ABI

AI Guide

Practical ideas for getting started with responsible AI

February 2024

Contents

Disclaimer

This Guide represents the views of the ABI and its members and is intended to support firms’ responsible use of AI. The adoption and use of AI will be a matter for individual firms. The information provided herein does not, and is not intended to, constitute legal or other professional advice, and should not be relied on as such. Nothing in this Guide is intended to replace or conflict with firms’ legal and regulatory obligations.

1. Foreword

Artificial Intelligence (AI) is not new.

Alan Turing envisaged the potential for computer intelligence as far back as 1947. It was another 50 years before Deep Blue, a computer built by IBM, beat the then world chess champion, Garry Kasparov.

Today, each of us has had some experience of engaging with AI. But it is only since recent developments in generative AI that its huge potential has started to be realised.

This brings with it numerous opportunities. It has the potential to transform the operations of our industry in the interests of customers, through improved affordability, accessibility and availability.

However, it does also raise questions and potential risks. No firm wants to be left behind, but we must ensure that we bring customers with us on the journey as we embrace and adapt to the benefits of AI. Their trust in its use is vital.

As the trade body for the insurance and long-term savings sector, the ABI is focused on ensuring the significant power of AI is harnessed in a responsible manner.

That's why we've worked with our members to develop this guide to help firms start on their AI journey and maximise the benefits for customers. It provides a practical approach to applying the five guiding principles set out in the UK AI Policy Paper, 'AI regulation: a pro-innovation approach', that underpin the responsible use of AI.

It will help firms understand the questions they need to ask themselves to make sure AI is being used responsibly, consider what they need to do to mitigate the risks of potential bias or exclusion and recognise examples of good practice from use across the sector.

Whether you're buying AI from a third party or building capability in house, we hope this guide can help set you on a path to using AI reliably and safely within your business. And as AI develops, we'll keep listening and learning so that together we can respond and adapt to this fast-moving environment.

Hannah Gurga
ABI Director General

2. Introduction

Using AI responsibly

Developed collaboratively by the ABI's AI Working Group, comprising experts from actuarial, data science, data protection, legal, regulatory strategy and compliance across the ABI membership, this guide aims to help firms start on their AI journey.

It aims to provide a practical approach to applying the five principles underpinning responsible AI (**Safety**, **Security & Robustness**, Appropriate **Transparency & Explainability**, **Fairness**, **Accountability & Governance**, **Contestability & Redress**), as set out in the UK's AI Policy Paper, "AI regulation: a pro-innovation approach".

The guide includes a set of AI use cases in insurance and long-term savings, some key questions firms should ask themselves to assist in the responsible use of AI, an AI Risk taxonomy, a set of good practice examples relating to AI, and an overview of existing regulations and legislation with application to AI.

Consumer-centric

Whilst AI brings opportunities and risks to a wide range of activity within the insurance and long-term savings context, ranging from the broad economic, political and environmental, to societal and legal, this guide focuses on ways in which the insurance and long-term savings industry can minimise and mitigate risks and maximise the benefits of AI to its consumers. It does not cover the vast range of other impacts of AI, or the wider effects of AI on business, the economy, society and environment.

> “This guide focuses on ways in which the insurance and long-term savings industry can minimise and mitigate risks and maximise the benefits of AI to its consumers.

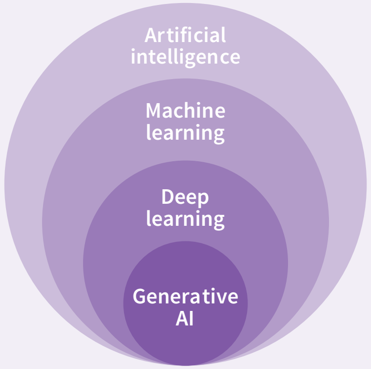
What do we mean by AI?

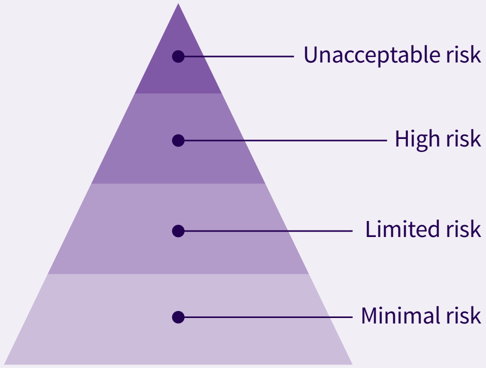
There is no universally agreed definition of AI, not least because the systems and their capabilities are developing so quickly. In line with the technology-neutral approach of UK digital regulators, this guide focuses on the outputs of the system and their potential effects, rather than the operation of the systems themselves.

To distinguish AI from other advanced data applications that have been in place for years, this guide is aligned with the UK AI policy paper which uses the term AI to mean technology that has the defining characteristics of:

- **adaptivity** (AI systems are trained and, through such training, develop the ability to perform new forms of inference), and
- **autonomy** (Some AI systems can make decisions without the express intent or ongoing control of a human)

It also takes into consideration the wider international AI definitions from the OECD, and the EU AI Act, which identifies and describes four risk levels **of minimal**, **limited**, **high**, and **unacceptable**.





Introduction (continued)

Changing environment

The fast pace of change in AI technology and need for regulations and laws to adapt means that the guide's references to current regulations and laws are likely to date quite quickly. Whilst this is the case, we expect the principles underlying the UK's AI Policy Paper to have longevity, and the risks and opportunities, and related questions and ideas, to remain relevant over a longer timeframe.

