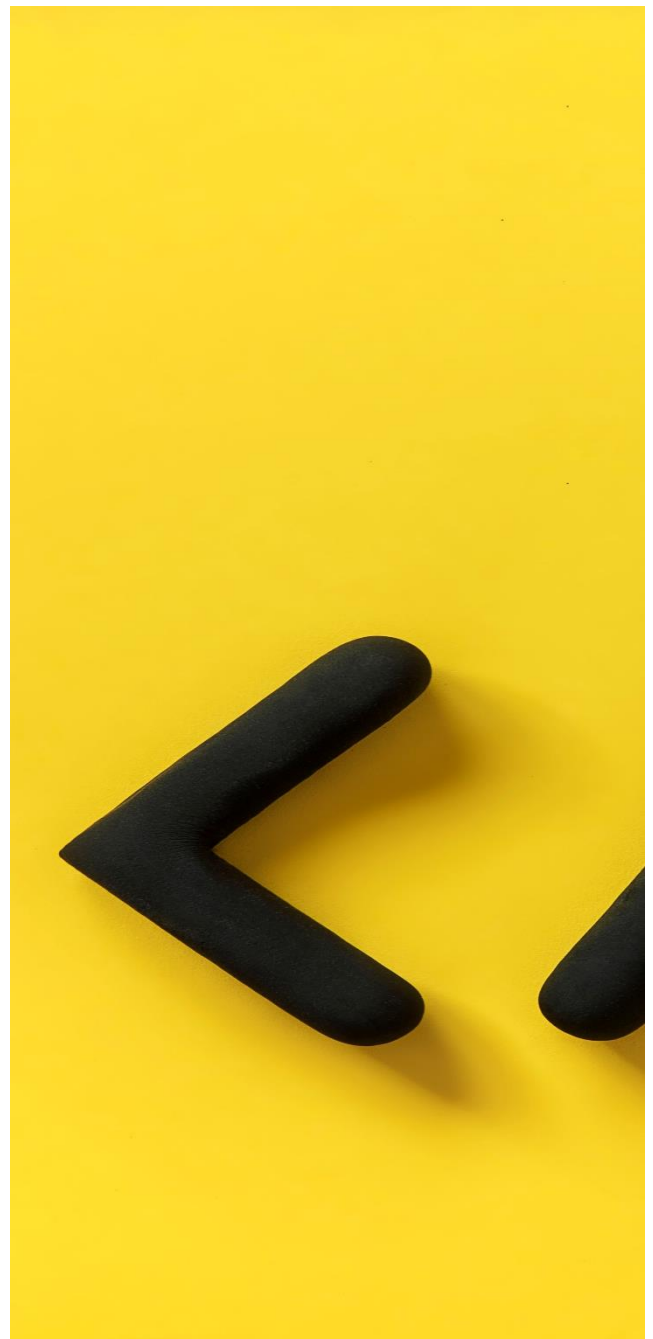
Aside from developers, users of generative AI systems also need to consider copyright risks relating to inputs (e.g. prompts) and outputs.

For prompts, AI systems will typically include language in their terms of use stating that the user must not through its use of the service infringe a third party's IP rights (e.g. by uploading input data in breach of a third party's IP). If the user does this, they will breach the terms and conditions and typically there will be a third party IP infringement indemnity in favour of the AI service provider linked to a breach of such a term that it can enforce against the user.

For outputs, users will need to navigate the risk that their output could be accused of reproducing all or part of a third party copyright work. Whether an output which looks similar to a piece of training data is an infringement of any copyright in that training data is a hotly contested legal and technical issue. However, in the past accusations of copyright infringement could be defeated by demonstrating that your work was independently created without reference to the earlier work. This becomes more challenging for a user of a generative AI system, as you are unlikely to know how the system has been trained and what training data has been used in the process. While users may want to seek protection from the developer under the terms of use, the current market position is generally for liability for any IP infringement by outputs to remain with the user.

Output issues also arise in relation to the copyright status of outputs from AI systems.

Copyright protects original intellectual creations. Entering a simple text prompt will typically be insufficient to grant the user copyright in the output. Without copyright, outputs would need to be protected through some other means such as imposing contractual and/or confidentiality restrictions on recipients. While terms of use typically provide that any copyright which does arise will belong to the user, this cannot by itself create copyright where the output is not an intellectual creation. Things may be different however if outputs are edited (either directly or via prompt engineering) or there is a selection or arrangement of multiple outputs.



