International Regulatory Strategy Group response to BoE / FCA DP5/22 –

Artificial Intelligence and Machine Learning

The International Regulatory Strategy Group (IRSG) is a joint venture between the City of London Corporation and TheCityUK. Its remit is to provide a cross-sectoral voice to shape the development of a globally coherent regulatory framework that will facilitate open and competitive cross-border financial services. IRSG welcomes the in-depth work that has gone into the Joint Bank of England (BoE) and Financial Conduct Authority (FCA) Discussion Paper on Artificial Intelligence and Machine Learning (DP5/22). DP5/22 is an important publication to help advance the policy debate on the use of AI in financial services and how policy and regulation can best support this.

The financial services sector has long been a major user of information and communications technology (ICT), using it to become better, faster, more efficient, and more effective for the end customer. In recent years, established financial entities have overhauled their business models to better integrate technology and digitise finance. They have also used technology to become more secure and resilient, offering their customers new tools and the system new protections.

Artificial Intelligence (AI) is a broad term covering a spectrum of techniques, including machine learning (ML) and deep learning, which may be used alone or together as part of a broader process. In any discussion about AI, it is therefore important to be precise, as the level of complexity and risk associated with simple automation or learning tasks will not necessarily be the same as more advanced deep learning techniques.

AI and ML are tools increasingly used in financial services and, as noted in the Discussion Paper, may enable financial services firms to offer better products and services to consumers, improve operational efficiency, increase revenue and drive innovation, leading to better outcomes for consumers, firms, financial markets and the wider economy.

AI also comes with a number of risks, including to consumers, market integrity and financial stability, and it is therefore vital that it is managed within an appropriate regulatory structure, in order to provide appropriate protections to individuals and markets whilst also ensuring the regulatory approach is sufficiently flexible and easily implementable for organisations.

Outcomes-based regulation of AI vs. process-based regulation of AI

As the BoE and the FCA rightly point out in DP5/22, a key question to consider is whether support for safe and responsible AI adoption is best delivered through a process-based framework (that seeks to describe the processes used in AI and specify how those processes should operate), an outcomes-based framework (that focuses on the outcomes that AI may produce for clients, employees and other stakeholders), or a combination of the two.

In the view of the IRSG, an outcomes-based approach is likely to be the most appropriate in practice, noting the risk that process-focused regulation of AI may stifle innovation by forcing AI applications and processes into static regulatory categories which are less able to adapt to new innovations. This comes with the important caveat that any regulatory expectations on AI need to be sufficiently precise for firms to understand which of their activities and processes are within scope. Process-focused

regulation of AI risks going against the principle of technological neutrality and, as a result, either constraining or skewing the innovative and beneficial ways in which AI is or may be used in financial services. Were the BoE and/or the FCA to introduce new process-focused rules on the use of AI applicable to authorised firms, this may in practice lead to:

- differences in the types of AI processes and techniques that can be developed and deployed by regulated firms compared to those operating outside the regulatory perimeter, resulting in uneven competition between regulated and unregulated firms operating within the financial services sector and to a risk of regulatory arbitrage;
- either under-capture or over-capture of processes and techniques which may be of legitimate supervisory interest to the BoE and the FCA, but which may inhibit the regulators in discharging their statutory responsibilities, either by requiring them to police processes that do not give rise to the principal risks associated with the use of AI, or by failing to capture processes that do give rise to a legitimate supervisory interest; and
- in the longer term, process-based elements of the rules may restrict innovation as processes need to be modified either to come within (or fall outside the scope of) BoE and/or FCA rules.

Guidance vs. hard rules

Closely linked to this is the question of whether new- AI-specific requirements are necessary to ensure the safe and responsible use of AI, or whether more general regulatory requirements can be effective to ensure that AI is used safely and responsibly. At a high level, we welcome the approach already begun in DP5/22 of clarifying how certain existing legal requirements and supervisory expectations apply to the use of AI, in preference to the development of new AI-specific rules. This approach is in keeping with many of the overseas initiatives from financial services regulators referenced in paragraph 4.4 and Appendix 2 of DP5/22 which, when read together, reflect a general preference amongst regulators globally for guidance and clarifications on the use of AI over the introduction of new process-focused requirements with the force of hard rules. This approach also allows for the regulatory approach to AI to evolve in line with innovation, without the need for wholesale review of rules in response to technological evolution.