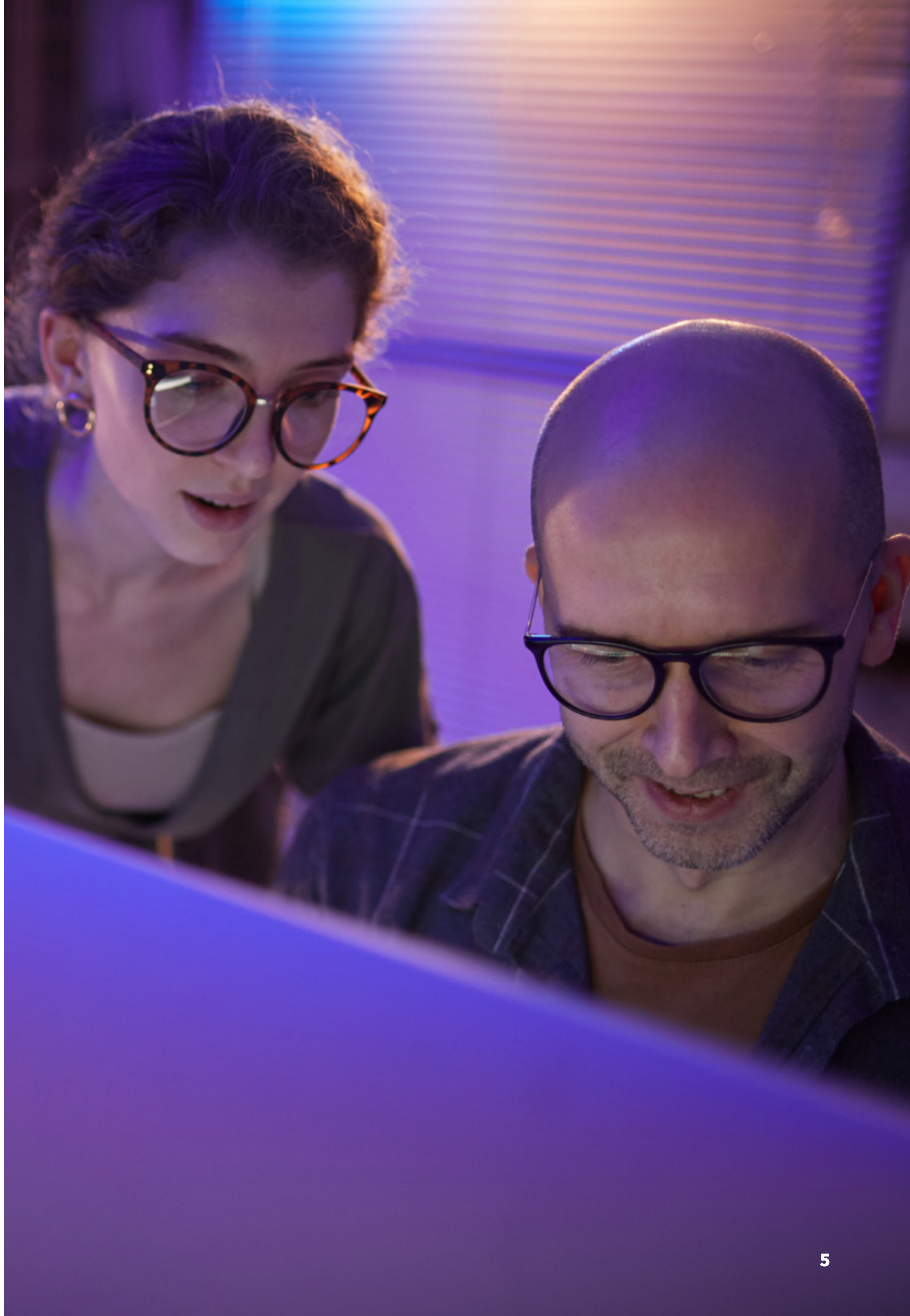
Engagement and continuous development

This guide aims to help firms start a conversation around AI. We hope it will:

- Provide a starting point for the industry to engage with government and regulators on AI.
- Provide relevant, adaptable resources, to help firms to develop a responsible AI strategy, governance, and oversight of use.
- Help drive inclusivity in firms' use of AI, empowering people in all roles to ask questions and not to make assumptions.
- Help firms consider AI in the context of relevant conduct, data, and consumer protection regulations and requirements.

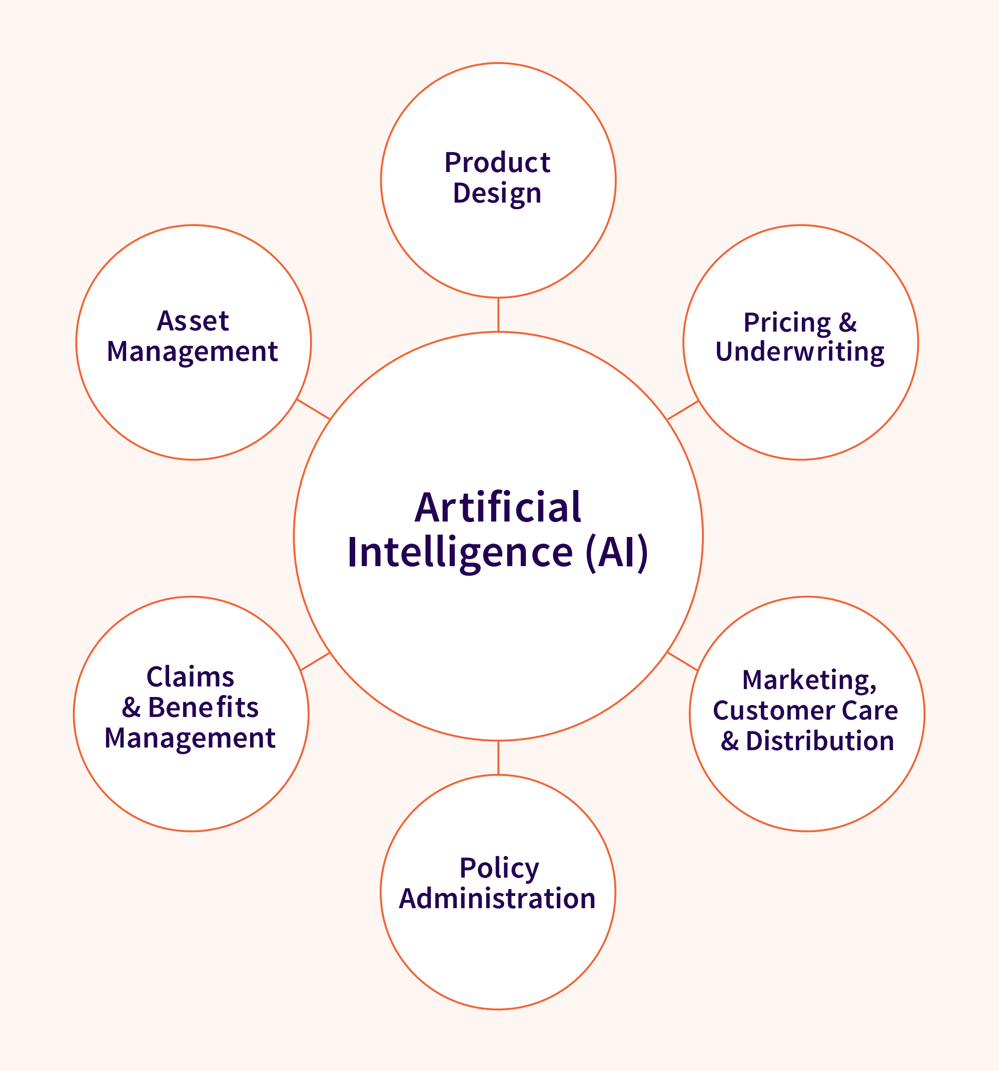
We think it is very important to keep on listening and learning so that we can respond and adapt to this fast-moving environment.

We look forward to continued close engagement with our members, the government and regulators as the technology and legal and regulatory framework around AI evolves.



3. AI use cases in insurance and long-term savings

This chapter provides some examples of applications and uses of AI in the insurance and long-term savings sector. The use cases are ordered and categorised broadly according to the stages of a typical product lifecycle.


Product Design
Pricing & Underwriting
Marketing, Customer Care & Distribution
Policy Administration
Claims & Benefits Management
Asset Management
Artificial Intelligence (AI)

Product Design

Portfolio optimisation	AI can identify patterns within underwriting portfolios, align with patterns in claims and identify areas where opportunities to rebalance exposure exist.
Elderly Care Planning	AI can assist in the development of long-term care insurance policies and financial planning for an aging population.

Pricing & Underwriting

Personalised Pricing	AI can analyse an individual's data to offer customised insurance premiums based on their risk profile, driving behaviour, health and more.
Predictive Underwriting	AI can be used to assess potential risk to policyholders and offer more tailored policies. This potentially expands access to insurance for individuals who might otherwise be denied coverage.
Mortality and morbidity modelling	AI can be used to augment current actuarial models including mortality and morbidity. It can also be used more broadly for lapse/ customer retention modelling.
Submission analysis and triage	AI can ingest emails and attachments, extract the relevant points as it relates to an underwriting appetite, validate these points, identify where further information is required and prioritise the workload for the underwriter.

AI use cases in insurance and long-term savings (continued)

Marketing, Customer Care & Distribution

Chatbots and Virtual Assistants	AI-powered chatbots and virtual assistants can provide instant customer support and guide policyholders through insurance processes.
Codified guidance and robo-advice	Use of AI to provide tailored advice and/or guidance to customers on insurance and long-term savings propositions. For example, later life planning.
Segmentation	Clustering and classification AI models can be used to group customers based on behavioural and/or attributional data, for more effective communications.