Garbage in, garbage out

The quality of the training data is very important as it will then impact on the quality of the outputs provided by the AI system. This is because the output of AI systems is based on the what the AI system has seen in its training data.

Linked to this point is the fact that most AI system terms will make clear the system is not 100% accurate and outputs can provide misleading or incorrect information. Users should be aware of these disclaimers meaning they will have limited rights to bring claims against the provider for their reliance on the outputs. In order to mitigate this, a human should be involved to carefully verify the accuracy of the outputs generated before relying on them e.g. before relying on a report generated by a AI systems trained to find incidents of suspicious activity the human should review the output and confirm its accuracy.

AI bias is also an important concern so appropriate technical due diligence should be undertaken on the AI systems including, where possible, its mathematical models and training data to ensure the data used to train it is not biased in a way that can lead to unintended consequences. For example, a machine learning AI system can be provided with CVs of previous, successful hires at a company and be trained to find more CVs like these hires – simple pattern recognition. However, what if all the CVs of previous hires were of men. This would mean, going forward, the AI system would be taught to find CVs that fit this pattern and this would risk it discounting CVs of very capable women as the CVs don't match the pattern it was trained on.

Regulation

In the UK alone there has been some interesting developments following the publication of a variety of papers over the last 3 years in relation AI adoption. An analysis of these papers is beyond the scope of this article but FS organisations should take time to familiarise themselves with them:

- In October 2020 the Bank of England and the FCA established the AI Public-Private Forum (**AIPPF**) to further dialogue on AI innovation and safe adoption in the FS sector. The AIPPF published its final report in 2022 exploring the various barriers to adoption, challenges and risks related to AI in the FS sector. The report made it clear that the private sector wants regulators to have a role in supporting the safe adoption of AI in UK FS sector.
- In October 2022, in response to the AIPPF final report, the Bank of England issued a Discussion Paper on AI and ML. The discussion paper raised some interesting points on the key question surround this technology – how do we regulate it? Can we rely on clarifications to existing regulations or does a new approach need to be adopted for this new technology. It also posed various questions to solicit feedback from the industry on the best next steps for regulating AI.
- Finally the UK government recently published a policy paper on “Establishing a pro-innovation approach to regulating AI” which advocates a light-touch approach by applying existing regulatory regimes to the regulation of AI rather than the creation of a new framework

The FCA and the European Securities and Markets Authority published a number of papers on the use of robo-advice tools which uses automation and algorithms to assess a person's suitability and can be used to assist firms with selecting certain financial instruments based on this assessment. A key risk identified by the regulators was in relation to herding.

Outside of these papers there are a variety of regulations and guidance that firms authorised in the FS sector (**Firms**) should consider prior to procuring services from AI providers. We've listed out below a selection of the key regulations and/or guidance to consider.

DORA

On December 2022 the EU's Digital Operational Resilience Act (**DORA**) was published in the Official Journal of the European Union and will enter into force on 16 January 2023. This is a far-reaching piece of legislation that will impose new obligations on both financial services organisations and critical third party service providers. DORA will require new systems and controls, risk frameworks and new contractual provisions to be included in ICT focused outsourcing agreements. Given the amount of time it will take for in-scope firms to be compliant with DORA, the regulation includes a two year implementation period with the rules coming into force on January 2025.

FS organisations will need to consider the extent to which they need to flow down contractual requirements under DORA to their vendors and this will also apply where AI is being supported to provide a service or function to an FS organisation.

Vendors that are categorised as “ICT-critical third party service providers” will also need to ensure they comply with DORA which requires registering with their local supervisory authority and agreeing to supervisory and access rights to such authority.

FCA’s and PRA’s papers on operational resilience

Both the FCA and the PRA has published papers on how firms need to ensure operational resilience (**Operational Resilience Requirements**). The Operational Resilience Requirements are different in nature and scope to the outsourcing rules. The Operational Resilience Requirements focus on the risk management processes of a firm in relation to identifying risks relating to third party dependencies, mapping these and identifying tolerances and ways to mitigate these risks. A firm is required to ensure that they carry out risk assessment on their use of AI from a third party. This may include applying certain controls, such as maximum recovery times, service levels, business continuity plans, agreed frequency of testing and other operational risk controls.

Outsourcing

There are a variety of rules and guidance relating to the utilisation of outsourced services by Firms. Each procurement will be fact-specific but, generally-speaking, it is likely that the procurement of AI systems will require Firms to consider relevant outsourcing guidance or rules applicable to them. We’ve highlighted a few key areas below.