As the FCA and PRA rightly point out in DP5/22, there are a range of systems, controls, governance, operational and management rules that already apply to the deployment of AI by regulated firms. For the reasons set out above, focusing on the overall adequacy and outcomes of firms' governance is preferable to introducing new process-focused rules. The FCA is already in the process of introducing detailed outcomes-focused rules applicable to consumers in the form of the FCA Consumer Duty, which may provide additional safeguards against inappropriate use of AI where end consumer outcomes may be impacted. This again comes with the caveat that any guidance produced by regulators in relation to AI needs to be precise enough for firms to understand which of their activities are within scope of the regulators' expectations on the topic.

Whilst the IRSG does not consider more regulation or hard rules are needed to meet the existing desired regulatory outcomes in relation to AI, as the use of AI in financial services develops there may be some specific areas where further guidance could be helpful, for example additional guidance on the application of certain existing rules to AI, building on the commentary already set out in DP5/22 and in the final report of the AI Public-Private Forum published in February 2022.

Introducing a sector-specific definition of AI

Linked to the question of whether the regulation of AI should be process-focused or outcomes-focused is the question posed in Q1 of DP5/22 of whether a sectoral regulatory definition of AI should be included in the supervisory authorities' rulebooks.

In the IRSG's view, a sector-specific definition of AI should not be necessary , on the assumption that an outcomes-focused approach to the regulation of AI would be preferable to a process-focused approach. In this context, we note that there are a wide range of different processes used by regulated firms today that are not expressly defined or referenced in either the FCA Handbook or the PRA Rulebook, but which support firms' compliance with the high standards set for those that operate in the sector. This noted, the processes and design elements involved in certain AI applications may result in particular oversight considerations that have not existed historically in relation to less novel technologies and the IRSG would encourage the BoE and the FCA to consider these differences in formulating their regulatory approach to AI. In particular, in considering the operation of any given AI system, we would encourage the BoE and the FCA to consider both:

- the level of autonomy of the system, i.e. the ability of a system operate and take decisions independent of human oversight; and
- the level of adaptability of the system, i.e. the ability of the system to modify the inputs and parameters applicable to any given task.

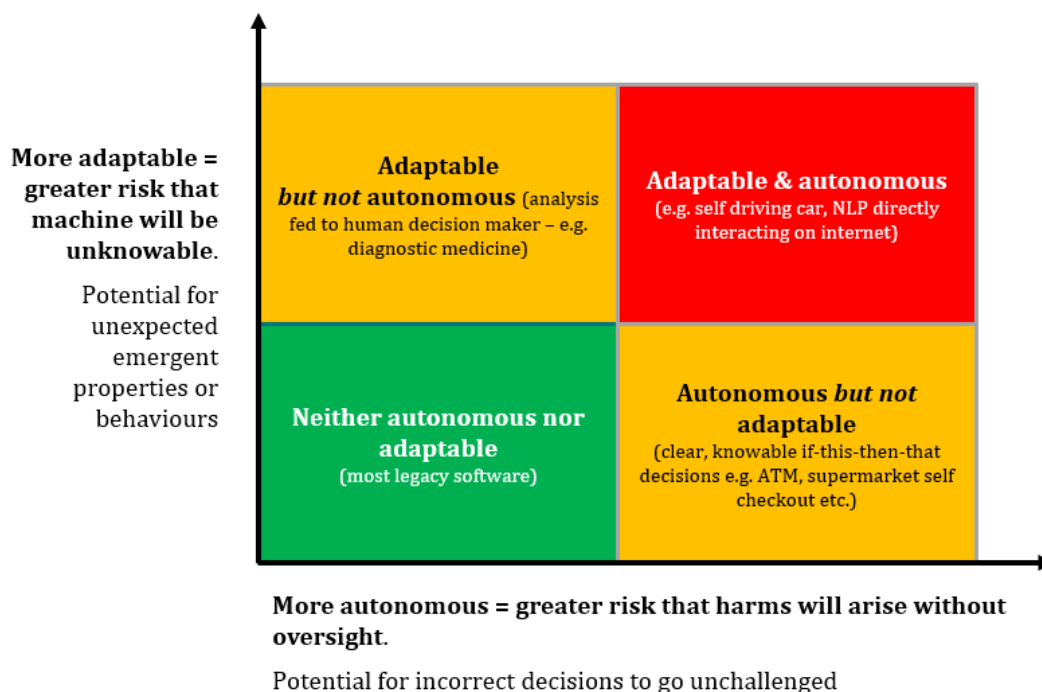In practice, both elements give rise to slightly different – though related – risks.