Policy Administration

Personal Productivity	Generative AI can be used to improve personal productivity by completing tasks such as: drafting and summarising emails, drafting and summarising content for reports, translating text into multiple languages, de-bugging and translating software code or transcribing and summarising online meeting minutes.
Content Creation	Generative AI can be used to generate written content, for instance hyper-personalised customer or broker correspondence or marketing communications. This enables the insurer to review and develop customer relationships.
Content Interpretation	NLP/Generative AI can be used to interpret unstructured information such as voice of the customer, contracts, claims submissions and complaints.

Claims & Benefits Management

Fraud Detection	AI algorithms can identify unusual patterns and behaviours to detect insurance fraud. They can read handwritten police notes, structured claims forms, watch video images and read photographic images of an incident. This results in faster claims handling and workload prioritisation for the fraud analysts.
Claims Automation	Automate claims processing and approvals by reading and validating claims against policy wordings and reading, analysing and validating images. This reduces the time taken to disburse funds to policyholders.
Natural Disaster Predictions	Use AI to predict natural disasters and assess the potential impact on the insurance industry, influencing long-term savings strategies and risk assessment.

Asset Management

Liquidity and balance sheet optimisation	AI can be used to improve liquidity forecasting and optimise balance sheet by making best use of available surplus liquidity.
Environmental Risk Assessment	AI can evaluate climate change-related risks and help insurance companies develop sustainable investment strategies for long-term savings.

4. Questions to assist in the responsible use of AI

This chapter considers the **five AI principles** within the context of a business operation. It sets out the statement of intent (the purpose or aim), some key questions firms can ask themselves to help consider the bigger picture, and some follow-up questions to help prompt further reflection and potential action for firms to take. The questions are not exhaustive, and firms will need to consider how to apply the principles within their own business models and risk appetites.

1 Safety, Security & Robustness

Reliability

We know the bounds within which our AI systems are expected to perform reliably.

- When does the system perform poorly?
- Can you provide examples of situations where your AI systems have failed to perform as expected?
- What is the process for addressing identified problems with your AI systems?
- How are you measuring/ identifying poor performance and/or accuracy?

Guide links

- **Consider chapter 6, Fairness subsection (3), point 6**. The most robust way to identify when AI systems are non-performant is to develop a clear test plan resulting in relevant KPIs.

Failure recovery

We anticipate when our system may fail and define a plan to recover.

- What are our disaster recovery and business continuity plans?
- How do you ensure that the AI system is secure and protected from cyber threats?
- How often do you test your disaster recovery and business continuity plans?
- What key performance indicators (KPIs) do you use to measure the effectiveness of your disaster recovery and business continuity plans?
- What contingency is built into third party supplier contracts who may be providing (or part providing) AI systems?
- What expertise and understanding of AI exists across the organisation in order that failures c an be identified and rectified?

Guide links

- **Chapter 6, sub-section 1, points 1-3** outlines the essential concepts to consider how you can firstly plan for potential AI failures and identify key contingency plans to protect customers and operational resiliency.

Monitoring

We monitor our AI systems to identify issues, manage, maintain and improve over time.

- How do we maintain the performance of our AI system?
- What metrics do you use to measure the performance of your AI systems?
- How do you ensure that your AI systems are performing as expected?
- What processes do you have in place to continuously improve the performance of your AI systems?

Guide links

- **Chapter 6, sub-section 3, point 7**. Building on the concept of a test plan. Baselining the model expectations enables any delta to be identified and drift detected.
- Additionally, **sub-section 4, point 3** discusses a catalogue of the underlying data. This is essential, as it enables audited and clear decisioning about what data is used to train AI algorithms.
- Equally, in **sub-section 5, points 2-3** highlight the importance of a challenge route and correction.

Questions to assist in the responsible use of AI (continued)

2 Appropriate Transparency & Explainability

Explainability

Where our AI systems support decision-making, they are designed to output intelligible results.

- How are the decisions made?
- What are the criteria for determining the intelligibility of the output results?
- Who provides the explanation – especially if third party AI systems are being used?
- How can we measure explainability to know that it is clear and not misleading, and ultimately understood by end users?

Guide links

- **Chapter 7 sub-section 1** references the UK GDPR guidance on the meaningful information about the logic involved in an AI-based decision. Special reference should be made to the AI and Data Protection risk toolkit.
- Additionally, **chapter 6 sub-section 5 point 1**: Using explainability techniques e.g. LIME and SHAP, where appropriate will further clarify the logic behind the AI reasoning.

Disclosure to users

We provide relevant information about our use of data, machine learning and AI.

- What do our users think is happening?
- How do you ensure that the explanations provided by your AI systems are accurate, relevant, and understandable for the intended users?
- When should you provide transparency statements to let users decide to interact or otherwise?

Guide links

- **Chapter 7 sub-section 2** provides the regulators' standpoint on when explanations should be provided to the user.
- When explanations are required, **chapter 6, sub-section 5 point 1** provides information on what explainability techniques can be engineered.

Purpose

We clearly understand the purpose and necessity of using AI to achieve that purpose.

- Is the data being used within this AI application collected for this purpose?
- How can we monitor (limit) usage to ensure purpose remains appropriate?

Guide links

- **Chapter 6 sub-section 3, point 1** asks the important question of 'why' it is appropriate to use AI, and therefore what data is necessary.
- **The following points 2-4** are useful in promoting data quality and bias awareness. Together these points ensure careful consideration of the appropriateness of using AI in any particular context.

Questions to assist in the responsible use of AI (continued)

3 Fairness

Equality of Service

Our AI systems offer comparable quality of service (QoS) for different demographic groups.

- How does QoS vary across demographic groups?
- How do you monitor and evaluate the performance of your AI systems to ensure that you are offering comparable QoS across different demographic groups?
- Can you provide an example of how your AI systems offer comparable QoS for different demographic groups?
- How are consumers advised on their right to challenge auto decisions and when/at what point?
- What impact will the absence of consent have on providing fair and equal outcomes in relation to AI auto decisions? For example, will users be excluded from pricing decisions if they choose not to consent?

Guide links

- Ensuring the identification of any quality parity issues over multiple demographic groups starts with a clear understanding of your model training data. **Chapter 6 sub-section 3, point 3** discusses how to introduce diversity into your training data. Building on this, **point 7** encourages clear causal relationships in the AI system. Understanding these relationships is the foundation to identifying quality variances.
- In **chapter 7 sub-section 5**. Clear rights to request information, challenge and correct decisions made by AI systems in relation to key regulation which addresses some key challenges of consent and fairness.

Bias Minimisation

We do our best to identify, understand and minimise unfair bias.

- How do we minimise unfair bias?
- What are the measures in place to ensure that the AI system is free from unfair bias and discrimination?
- Can you provide an example of how you identify unfair bias in your AI systems?

Guide links

- **Chapter 6 sub-section 3 points 1-4** provide information on both identifying and minimising unfair bias in AI systems.

Consistency

Where possible, we ensure that our models produce consistent and repeatable results.

- How and why do the results/decisions vary?
- Can you provide an example of a decision that was made by the AI model that was inconsistent with the expected outcome?
- How do you ensure that the data used to train the AI model is representative of the population it was intended to serve?
- What measures do you have in place to detect and mitigate inconsistency in the AI system results?