The key regulatory requirements relating to outsourcings by banks, payment institutions and electronic money institutions are now set out in the PRA Supervisory Statement of March 2021 (**PRA SST**) and the EBA Outsourcing Guidelines dated September 2018 (**EBA OG**). There are similar requirements for insurers under the Solvency II regulation and the EIOPA guidelines.

Under the PRA SST and EBA OG certain Firms are required to impose certain rights and obligations in their contractual requirements with service providers where arrangement with service providers are viewed as critical or important outsourcing (a term used under the EBA OG) or as material outsourcing (a term used in the PRA SST). The term critical or important outsourcing and material outsourcing mean the same. The PRA SST applies to Firms that are authorised by the PRA and so this covers UK authorised banks, building societies, insurers and PRA authorised firms (i.e. the largest investment firms that can

deal on their own account). The EBA OG applies to banks, investment firms, payment institutions and electronic money institutions.

So, when such Firms are utilising AI systems from service providers they need to consider if the provision of the AI system by the service provider to them constitutes a critical or important/material outsourcing or not. If it does they need to take account of either the PRA SST or EBA OG and flow down certain terms into their contracts with the service provider. (If the provision of the AI systems is a non-critical or important outsource/non-material outsource then the PRA SST and EBA OG applies as best practice guidance rather than mandatory requirements to consider to flow down into contracts.)

Even where there is no outsourcing but only a ‘third party arrangements’ (for example, buying an AI model but deploying it in the Firm’s own servers), regulatory rules should still be taken into account.

General regulatory rules

To our knowledge, there are not currently any UK rules which are specific to the use of AI by Firms in the FS sector. However, there are a number of existing regulatory rules which may be relevant:

- The FCA’s Principles including Principle 2 (conducting its business with due skill, care and diligence), Principle 3 (using reasonable care to organise and control its affairs responsibly and effectively, with adequate risk management systems) and Principle 6 (the requirement to treat customers fairly). Firms should ask themselves whether their use of AI is consistent with these high-level rules and what controls they should put in place.
- Similarly, the Systems and Controls (**SYSC**) part of the handbook will have general regulatory obligations which should be considered by a Firm using AI. So, for example, the requirement to maintain proper systems and controls as are appropriate to its business (what controls are being put in place in respect of the use of AI) and to have proper oversight of the activities being carried on (what oversight is there of the use of AI given some of the inherent risks with the technology).
- The FCA’s Consumer Duty requires Firms to ensure good outcomes for consumers. Firms providing services to consumers will need to consider how AI can support good outcomes for consumers (such as financial inclusions) but also what risks they may pose (such as bias) and how those risks can be mitigated.

Some of the factors to be considered in relation to the use of an AI-based system within a regulated firm can be drawn from the Ethics Guidelines for Trustworthy Artificial Intelligence created by the High-Level Expert Group on AI brought together by the EU Commission. These (non-exhaustively) include:

- **Lawful and Fair:** Are the results it produces fair or could they be biased or unlawfully discriminatory? For example, giving lower insurance premiums to young women drivers may reflect their risk profile, but be unlawful sex discrimination. Do you know whether the training data could be biased or the weightings it has used when working with that data to produce its conclusions. Did it take into account factors that it should not properly have considered?
- **Accuracy:** is the training data used comprehensive and up-to-date but also has the machine learned wisely or made logical mistakes. Although an AI system will often be smarter than humans in drawing conclusions from large quantities of data, because of the way it “thinks”, if it does make a mistake, that mistake may be a serious one that a human would never make – e.g. describing a picture of a bus as a representation of an ostrich.
- **Transparency:** One of the key problems with machine learning systems is that they often operate as “black boxes” which make decisions in a way which cannot easily be interrogated. This is likely to be an area of significant regulatory concern in relation to auditing compliance by a regulated firm with its obligations.
- **Ethics:** is the use of an AI and the data gathering methods it uses ethical? For example, is it fair to look at someone’s social media feed to ascertain their lifestyle for the purposes of underwriting a life insurance policy?
- **Robustness and Reliability:** the system reliable or could it be vulnerable to catastrophic failure, cyber-attack or manipulation?

It should also be borne in mind that the use of an AI system by a business may lead to regulated activity being carried on unintentionally. For example, if an AI is asked to compare the products of a number of market providers of a particular financial service (e.g. a pension, insurance contract or credit card), any response it gives could constitute “investment advice” which can only be given by a regulated entity. Equally, if asked to put a questioner in touch with sellers of particular financial assets, it could start to become

involved in the regulated process or arranging deals in investments.