Fig. 1 – Graph showing different oversight considerations arising out of different AI applications

As illustrated in Figure 1 above, AI systems may have varying degrees of both autonomy (giving rise to a greater risk that incorrect decisions will go unchallenged in the absence of appropriate human

oversight, thereby causing harm) and adaptability (creating a higher risk of unexpected emergent properties or behaviours, leading to an increased risk that the system will be poorly understood by those attempting to oversee it and that its decisions may be opaque and therefore difficult to explain to customers, regulators or other stakeholders). Some systems may show more of one trait than the other, giving rise to a slightly different set of risks; however, both factors will ultimately be relevant to the appropriate controls that should be put in place around any given system.

In our view, the regulatory approach should focus on giving guidance to help firms apply existing rules to new technologies effectively by illustrating how to tackle identified risks arising out of the use of certain types of AI systems, notably those involving degrees of autonomy and adaptability. Guidance should focus on the specific identified risks arising out of the deployment of systems that are autonomous and/or adaptable, rather than setting out detailed expectations on the use of AI more generally.

As AI technology matures, and more companies provide AI products, the barriers to entry will likely decrease and many more out-of-the-box solutions are likely to be widely available to companies operating in the sector. This may bring some benefits in terms of not needing to employ specialised teams to build AI systems from scratch, a lowering of the cost of adopting and creating products based on AI systems and enabling faster product-to-market timelines. However, out-of-the-box solutions may also be more susceptible to the 'black box problem' of the AI decision-making process becoming opaque. Widespread availability of open-source systems may also result in adapted AI products that are spin-offs of the original model, leading to usage of certain AI solutions in areas beyond that the parent model was intended to be used in, giving rise to unforeseen performance issues or even security vulnerabilities.

Focusing on the risks arising out of any given system is, in IRSG's view, preferable to introducing a wider, catch-all definition of AI which may capture systems that lack the traits that generate the novel risks associated with AI. For example, whilst the Financial Stability Board's (FSB) definition of AI – "*the theory and development of computer systems able to perform tasks that traditionally have required human intelligence*" – has found favour with some policymakers, including De Nederlandsche Bank,[1] in IRSG's view such a broad definition risks being overinclusive, by omitting both the principle characteristics exhibited by AI and the main policy considerations that AI gives rise to and therefore applying to systems that do not give rise to the same risks. To take a crude example, an ATM performs tasks that were previously undertaken by human bank tellers, but need not involve the use of AI and requires very different types of oversight; however, a literal application of the FSB definition would bring ATMs into scope, even though they lack the characteristics of autonomy and adaptability.[2]

Another important consideration in this context is the importance of international alignment. Firms within the regulated sector that operate globally can find it challenging to meet overlapping requirements on topics such as AI where the approach to regulation and expectations of regulators in

---

[1] *General principles for the use of Artificial Intelligence in the financial sector*, at p.9.
[2] IRSG notes that, during the course of the development of the proposed EU regulation on artificial intelligence (the AI Act) from the initial proposal from the EU Commission, the current text of the AI Act has moved away from a more specific process-focused list of techniques and approaches listed in Annex 1 to the draft Regulation to a policy-focused definition of a system designed with elements of autonomy.

different jurisdictions differs. It is therefore very positive that the BoE and the FCA have studied closely the regulatory developments in AI in other jurisdictions in their work on DP5/22 and we would encourage a continuation of this approach in any policy initiatives that either regulator develops following the responses to the consultation. Such an approach would better focus on key risks arising out of the deployment of AI and will generate less risk of tension with other jurisdictions.