Guide links

- Building on the transparency measures detailed in **chapter 6, sub-section 5 point 1**, perhaps the most important point is made in **chapter 6 sub-section 1 point 5**. Training staff to be aware, challenge, and question the outcomes of an AI system is often your first line of defense against the detrimental use of AI.

Questions to assist in the responsible use of AI (continued)

4 Accountability & Governance

Impact

We care how our AI systems impact people, organisations and the environment.

- What impacts will this AI system have?
- What specific measures do you have in place to ensure that the impact of your AI systems is positive?
- How do you measure the impact of your AI systems on people, organisations, and the environment?
- What steps do you take to address any negative impacts of your AI systems on people, organisations, and the environment?
- How would cross border usage impact if we have different regulatory regimes?

Guide links

- **Consider chapter 7** to judge the impact of your systems against current regulatory policies. Clear levels of transparency (made possible by explainability techniques such as LIME & SHAP) of what the AI is doing enables clear KPIs on people, environment, and negative impacts of the AI in use.

Adverse effects

We identify where AI systems have adverse effects on people, organisations and environment.

- What could go wrong?
- Can you provide an example of an adverse effect that your organisation has identified in the past?
- How do you ensure that the adverse effects of AI systems are mitigated or eliminated?
- What is the process for reviewing and updating your principles to ensure they remain relevant and effective?
- Where will liability rest – with the firm deploying the AI or with a third party AI provider of the solution?

Guide links

- Again, consider **chapter 6 sub-section 1, point 5**. Having well-trained, data- literate staff enables clear decisions to be made on where AI works, and where it doesn't.
- Additionally, if something should go wrong, clear routes of accountability are essential. **Chapter 6 sub-section 4, point 4** discusses this point.

Level of oversight

Wherever possible, we enable appropriate levels of human oversight.

- How can we challenge a decision?
- Do we understand the limits of the AI system?
- What specific measures do you have in place to ensure that the level of human oversight is appropriate?
- How do you ensure that the level of human oversight is appropriate?
- How do you ensure that the human oversight is effective in identifying and mitigating any adverse effects of AI systems?
- What is the process for challenging a decision made by an AI system, and how do you ensure that it is transparent and fair?
- Who holds ultimate responsibility for the model?

Guide links

- A combination of chapter 6, specifically the transparency mechanisms detailed in **sub-section 2, point 3**, and **sub-section 5**. The principles set out allow customers to challenge AI-based decisions.

Questions to assist in the responsible use of AI (continued)

4 Accountability & Governance (continued)

Lawful basis

We will always act within the guidelines and limits of the relevant legal and regulatory frameworks. We have a clear, documented and evidential consent position in relation to auto decision making.

- How have we demonstrated lawful basis?
- What specific legal and regulatory frameworks do you follow to ensure that your use of AI is lawful?
- How do you ensure that your use of AI is compliant with the relevant legal and regulatory frameworks?
- What measures do you have in place to ensure that your use of AI is transparent and accountable?

Guide links

- **Reference chapter 7 sub-section 1**, the ICO toolkit on AI and data protection. Ensuring before any AI development takes place, an appropriate risk assessment on the use of AI regulatory frameworks goes some way to evidence tour lawful and compliant use of AI.

Risk ownership

We know who is accountable for our AI systems.

- What is the RACI for this system?
- How is the risk ownership distributed across the organisation?
- What are the consequences of non-compliance with the principles guiding the use of AI?

Guide links

- **Chapter 7 sub-section 4** shows us the importance of AI ownership from the outset. Having clear owners of AI risk prompts careful consideration of AI use.

5 Contestability & Redress (C&R)

Clear, accessible point of contact for all actors in the AI lifecycle

There is a single, clear point of contact or mechanism for users and affected parties to seek redress for adverse outcomes.

- How does someone appeal, complain or seek redress for adverse outcomes or impacts, or to correct or delete inaccurate or harmful data?
- What are our processes for resolving the harmful AI decision(s) or outcome(s) quickly and effectively?
- How do we help those who may be affected with a potentially harmful AI outcome or decision to raise awareness, contest or seek redress for adverse outcomes?
- How do we resolve any tension between a right outcome/decision that a user perceives to be unfair or incorrect? (i.e. if the computer rightly said no).
- Who is responsible for harmful outputs that arise from misuse, malfunction, flawed model or ML bias or inaccurate data – especially if data was sourced elsewhere?

Guide links

- A clear data privacy policy is the foundation stone of minimising customer harm. Beyond that **chapter 7 sub-section 4** details clearly how participants in the AI journey have routes of accountability. This should be both technical and business stakeholders.
- **Sub-section 5** details the rights of the individual in an AI system. Reference should be made to the ICO subject access request as this is a likely mechanism a customer may raise an issue with AI decisions.

5. AI Risk Taxonomy

This chapter aims to provide an overview of some key risks associated with the use of AI in the context of consumers and conduct regulation, and some suggestions for some steps that can be taken to mitigate these risks.

Safety, Security & Robustness

Risk & Description	Mitigating actions
Inadequate data quality Risk that AI systems are not robust or produce outputs that are inaccurate	• Human Intervention • Data governance policies and procedures to maintain data quality through its lifecycles • Algorithm selection and validation • Anomaly detection • Monitor and feedback • Data quality metrics
Internal and external fraud Risk that vulnerabilities in AI systems can be exploited for malicious intent	• Ethical AI development • Align to regulation and policy • Education and awareness of developers and users of AI
Model drift Risk that over time models may become less accurate due to changes in operational data distribution or environmental factors	• Continuous monitoring • Data quality assurance • Regular retraining of the model with updated datasets • Automated testing to evaluate model performance • Feedback loops – to collect insight from end users

Risk & Description	Mitigating actions
Data security Risk that unauthorised access, data breaches and data manipulation can compromise the integrity of the AI Systems	• Data security policy • Training data integrity checks • Data access audit trails • Secure coding practices • Secure communication channels • Regular security updates • Training and awareness • Incident response plan
Cybersecurity threats The risk that AI is used to develop sophisticated cyberattacks, posing risks to critical infrastructure and information systems	• Security awareness training to users on the risks and threats of AI cyber attacks • Robust authentication and authorisation mechanisms • Regular security audits and penetration testing • Implement intrusion detection and prevention systems • Collaborate on threat intelligence

AI Risk Taxonomy (continued)

Transparency & Explainability

Risk & Description	Mitigating actions
Lack of interpretability & explainability Risk that AI models and systems are complex and difficult to understand in order to show how decisions are made	• Use model simplification techniques • Use interpretable models, decision trees or linear regressions as proxies • Model validation and testing • Maintain documentation on model architecture and decision-making processes • Human oversight and intervention • Continuous monitoring • User training on model use

Fairness

Risk & Description	Mitigating actions
AI bias Risks that AI systems amplify the bias in the training data and are applied in a way that is unfairly discriminatory towards a particular community or group	• Diverse representative data • Bias detection and evaluation • Data Ethics Frameworks • Diverse development teams • Audits and Monitoring • Clear accountability
Privacy violations AI systems may collect and misuse personal data	• Data minimisation • Anonymisation and pseudonymisation • Data governance policy • Accountability and oversight • Data Ethics Frameworks • Regular audits • Continuous training

Accountability & Governance

Risk & Description	Mitigating actions
Management of third party and suppliers Risk that third parties develop AI, that acts on our behalf without our awareness, resulting in regulatory exposure	• Reputable vendor selection • Contractual agreements • Risk and control oversight • Clear guidelines requirements • Oversight and Audit of third parties • Legal and compliance oversight
Dependency on third party providers Risk that third parties' reliance on AI technologies and services impact service quality, reliability and continuity	• Reputable vendor selection • Contractual agreements • Risk and control oversight • Oversight and audit of third parties • Third party supplier frameworks • Customer service standards
Skills gap Risk that inadequate training and expertise among employees can lead to misuse or misinterpretation of AI outputs	• Continuous learning and development • Diverse multidisciplinary team • Audit and reviews • Guidelines and documentation on responsible AI

6. A set of good practice examples relating to AI solutions

Overview

This chapter provides a set of good practice examples, structured according to the five principles, for firms to consider when using AI solutions. They are designed to help firms make the most of the AI opportunity whilst minimising the risks to their business and most importantly their consumers.