Specific Regulatory Rules

There are likely to be a number of specific UK regulatory rules depending upon the product or service that it is using AI to facilitate. For example:

Consumer credit – if AI is being used to decide whether a loan is affordable, does the AI system take into account the necessary information to make the decision?

Investment services – if AI is being used to recommend investment products, do these recommendations take into account requirements as to suitability and appropriateness?

AI Act

Currently there is no harmonisation with different countries adopting different approaches from the lighter touch approach advocated by the UK (as described above) and the US to the sterner approach advocated by the EU.

There is no doubt that EU’s approach will have far reaching repercussions. As one of the first jurisdictions to propose specific legislative focused on AI, the EU has the ambition to set a global benchmark for regulation of AI systems. The European Commission launched its Proposal for an AI Act in April 2021. It is designed to create a regulatory framework for the development and use of AI in the European Union that ensures safety, transparency and respect for fundamental rights. This proposal is currently in the process of being discussed and amended by the EU institutions, with the aim of reaching agreement on a final text by the end of 2023. At present, debate is focusing on the potential inclusion of new provisions focused on the rise of generative AI systems and on banning the use of real-time biometrical identification systems in public places.

Overall, the proposed AI Act takes a human-centric and risk-based approach to the regulation of AI systems. Rather than regulate a technology as such, the proposal addresses the perceived risks of specific uses of AI by categorizing them into four different levels: unacceptable risk, high risk, limited risk and minimal risk. This legal framework will apply to both public entities and private companies established inside and outside the EU, as long as the AI system is placed on the European Union market or its use affects people located in the EU.

Once adopted, the Act will be directly applicable across all 27 EU Member States and its obligations are expected to apply three years after the regulation's entry into force. Member States will hold a key role in the enforcement of the Regulation, with each country set to designate one or more competent authorities to supervise the implementation of the new rules and to carry out market surveillance activities. Member States will also have to lay down dissuasive penalties in the event of non-compliance, which may be up to EUR 30M or 6% of the total worldwide turnover in the preceding financial year for infringements on prohibited practices or non-compliance related to requirements on data.

Next steps?

AI is moving fast and legal teams need to keep up. Firms need to carefully consider the risks and balance them against the opportunities. As a first step, the development of a generative AI policy or guideline for its use within the business is a good first step to focus minds and ensure use of these novel technologies are aligned with the strategy of the firm and compliance with regulatory requirements.

See our article on AI policies and guidelines [here](#):

Get in touch if you need support.



Gavin Punia

Partner

+442030176884
gavin.punia@twobirds.com



Jonathan Emmanuel

Partner

+442074156052
jonathan.emmanuel@twobirds.com



Toby Bond

Partner

+442074156718
toby.bond@twobirds.com



Sanjana Sura

Legal Director

+442074156658
sanjana.sura@twobirds.com



Tom Hepplewhite

Senior Associate

+442074156777
tom.hepplewhite@twobirds.com

twobirds.com

Abu Dhabi • Amsterdam • Beijing • Bratislava • Brussels • Budapest • Casablanca • Copenhagen • Dubai
• Dublin • Dusseldorf • Frankfurt • The Hague • Hamburg • Helsinki • Hong Kong • London
• Luxembourg • Lyon • Madrid • Milan • Munich • Paris • Prague • Rome • San Francisco • Shanghai
• Shenzhen • Singapore • Stockholm • Sydney • Warsaw

The information given in this document concerning technical legal or professional subject matter is for guidance only and does not constitute legal or professional advice. Always consult a suitably qualified lawyer on any specific legal problem or matter. Bird & Bird assumes no responsibility for such information contained in this document and disclaims all liability in respect of such information.

This document is confidential. Bird & Bird is, unless otherwise stated, the owner of copyright of this document and its contents. No part of this document may be published, distributed, extracted, re-utilised, or reproduced in any material form.

Bird & Bird is an international legal practice comprising Bird & Bird LLP and its affiliated and associated businesses.

Bird & Bird LLP is a limited liability partnership, registered in England and Wales with registered number OC340318 and is authorised and regulated by the Solicitors Regulation Authority (SRA) with SRA ID497264. Its registered office and principal place of business is at 12 New Fetter Lane, London EC4A 1JP. A list of members of Bird & Bird LLP and of any non-members who are designated as partners, and of their respective professional qualifications, is open to inspection at that address.