Approach to guidance

In IRSG's view, any new guidance on the use of AI within the financial sector should consider carefully how to address specific risks inherent to AI which might impact adversely on existing desired regulatory outcomes. In particular, if additional guidance is considered necessary the IRSG would advocate care to ensure it remains consistent with other existing guidance. For example, in the context of outsourcing regulation and in addition to guidance produced by other policymakers including the European Banking Authority (EBA), firms are also subject to detailed guidance, which go beyond outsourcing to cover any third party dependencies that could impact its statutory dependencies (i.e. third party dependency management). Clearly this guidance could apply to the use of the third party AI solutions by regulated firms, and so it would be unhelpful to create duplicative guidance if AI were to be subject to a separate approach in its own right; in fact, the PRA in its Supervisory Statement on Outsourcing and third party risk management (SS2/21) identifies "*off the shelf*" machine learning models as being capable of falling into this category.

SS2/21 allows for the application of proportionality (which we support), but advocates for controls for "*material*" or "high risk" non-outsourcing third party arrangements. It would clearly create a challenge for firms if an alternative category of "*high risk*" were to arise under separate guidance for AI, if in fact an AI provider was already captured by following existing guidance, especially since SS2/21 encourages a holistic approach to managing risk across outsourced and non-outsourced activity.

Housekeeping of the SMCR regime

The BoE and the FCA have quite rightly asked the question of whether amendments to the Senior Managers and Certifications Regime (SMCR) are necessary to ensure adequate governance of AI and reduce the knowledge gap between senior officials in management bodies and often more junior staff involved in the deployment of AI. As noted in DP5/22, there is an existing certified function in the FCA's Senior management arrangements, Systems and Controls Sourcebook (SYSC) covering the approval and oversight of algorithmic trading (as well as deciding whether or not a trading algorithm is compliant with a firm's obligations); however, the definition of an 'algorithm' for this purpose is defined broadly as covering any "*computer algorithm used in algorithmic trading*", so potentially already captures algorithms incorporating AI whilst not extending to the use of AI for purposes other than trading.

Linked to this is how the BoE and the FCA should ensure appropriate oversight of AI in firms' management bodies. In practice, a range of different individuals will be accountable for their firm's management of AI risks arising within their perimeter and we do not think it is realistic to have a single individual responsible for all 'AI risks', given how diverse these are. In the IRSG's view, concentrating the oversight of AI into a single SMF function or a single individual may be challenging with reference to the varied applications of AI throughout regulated firms, and could in fact undermine one of the key purposes of AI in ensuring that business line and function-aligned SMFs have individual responsibly

and accountability for activities carried out within their perimeter, e.g. in the case of AI driving trading decisions within a secondary markets business, it may be proper for the business line SMF (e.g. SMF3 or SMF21) responsible for overseeing that business to have personal responsibility for ensuring the proper oversight of AI developed within the business for which they are responsible, rather than their individual accountability being superseded by the concentration of AI accountability in a different, single individual.

Responsibility will, in our view, be clearer and more effective when focused on specific risks arising out of the application of AI to particular areas of firms' activities, rather than being focused on specific technologies, which may apply across a range of different business areas. In line with this, requiring a de facto '*AI czar*' analogous to a Data Protection Officer (DPO) under the UK onshored version of the General Data Protection Regulation (GDPR) may not neatly fit the skill set of available candidates in many cases, noting that the DPO is in any case not required to be a board-level role. With this in mind, the IRSG would not advocate either:

- creating a dedicated SMF role for AI; or
- making AI a prescribed responsibility (i.e. a specific responsibility that must be allocated to at least one senior manager), recognising that some firms may not be users of AI at all, whilst others will be heavy users of AI.