1 Safety, Security & Robustness

1.1 **Consider the impact of AI on security –** Data systems should be subject to stringent cybersecurity processes to identify and mitigate security threats that could apply at different stages of the AI lifecycle. The RACI (Responsible, Accountable, Consulted, Informed) for each use of machine learning or AI should include details of cybersecurity controls and actors. There is already information available that can be considered such as the National Cyber Security Centre (NCSC) principles for securing machine learning models and the Guidelines for secure AI system development.

1.2 **Data Protection and Privacy –** Processing of personal data involved in the design, training and use of AI systems should be compliant with the requirements under the UK General Data Protection Regulation (GDPR) and the Data Protection Act 2018, particularly around solely automated decision-making. High-risk processing would require a Data Protection Impact Assessment to help minimise the risk of non-compliance with data protection legislation that the use of AI poses.

1.3 **Data environments should focus on privacy –** Data environments should expose the minimum data to analysts. The use of IDs, hashing, and other techniques should be utilised to ensure limited access to data, and appropriate access controls should be in place.

1.4 **Consider Privacy Enhancing Techniques –** Consider the granularity of data included in Data Science environments. For example, some postcodes have a single household in them, therefore, this postcode could be used to identify an individual. Also, for smaller areas a single postcode and age could be sufficient to identify an individual. In these instances, consider aggregation techniques to avoid using data which could directly identify an individual.

1.5 **Train staff to be alert and sceptical –** AI can seem to have amazing powers to improve our lives. However, that convenience can also make people over-reliant, too accepting of its results and miss situations where things have gone wrong. Therefore, staff need regular training to ensure awareness of AI's limits and the need to maintain human judgement. Firms need to create processes to detect when things have been substandard. Staff using AI should be responsible for its output and have ways to report when things have not worked to developers.

2 Appropriate Transparency & Explainability

2.1 **Champion the use of "good enough" techniques –** When building an in-house model, we recommend that you use the simplest model in terms of technique and variables. There is often a trade-off between accuracy and explainability, therefore, different techniques and models should be tried to ensure there is an optimal balance between these two considerations. Independent technical reviewers (i.e., not the person or team building the model) should be utilised to ensure the correct balance has been struck. It is unlikely that third-party suppliers will divulge full details of their models, however, requesting general information about their approach is recommended.

2.2 **Visualise Blackbox machine learning –** Visualisations will help fully understand the outcomes of predictive models and will ensure that fair decisions are made. Any machine learning models using Blackbox techniques (e.g., neural network, tree ensemble models etc), should be fully visualised through methods such as SHAP/LIME/ALE. Common English descriptions of models should also be available for reviewers and governance functions. Increasing internal understanding of models gives governance functions the ability to provide an appropriate level of challenge. Third-party models should be tested to ensure that an appropriate level of understanding can be established.

A set of good practice examples relating to AI solutions (continued)

2.3 **Work within customers' realistic expectations –** Data should not be used in a way that undermines customer trust. Therefore, we need to consider customer expectations and whether they will be impacted in a way they would not predict or expect. We recommend that a statement about how the use of data meets customer expectations is included in the development of models and within AI use cases more generally.

2.4 **Use sensitive information carefully and sparingly –** We should use the minimum and the least sensitive information available. Justification for using this data should be documented and assessed to ensure that it is lawful and proportional.

2.5 **Provide adequate explanations –** There should be an appropriate level of transparency with customers when capturing and using customer data. Privacy functions should advise on levels of transparency required for each use case.

3 Fairness

3.1 **Ensure customer benefit –** Customers should receive a discernible benefit in return when their data is used. The benefits of AI should be documented, and a risk assessment carried out to balance opportunities against any potential risks that AI can generate.

3.2 **Ensure diversity in the training data –** Consider bias in data. Bias in data can be caused by a range of reasons, including lack of representation (e.g. through lack of access to, or use of a service), poor design, human bias, historical or societal bias that is repeated and reinforced. This can result in harmful outcomes. The process of detecting bias is not easy, and we recommend introducing reviews where relevant experts not directly involved in the build consider what biases may exist. If bias is detected, take steps to consider what can be done to mitigate it. Stakeholders should be made aware of potential biases and any and all steps that have been taken to limit them.

3.3 **Ensure high standards of data quality –** Models need to be built on reliable and clean data to ensure accurate and fair outcomes. Data used in models should be rigorously tested for accuracy, timeliness, completeness, consistency, and validity. All data quality issues should be reported, documented, and managed. Processes should be introduced to prevent poor quality data from being captured and used in scoring. When using third-party models, consider whether controls in place are appropriate to ensure their models are built on clean and reliable datasets.

3.4 **Treat protected characteristics carefully –** Although in a limited number of circumstances, models may use protected characteristics (e.g., age in a medical risk model), consider if the use of protected characteristics and their close proxies is appropriate in each context. Any justification for the use of protected characteristics or their proxies should be documented, including consideration of legal basis and proportionality. When using third-party models, consider if appropriate controls are in place to identify and justify the use of protected characteristics or their proxies, which may lead to discriminatory AI decisions and outcomes.

3.5 **Understand who is impacted –** Machine learning models should be visualised to understand who is impacted to inform decision-making. The visualisations should provide information such as: who has the highest score, who will be charged the most, and which customer groups might have inaccurate scores. Scores and error rates for different customer groups should be reviewed to ensure no group is negatively penalised.

3.6 **Have a test plan –** Before building or deploying AI or machine learning models, a test process needs to be considered to document strengths and weaknesses, which should be communicated to decision-makers. The results of testing should focus on the limitations of AI and where it makes mistakes, and the testing should occur for both in-house and third-party models and tools. Generative AI has several known limitations such as bias (e.g., gender, racial) and hallucinations (inaccuracies) therefore, it is important that any use of generative AI is subject to both human oversight and regular testing.

3.7 **Focus on causal links not merely correlation –** Consider if there is a reasonable causal link between the model variables and the outcomes of the model. As an industry, we should ensure the relationships in the models used are logical and causal. A written description of the direction and causality of each variable in machine learning models should be produced and independently reviewed before a model is deployed.

3.8 **Consider an appropriate level of accuracy –** Levels of required accuracy should be established pre-build or during the procurement stage, and only models meeting this level of accuracy should be used. This will mitigate the risk of using poor performing models (over or underfitting), which can result in unfair outcomes. We have a duty to be as accurate as possible when predicting risk, therefore, machine learning models should be monitored to ensure high performance and detect drift throughout the lifecycle. For third-party models and generative AI, we recommend randomised testing to ensure performance remains high.

A set of good practice examples relating to AI solutions (continued)

3.9 **Human review and independent challenge –** Fairness is a human and not purely a mathematical concern. Appropriate human oversight and monitoring mechanisms should be in place, and we recommend that all models are subject to independent challenge from technical experts. Technical experts should act as the independent voices of the customer.