In practice, a degree of oversight responsibility for processes involving AI may already be included within the scope of the Statements of Responsibility (SoRs) for particular SMFs, noting that it may not always be appropriate to include express, process-focused references to AI over and above other, important processes (e.g. manual processes) which also have the potential for significant stakeholder outcomes.

We note that, as part of the "Edinburgh Reforms", the UK government is due to commence a review of the SMCR in Q1 2023.

Ethical use of AI

Separate from the management of risks arising out of AI processes is the question of when the deployment of AI is ethical, particularly where AI is applied to personal data or may have a potential impact on consumer outcomes. The ethicality of any proposed AI solution is one that needs to be considered carefully, both by companies operating in the sector and by regulators, noting that ethical questions intersect with the management of AI risk. To take an example, at least some customers may have ethical concerns regarding the deployment of AI based on data from their current account, even if the deployment of AI using data inputs from current account transactions could have a theoretically positive outcome from the perspective of the customer's overall financial management.

In this context, the IRSG echoes comments already made in the final report of the Public-Private Forum[3] on the role of the Information Commissioner's Office (ICO) and we would encourage the BoE and the FCA to proactively engage with the ICO given the multiple intersections with privacy and data protection laws. This should help to ensure consistent and coherent policy making in relation to the

---

[3] https://www.bankofengland.co.uk/-/media/boe/files/fintech/ai-public-private-forum-final-report.pdf?la=en&hash=F432B83794DDF3F580AC5A454F7DFF433D091AA5

use of AI. Examples of areas where the supervisory interests of the FCA and the ICO in particular may overlap could include:

- how the use of AI may impact groups sharing protected characteristics;
- whether particular applications of AI could be discriminatory, e.g. when certain data or data sets are used as AI inputs in respect of credit decisions; and
- how any such impacts might be mitigated either by companies deploying AI or by the supervisory authorities.

Potential mitigations of these risks could the anonymisation of data, restrictions on the tracking of data subjects across different applications, restrictions on the use of protected characteristics during training of deep learning models and not using Ai models that are fully automated for particular decision-making applications.

Data standards and curation

The IRSG agrees that data curation is a cornerstone to safe and responsible AI. Like many existing (non-AI) systems, the quality of outputs from AI very much depend on the quality of the data input to train the AI.

We agree that we are likely to see an increase in the complexity of:

- the data that AI systems are able to interrogate, including higher volumes of unstructured data;
- chains of interlinked AI solutions which may allow for increasingly complex use cases to be tackled by AI, but which may also give unpredictable and unexplainable results that may not have been inherent in the individual AI models used to build the combined solution; and
- the power of computing including quantum and analogue computing, offering greatly enhanced computational power but potentially with more challenges for explainability of their logic processes.

These various complexity related challenges may make effective "human in the loop" oversight, transparency and explainability more challenging to achieve. Even with the current incarnations of AI solutions, data science and statistical analysis skills may be required to be able to effectively detect bias (and other data risks such as incomplete data, erroneous data, unrepresentative data, etc) and explain the logic used by AI to reach any given decision.

We agree that there is a potential for a knowledge gap in firms including at the board level and that new skill sets are likely to be needed to ensure that there is effective oversight. With ever increasing computing power and more complicated and more autonomous AI, there is also a wider challenge as to how effective human oversight can be maintained and how these highly complex solutions can be explainable and transparent in plain English when processing personal data (or otherwise).

The next generation of AI models are being designed to match the number of trainable parameters in the human brain. OpenAI's next generation natural language model, GPT-4, is expected to have around *100 trillion* trained parameters – and while it is possible to determine the state and logic used by these systems at any point in time, few have the skills to be able to interpret that logic and determine if it is behaving as predicted and lawfully.

To reduce AI solutions from the size of a large building to closer to the "wetware" size of the human brain, new technologies will be required, including analogue computing which rather than using circuitry to add values together as in traditional, digital computing, use much simpler structures to add by accumulating electric charges and multiply by amplifying electric charges. While analogue computing has exciting potential for power and space efficient processing, it adds considerably to the challenge of transparency and explainability. Novel methods of interrogating and interpreting the operation of such systems and to report in a manner which allows for effective "human in the loop" oversight will be required if they are to meet these requirements[4].

Firms are also mindful of their obligations under the increasingly complex web of global data protection and privacy laws. At least some AI systems which process personal data are likely to meet the threshold for high risk processing under Article 35 of the UK GDPR (and EU GDPR), requiring firms to carry out a data protection impact assessment. Any sectoral guidance and requirements for AI should therefore align with these existing legal requirements to avoid duplication and inconsistency, which would create a disincentive to the development of safe and responsible AI.

It will be important for firms to manage bias in their data sets – as well as in their processes and the oversight of those processes – particularly as more unstructured data are used to train AI solutions. For example, researchers in the US have drawn attention to risk of racism in lending practices[5] and that it should not be assumed that use of AI automatically prevents racist practices in lending.[6] The context of the use case is always important to consider, as a biased data set may nevertheless deliver unbiased and explainable results if the bias characteristics of the data set do not have any impact on the logic used and outcomes. On the other hand, data sets which do not appear to include any obvious bias may produce (unlawfully) biased results for specific use cases either on their own or when linked with other AI solutions. This means that care will need to be taken when tackling the challenge of unlawful bias, so as not to impose disproportionate regulatory burdens on the one hand but also to ensure that all unlawful bias is detected and prevented on the other hand. Balanced against this, it is also important to recognise there will also be cases in which narrow data sets are used intentionally with a view to ensuring positive consumer outcomes in particular subsets of customers, e.g. vulnerable customers, where the deployment of AI could be used to take better and more informed decisions that lead to an increase in service provision to customers whose available options in the market might be more limited.

International considerations

Introducing a sectoral regulatory framework involves consideration of the direct and indirect impact of the (cross-sectoral and economy-wide) EU AI Act and other significant international AI legislative developments. Alignment is relevant to firms that do business in Europe and around the world. We would also suggest that it would be prudent to monitor and benchmark in parallel how the implementation of broader AI legislative developments will reverberate, potentially setting standards

---

[4] For a longer discussion of these challenges, see "Immeasurably better? Non-digital AI and the regulatory challenge" by Gareth Stokes, November 3, 2022, DLA Piper Technology's Legal Edge;
[5] A comment on Bank of America/Countrywide's discriminatory mortgage lending and its implications for racial segregation | Economic Policy Institute (epi.org)
[6] See e.g. the February 2019 paper Consumer-Lending Discrimination in the FinTech Era.

that will become a market expectation (which, as the Discussion Paper identifies, are predominantly non-sectoral, focussed on governance considerations from a statutory[7] perspective, which as outlined above we consider are covered by existing UK financial services frameworks, or guidance, in line with our comments above[8]).

It is also relevant to consider from a UK and international business perspective the context of firms that operate different business lines both within and outside the financial services regulatory perimeter (e.g. business lending in the UK) and in addition uses of AI for non-customer-facing purposes (e.g. internal talent management). Care around a sector-specific regulatory framework is needed to minimise the potential conflicts of markedly different regulatory frameworks applying to different parts of the same entity, whether in the UK or abroad.

Concluding comments and membership

We stand ready and willing to continue to engage with the Bank and the FCA on this important project and look forward to opportunities to exchange more detail in the future.

The IRSG organised a workshop on DP5/22 in November 2022 with interested Members and has conducted a number of follow-up discussions with members to discuss and collate their feedback. The IRSG wishes to thank those who have overseen production of this response, in particular DLA Piper.

We thank you for considering this submission.

Contact address:
IRSGSecretariat@cityoflondon.gov.uk

---

[7] The US Algorithmic Accountability Act
[8] EG Netherlands, Hong Kong, Singapore – As per Appendix 2 of the Discussion Paper