4 Accountability and Governance

4.1 **Consider whether the use of AI is necessary –** Not every problem can be solved with a model or algorithm, and sometimes the fairest approaches avoid the uses of predictive analytics and AI entirely. Before beginning a project, consider the benefits and limitations of using AI in that context. Also, document your justification for the use of AI to show that it was the most appropriate and proportionate way of achieving a specific purpose.

4.2 **AI catalogue / inventory –** All uses of machine learning should be fully documented within an inventory/product management system. Documentation should have sufficient detail to allow another data scientist or developer to replicate the model from scratch. Having an individual location allows an organisation to access this documentation easily and efficiently and facilitates more effective audit and governance.

4.3 **Data catalogue –** All data used within machine learning models or used within Generative AI should be documented within a data catalogue, which will allow developers to understand how data should be utilised, its limitations, data owners and acceptable uses.

4.4 **Documented roles and responsibilities (RACI) –** When building or using AI, it is important to have documented roles and responsibilities to ensure accountability and effective oversight of management of AI risks throughout the AI lifecycle. For in-house models, this will be a named list of individuals designed in the build and utilisation of a model (e.g., data scientists, data engineers, data owner, management, testers, stakeholders, governance etc). For Generative AI and third-party AI, the list should include all individuals involved in procurement, internal implementation, testing and utilisation.

4.5 **Training and awareness –** Staff and stakeholders should be educated about the ethical considerations and risks associated with AI to reduce the likelihood of unintended consequences and ensure responsible AI development and governance. Dedicated training should be delivered to staff developing models and those responsible for reviewing and signing off models.

4.6 **Technical support for decision makers –** Independent technical experts should be able to provide challenge and recommendation on relevant working groups and committees upon request.

4.7 **Escalation and communication channels –** Machine learning and AI projects require clear escalation and communication channels to senior management to ensure awareness, accountability, and control.

4.8 **Only work with trusted third-party suppliers –** When using customer information from third parties, we will have imperfect knowledge about the processes that information has gone through. It could be possible that the data has not been collected with the ethical standards we hold ourselves to. Therefore, it is important that we work with suppliers who are transparent about their practices and have taken proactive steps to ensure fair customer outcomes.

5 Contestability and Redress

5.1 **Empower query teams –** Staff dealing with customer queries regarding AI decisions and outcomes should have the tools and training necessary to explain AI outputs to customers and regulators. Techniques to explain and visualise AI solutions (e.g. SHAP, LIME) should be developed, and query teams should be provided with full training on the operation of predictive models.

5.2 **Processes for challenging outcomes –** Processes should be made available to customers to allow them to challenge AI decisions and outcomes which negatively impact them.

5.3 **Processes for correcting data –** Processes should be made available to customers to allow them to flag inaccurate information used in AI decision-making. Any inaccurate information should be updated to ensure accurate outcomes.

7. Overview of regulations and legislation with application to AI

Overview

This chapter is designed to highlight the existing UK regulations that already align with the UK government's five principles of AI Regulation. These regulations will change over time, so this chapter only acts as a point in time reference, as of February 2024.

Background

There are multiple regulations and legislation that govern and guide our industry in the United Kingdom. Many of these have an impact on how and where we apply technology. In the context of the application of AI there are a few that are consistent across the five principles. Specifically, these include:

- FCA Consumer Duty
- FCA Handbook
- ICO – UK GDPR and Data Protection Act
- PRA – Solvency II – Systems of Governance
- Equality Act 2010

Outside the UK, the European Union is in the final stages of agreeing the EU AI Act. The principles in the EU AI Act are broadly comparable to the UK's five AI principles, and focus on addressing data quality, transparency, human oversight and accountability. If a business operates within the EU, then it will be subject to the EU AI Act.

The AI Act introduces an AI Classification system that determines the level of risk an AI solution could present to individuals. The four risk classifications are:

1. **Unacceptable risk –** Application of AI that is banned within the EU, for example social scoring and monitoring of people.
2. **High risk –** AI that controls access to financial services, critical infrastructure or employment is subject to strict conformity assessment and monitoring.
3. **Limited Risk –** Examples include chatbots where there are specific transparency obligations. For example users should be aware that they are interacting with AI.
4. **Minimal or no risk –** For example spam filters.

AI solutions within the Life and Health Insurance industry will be classified as High risk. They will be subject to strict conformity assessments to determine that they meet the requirements of the AI act.

The EU AI Act also includes provision against the threats of General-Purpose AI (GPAI) systems. These are powerful AI systems that could present a systemic risk. They are subject to additional regulation. Whilst OpenAI GPT-3.5 would not be in scope of this regulation, GPT-4 would be in scope.

Overview of regulations and legislation with application to AI (continued)

1 Safety, Security & Robustness

Addressing the operational practices of safety and security challenges of complex AI systems is critical to fostering trust in AI. In this context, robustness signifies the ability to withstand or overcome adverse conditions, including digital security risks.

This principle further states that AI systems should not pose unreasonable safety risks including to physical security, in conditions of normal or foreseeable use or misuse throughout their lifecycle.

> ***"AI systems should function in a robust, secure and safe way throughout the AI life cycle, and risks should be continually identified, assessed and managed."***
>
> – UK AI Regulation Policy Paper

Existing laws and regulations in areas such as consumer protection already identify what constitutes unreasonable safety risks and GDPR/DPA2018 covers security.

How do existing regulations support this principle?

ICO – UK GDPR – Security and data minimisation in AI

The ICO publishes guidance on the security and data minimisation of AI for Data Protection Officers. This includes guidance on what steps should be taken to manage the risk of privacy attacks on AI models and data minimisation and privacy preserving techniques for AI systems.

The ICO provides a useful AI and Data Protection risk toolkit to help organisations to consider the complete risks when developing an AI solution and what practical steps to mitigate the risks.

Overview of regulations and legislation with application to AI (continued)

The primary articles relevant to UK GDPR include:

- Article 32 - requires Data Controllers and Data Processors to implement technical and organisational measures that ensure a level of data security appropriate for the level of risk presented by processing personal data.
- Article 33 - mandates that in the event of a personal data breach, the data controller must notify the supervisory authority without undue delay, and where feasible, not later than 72 hours after becoming aware of it, unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons.
- Article 35 - requires organisations to carry out a Data Protection Impact Assessment (DPIA) prior to the processing, where a type of processing, in particular using new technologies, is likely to result in a high risk to the rights and freedoms of natural persons.

FCA Consumer Duty

As part of operating within the Consumer Duty, organisations are required to have a robust governance model in place. Whilst no reference is made to AI Safety, Security and robustness inherently, the nature of this governance model will require the maintenance of a secure and robust AI capability.

Further guidance

The UK government has published a set of AI Assurance techniques for anyone involved in the designing, developing, deploying or procuring AI enabled systems.

The National Cyber Security Centre's Principles for the security of machine learning, and industry standards/ frameworks such as ISO270001 should be applied to AI. They provide context and structure to help data scientists, engineers, business owners and risk owners make educated decisions about system design and development processes, helping to assess the specific threats to a system.

2 Appropriate Transparency & Explainability

Organisations developing and deploying AI should be able to communicate when and how it is used and explain a system's decision-making process in an appropriate level of detail that matches the risks posed by the use of AI.

> ***"AI systems should be appropriately transparent and explainable. Transparency refers to the communication of appropriate information about an AI system to relevant people (for example, information on how, when, and for which purposes an AI system is being used). Explainability refers to the extent to which it is possible for relevant parties to access, interpret and understand the decision-making processes of an AI system."***
>
> – UK AI Regulation Policy Paper

How do existing regulations support this principle?

FCA Consumer Duty

The FCA's Consumer Duty rules (FG22/5: Final non-Handbook Guidance for firms on the Consumer Duty (fca.org.uk)) include a requirement to act in good faith. The FCA give as an example of not acting in good faith:

> ***"Using algorithms, including machine learning or artificial intelligence, within products or services in ways that could lead to consumer harm. This might apply where algorithms embed or amplify bias and lead to outcomes that are systematically worse for some groups of customers, unless differences in outcome can be justified objectively."***

Clear explanations to justify an outcome objectively in this situation would require decisions being made by the machine learning or AI tools.

Overview of regulations and legislation with application to AI (continued)

ICO Data Protection

The ICO has produced detailed guidance on explaining decisions made with AI – Explaining decisions made with AI and the Guidance on AI and Data Protection.

In this paper, the ICO refers to various categories of explanation: Process-based vs Outcome-based; Rationale explanation; Responsibility explanation; Data explanation; Fairness explanation; Safety and performance explanation; and Impact Explanation.

The guidance makes reference to the UK GDPR and the Data Protection Act, noting the particular importance of explaining decisions where personal data is being used by the AI model.

In the guidance they note that some firms may be concerned about commercial sensitivity, but they do not expect decisions to be at the sort of level of detail that would leak any proprietary information.

Equality Act 2010

The Equality and Human Rights Commission has published guidance regarding unlawful discrimination in relation to financial services providers: Equality law – Banks and other financial services providers | EHRC (equalityhumanrights.com).

If you are using an AI system in your decision-making process, you need to ensure, and be able to show, that this does not result in unlawful discrimination.

Overview of regulations and legislation with application to AI (continued)

3 Fairness

AI should be used in a way which complies with the UK's existing laws, for example the Equality Act 2010 and UK GDPR, and must not discriminate against individuals or create unfair commercial outcomes.

> ***"AI systems should not undermine the legal rights of individuals or organisations, discriminate unfairly against individuals or create unfair market outcomes. Actors involved in all stages of the AI life cycle should consider definitions of fairness that are appropriate to a system's use, outcomes and the application of relevant law."***
>
> – UK AI Regulation Policy Paper

The concept of Fairness is embedded across many areas of law and we already have some guidance from the existing regulatory framework (Human Rights Law, Consumer Protection & Data Protection Law).

How do existing regulations support this principle?

FCA Consumer Protection

- The FCA's approach to consumer protection is founded in their Principles for Business, guidance and both high-level & detailed rules.
- The Consumer Duty establishes a higher standard than the existing Principles regarding how firms should treat retail customers – with a requirement to deliver "good customer outcomes".
- Consumer Duty also seeks to address discriminatory harms by requiring firms to consider the diverse needs of their customers, including those with characteristics of vulnerability or protected characteristics.
- AI-derived pricing strategies that differentiate between different groups of customers could breach the requirements if they result in poor outcomes for particular groups of retail customers.
- Consequently, firms should monitor, explain, and justify if their AI models result in differences in price and value for different cohorts of customers.

Equality Act 2010

- Where firms utilise AI in their decision-making process, they must ensure it does not result in unlawful discrimination based upon nine protected characteristics.
- The FCA's Vulnerable Customer Guidance notes that firms must give regard to the Equality Act and many of the characteristics of vulnerability overlap with protected characteristics.
- As such, any breach of the Equality Act (such as discriminatory decisions made by AI systems) could violate FCA rules and be subject to action from the regulator.

ICO – Data Protection

- If firms use AI to process personal data, they must comply with regulatory obligations set out under UK GDPR and the Data Protection Act 2018.
- The Information Commissioner's Office (ICO) has responsibility for enforcing compliance with data protection measures.

The ICO has published guidance on how to interpret the UK GDPR's Fairness principle as it applies in an AI context. It notes the importance to fairness of UK GDPR Article 22 safeguards on solely automated decision making and profiling. It also notes that data protection requirements should be read in conjunction with other legislation and regulations where fairness and discrimination exist as concepts.

Overview of regulations and legislation with application to AI (continued)

4 Accountability & Governance

AI oversight should be appropriate based on the way AI is being used and there should be clear accountability for AI outcomes.

> *"Governance measures should be in place to ensure effective oversight of the supply and use of AI systems, with clear lines of accountability established across the AI life cycle. AI life cycle actors should take steps to consider, incorporate and adhere to the principles and introduce measures necessary for the effective implementation of the principles at all stages of the AI life cycle."*
>
> – UK AI Regulation Policy Paper – page 30

How do existing regulations support this principle?

FCA Consumer Duty

Firms need to ensure that their Consumer Duty governance and oversight structures act to identify foreseeable harm and that the technology is used in good faith.

At each stage of development and deployment firms will need to collect and analyse Management Information (MI) to detect, identify and rectify any poor outcomes experienced by customers arising from the use of AI. Firms will need to consider potential harm under the four stated outcomes for consumers, which relate to products and services, price and value, consumer understanding, and consumer support.

For example, what governance and testing is in place to ensure that content created with the use of AI is understandable to customers? What MI is collected to ensure that the deployment of chatbots meets the support needs of customers, especially those who may be vulnerable?

Within the annual report and attestation process, the firm should be able to demonstrate effective management of these risks along with the prevention and remediation of poor customer outcomes.

FCA Handbook – Threshold Conditions (COND)

A firm is required to meet the FCA's Threshold Conditions as a requirement of their authorisation to carry out regulated activity. A requirement of threshold conditions is that a firm must possess adequate non-financial resources.

In order to exercise its governance and oversight of the use of AI within the firm, it should give consideration as to whether it has sufficient resources to do so as required by the Threshold Conditions. This may include whether the firm has the required number of people and the necessary skills and competence.

FCA Handbook – Systems & Controls (SYSC)

SYSC requires a firm to take reasonable care to establish and maintain such systems and controls as are appropriate to its business.

The existing systems and controls that a firm is required to implement, of which the key ones are outlined in its appendices, would also extend to the use of AI. A firm is required to enact appropriate policies, procedures and controls to identify and mitigate the risks associated with the use of the technology.

The firm is required to have sufficient understanding of, and access to, the technology to provide effective oversight. Firms seeking to implement effective model risk management can integrate this into their current control environment and leverage the risk management and oversight required by SYSC at all stages of the AI lifecycle.

Senior Manager & Certification Regime (SMCR)

SMCR is the regulator's primary tool for establishing whom within a firm is responsible for a particular area or function. The regulators have confirmed that they will seek to utilise these rules to assign specified responsibility and accountability for the use of AI within firms in financial services.

Firms will need to consider with whom accountability for the oversight of the use of AI should sit, ensuring that the individual has sufficient training and competence to discharge the role. This should be documented within the firm's Responsibilities Map, with the accountable senior manager's Statement of Responsibility defining their responsibilities in this role.

Other senior managers other than the individual accountable SMF holder may be involved in the deployment and use of AI. In such instances thought should be given to updating the Statement of Responsibility for those who are also responsible for the use of these systems.

Overview of regulations and legislation with application to AI (continued)

Prudential Regulation in the UK

Insurers regulated by the PRA are required to adhere to the requirements of the retained EU Solvency II Directive, which has undergone some key UK-specific reforms since the UK left the EU in January 2020. The Solvency II regime will be replaced by Solvency UK on 31 December 2024, when the remaining retained EU Solvency II legislation will be deleted, and the majority of the rules will be transferred to the PRA's Rulebook. A key part of the qualitative elements of pillar II of Solvency II is the Systems of Governance requirements. It will be necessary for a firm to, amongst other things ensure that the board has an effective Governance system in place that allows for the sound and prudent management of the business. This would capture oversight of the use of AI.

To comply with the above requirements, firms will need to

- identify the committees and subcommittees that will provide appropriate checks on the use of the technology.
- ensure that appropriate policies and procedures for the use of AI is in place, that they are reviewed annually and approved by the governing bodies; and
- ensure that the Risk and Compliance and third line Internal Audit teams are effective in their assurance and advisory roles to support and oversee the use of AI.

When Solvency II is replaced by Solvency UK, the fundamental principles for the systems of governance within a firm will remain.

Further guidance

The FCA and PRA provide guidance on preparing, preventing and recovering from operational disruption.

Consideration should also be given to the role and responsibilities of third party outsources and service providers in the design, training, delivery and maintenance of systems that contain AI capabilities. The Financial Services and Markets Bill has provisions for the oversight of designated Critical Third Parties (CTPs) in Chapter 3C.

Overview of regulations and legislation with application to AI (continued)

5 Contestability & Redress (C&R)

People need to have clear routes to dispute harmful outcomes or decisions generated by AI.

> *"Where appropriate, users, impacted third parties and actors in the AI life cycle should be able to contest an AI decision or outcome that is harmful or creates material risk of harm. Regulators will be expected to clarify existing routes to contestability and redress, and implement proportionate measures to ensure that the outcomes of AI use are contestable where appropriate."*
>
> – UK AI Policy Paper, page 31

How do existing regulations support this principle?

Consumer Duty

The Consumer Duty stipulates that firms must design products and services that aim to secure good consumer outcomes. And they must demonstrate how all parts of their supply chain deliver these throughout the customer journey.

ICO – UK GDPR – Guidance for AI

The UK GDPR specifies that an organisation must:

- Be proactive in giving individuals meaningful information about the logic involved, as well as the significance and envisaged consequences.
- Give individuals at least the right to obtain human intervention on the part of the controller, to express their point of view and to context the decision.

ICO – Individual rights in an AI system

The ICO sets out the rights of the individual within an AI system and offers guidance on the right to rectification. This requires companies to check and potentially rectify the accuracy of personal data used in training data as it relates to AI outcomes.

ICO – ICO25

The ICO is putting in place a Subject Access Request tool to help individuals make requests in ways that enable organisations to respond effectively. This tool will help people identify where to send their requests and explain what they should expect. The receiving organisation will receive information from the ICO to help them respond quickly and simply.

Parts 3 and 4 of the DPA 2018

Protection for solely automated decisions that have an adverse legal effect or significantly affect the data subject, and which are carried out for law enforcement purposes. Individuals can obtain human intervention, express their point of view, and obtain an explanation of the decision and challenge it.

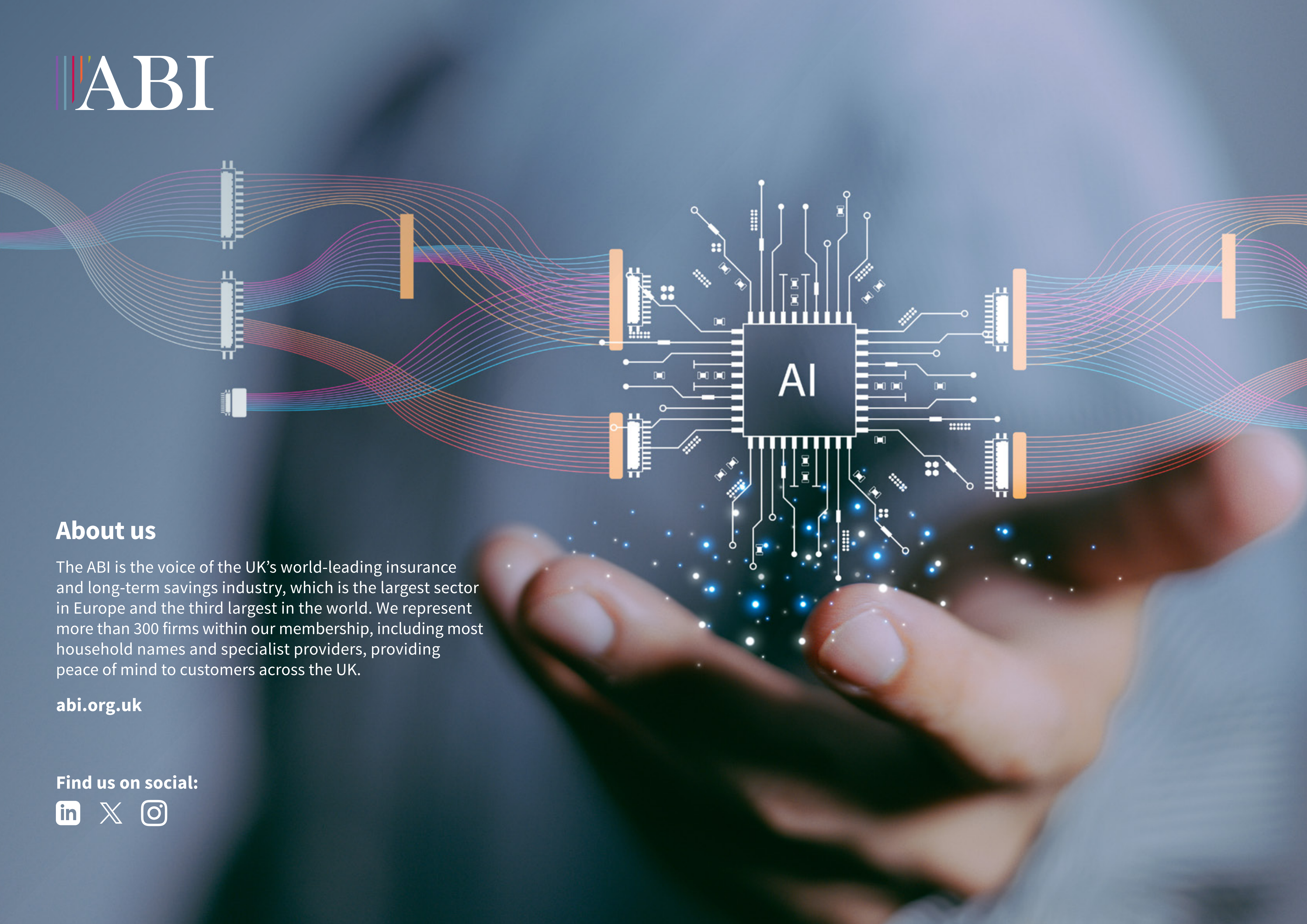
FCA Handbook – DISP complaint handling rules

The DISP rules require a firm to have processes and controls to effectively identify, investigate, manage and resolve complaints in a timely way. The individual raising a complaint may also have recourse to the Financial Ombudsman Service (FOS). Eligibility and other requirements for dealing with complaints are further outlined in the DISP rules.

Equality Act 2010

If you are using an AI system in your decision-making process, you need to ensure, and be able to show, that this does not result in discrimination that:

- causes the decision recipient to be treated worse than someone else because of one of these protected characteristics; or
- results in a worse impact on someone with a protected characteristic than someone without.



About us

The ABI is the voice of the UK’s world-leading insurance and long-term savings industry, which is the largest sector in Europe and the third largest in the world. We represent more than 300 firms within our membership, including most household names and specialist providers, providing peace of mind to customers across the UK.

abi.org.uk

Find us on social: