CMORG
CROSS MARKET OPERATIONAL RESILIENCE GROUP

CITY OF LONDON

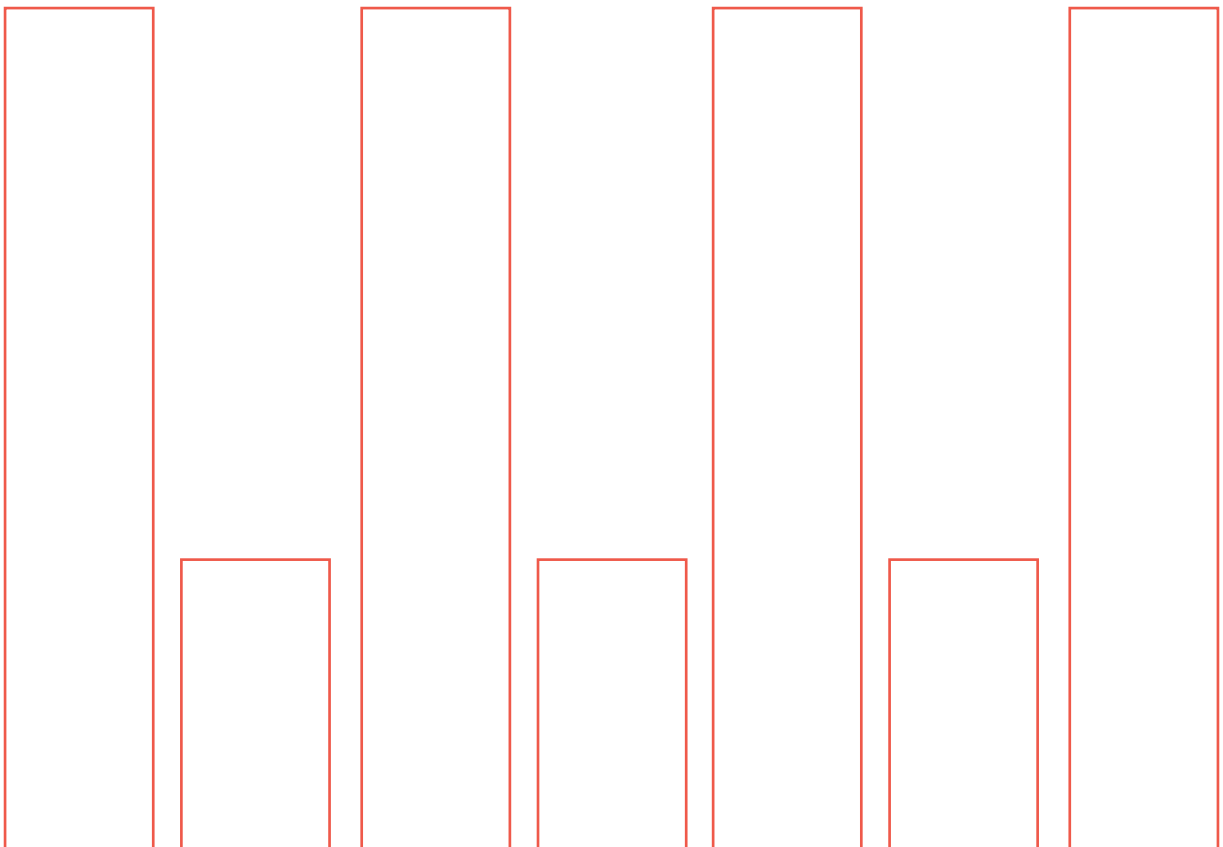
FS-ISAC

UK Finance

AI BASELINE GUIDANCE REVIEW

AI TASKFORCE

VERSION 1.0 | JANUARY 2025 | TLP CLEAR



CONTENTS

1 EXECUTIVE SUMMARY

1.1 BACKGROUND AND OBJECTIVES

The Cross Market Operational Resilience Group's (CMORG) AI Taskforce was created in 2024 as a joint initiative of the CIO Forum and Cyber Coordination Group, in response to concerns relating to sector-level risk introduced by the rapid adoption of Generative AI (Gen-AI).

An immediate priority activity identified by the Taskforce was to conduct a baseline review of existing Gen-AI risk mitigation guidance and develop a composite financial sector-specific view of good practice, supporting capability building among firms. This review has been undertaken by subject matter experts (SMEs) from CMORG member firms, alongside representatives from UK Finance, City of London Corporation (CoLC) and the Financial Services Information Sharing and Analysis Center (FS-ISAC).

The output of this document is a thematically focused guide to support finance sector firms in responding to the potential risks posed by Gen-AI. The content has been endorsed by CMORG, with additional input and support from UK Finance, CoLC and FS-ISAC, as a summary guide to good practice in respect of Gen-AI risk management for firms.

Alongside a quick reference list of key takeaway risk mitigation actions for firms, the document provides a useful reference point for the most relevant reading materials available to build deeper understanding of this complex and fast evolving risk area.

1.2 HOW TO USE THIS DOCUMENT

To help with navigation, findings are split out under the following thematic categories:

1.Government & Regulatory	2. Risk Management Principles & Frameworks	3. Technical Implementation Guidance	4. Third Party & Legal Considerations	5. Education & Awareness
High-level overview of how authorities are balancing the competing forces of Gen-AI opportunity and risk, including a snapshot of emerging regulation.	Gen-AI principles and risk frameworks firms may wish to adopt to provide confidence related risk is being managed effectively.	Guidance and standards firms should consider when deploying control frameworks to manage the risks associated with Gen-AI adoption and implementation.	Considerations for firms in respect of managing third party and legal risk arising from Gen-AI usage.	Guidance on how to build and embed a 'responsible AI' culture and upskill colleagues to mitigate Gen-AI risks and threats.

Within each section you will find:

- A summary of key observations from the review of related documentation.
- Key takeaways for finance sector firms to consider in their own organisational context.
- A list of further reading for reference and deeper understanding.

Additional 'spotlight' sections provide further insights into pertinent areas of Gen-AI thinking. Note that this document is focused on Gen-AI specifically, so not all generic AI resources are included.

1.3 KEY TAKEAWAYS

The below tables set out key takeaway actions for firms striving to implement effective management of Gen-AI risk.

1. Government & Regulatory	2. Risk Management Principles & Frameworks	3. Technical Implementation Guidance	4. Third Party & Legal Considerations	5. Education & Awareness
○ Assess impact of emerging regulatory compliance timelines, notably the EU AI Act. ○ Map out the Gen-AI regulatory regimes that impact you, noting any tensions and inconsistencies between requirements. ○ Participate actively in sector engagement to influence emerging regulation and manage broader risk.	○ Agree your firm's position on Gen-AI adoption in consultation with your board, reflecting risk appetite, current / expected usage, desired benefits, compliance and regulatory requirements. ○ Update governance and risk frameworks to include adequate coverage for Gen-AI related risks, aligned to firm risk appetite. ○ Reference industry frameworks to ensure comprehensive coverage of novel risks, e.g. ISO, NIST, FS-ISAC.	○ Implement robust controls with reference to industry good practice guidance and standards for Gen-AI adoption including: a) Data Protection & Privacy. b) Cyber & Information Security. c) Model Risk. ○ Continuously review and update controls in response to rapidly evolving regulatory and technology landscapes. *See the next figure for further detail.*	○ Update third party risks & control frameworks to mitigate Gen-AI specific risk. ○ Identify all third party Gen-AI use, for both existing and new / proposed suppliers and enforce mandatory notification for new adoption. ○ Implement processes to ensure permitted uses of third party Gen-AI are defined understood and adhered to. ○ Ensure legal experts are consulted in respect of Gen-AI adoption, notably re IP, liability and contractual risk.	○ Agree and communicate an Acceptable Use Policy (AUP) for Gen-AI to clarify colleague accountability and set guardrails, aligned to firm risk appetite. ○ Undertake extensive training and awareness for colleagues, including AUP, vital importance of human-in-the-loop, and Gen-AI technical skills. ○ Uplift existing security education & awareness programmes to address risk of Gen-AI powered attacks, e.g. phishing, deepfakes.

3a. Data Protection & Privacy	3b. Cyber & Information Security	3c. Model Risk
○ Identify and understand all data storage, retention, third party transfers and international transfers involved in Gen-AI processing. ○ Set and enforce appropriate use case guidelines for Gen-AI models in the development phase, based on strong data protection principles. ○ Understand, limit and protect any personal data used to train or finetune large language models (LLMs), setting appropriate controls to restrict access to datasets based on user roles and responsibilities. ○ Implement assurance mechanisms, including regular audits and model performance testing, and carry out exercises to identify unintended outcomes, such as bias or discrimination. ○ Consider third party management practices, clearly defining the accountability and liability of developer vs deployer.	○ Expand established Secure Systems Development Lifecycle (SSDLC) good practices to adopt a Gen-AI specific threat modelling approach and mandate its usage during design stage to combat novel threats e.g. jailbreaking, prompt injection. ○ Ensure that Gen-AI capabilities inherit the access rights of the user's session rather than any broader access rights and implement procedures to report any inappropriate access rights surfaced via Gen-AI usage. ○ Maintain a watching brief on emerging Gen-AI tooling, to address novel Gen-AI risks and to augment defensive teams in addressing both traditional and Gen-AI related threats.	○ Expand established model risk management processes to ensure Gen-AI specific risks are addressed, notably: accuracy, explainability and bias. ○ Implement measures to ensure that quality data inputs underpin any Gen-AI systems. ○ Ensure Gen-AI systems are subject to ongoing testing and monitoring against model risk, including ensuring false inferences are not drawn. ○ Prioritise upskilling to ensure all colleagues understand the critical importance of human interaction to mitigate Gen-AI model risk.

1.4 ESSENTIAL READING

This section highlights five key documents, one for each thematic, which we have recommended for senior individuals looking to develop a strategic view of key guidance. While a full read-through of this document is encouraged, these five resources provide critical insight into the potential risks posed by Gen-AI and essential information to help you quickly understand this fast evolving risk area.

1. Government & Regulatory

UK AI Whitepaper: Policy paper: A pro-innovation approach to AI regulation March 2023

- https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper#executive-summary
- The UK AI Whitepaper emphasises a pro-innovation approach, aiming to create a flexible regulatory framework and introduces no new penalties. It relies on existing regulators to prepare any necessary guidance to manage AI risks in their domains.

2. Risk Management Principles & Frameworks

AI Risk Management Framework: Generative Artificial Intelligence Profile | NIST July 2024

- Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile | NIST
- The Gen-AI Risk Profile is designed to help organisations to decide how to best manage risks which are novel to or exacerbated by the use of Gen-AI, including suggested mitigations.

3. Technical Implementation Guidance

Cyber Security risks to AI | UK Department for Science, Innovation and Technology (DSIT) & Mindgard February 2024

- https://assets.publishing.service.gov.uk/media/663cf205bd01f5ed32793891/Cyber_Security_for_AI_recommendations_-_Mindgard_Report.pdf
- Makes both technical and general recommendations to address cyber security risks across different phases of the AI development life cycle based on comprehensive review of existing guidance documentation and reported cyber-attacks against AI.

4. Third Party & Legal Considerations

Assurance of Third-Party AI Systems for UK National Security | Alan Turing Institute January 2024

- Assurance of Third-Party AI Systems for UK National Security | Centre for Emerging Technology and Security
- The Centre for Emerging Technologies and Security research report is intended to inform the reader of the risks associated with the evolving design and development of AI capabilities and support the mitigation of these risks through and AI assurance framework.

5. Education & Awareness

Framework of an Acceptable Use Policy for External Generative AI | FS-ISAC September 2023

- *FrameworkOfAnAcceptableUsePolicyForExternalGenerativeAI.pdf*
- Offering guidance on both permissive and stringent approaches, this adjustable policy helps organisations to arrive at the right balance in designing a Generative AI Acceptable Use Policy that aligns to their specific circumstances.

2 REVIEW SUMMARIES

2.1 GOVERNMENT AND REGULATORY APPROACHES

Summary:

- Governments worldwide are exploring the potential risks posed by the emerging use of Gen-AI at sectoral, national and global levels. They are working with industry and relevant authorities to shape approaches to managing these risks and building public trust.
- There is broad agreement that a balanced approach is required to avoid stifling the innovative potential of Gen-AI technology to drive growth and solve some of humanity's most pressing problems.
- Regulations are emerging, notably the EU AI Act, with compliance deadlines through 2025 and 2026.

Overview:

Governments and regulators worldwide are trying to strike a balance between harnessing the capability of Gen-AI to drive economic growth and scientific progress whilst managing the associated risks and building public trust in this emerging technology. Although this document focuses on Gen-AI specifically, policy development often encompasses AI more broadly.

There is enthusiasm for AI technology among policy makers, including to solve pressing global problems, such as climate change. At the same time, there is broad recognition of the risks at sectoral and national levels, with developers urged to take care that their products respect the rule of law and human rights.

Whilst international cooperation is championed in principle to avoid regulatory fragmentation, governments naturally differ in approaches at this relatively early stage of the technology's development. Some prefer guidance to promote safe, secure, and trustworthy development of advanced AI. These regimes tend to rely on existing regulation, rather than new legal instruments. In other jurisdictions, bespoke 'hard law' AI regulation is emerging, notably the EU's AI Act.

Policy papers differ not only in their overall approach to managing AI risks but also in key details, such as their definitions of AI, which can increase compliance complexity for firms. Despite these differences, there are also common points between national and regional regimes. Generally speaking, a collaborative approach that advances public-private dialogue is championed even where hard law is also being developed. Risk-based approaches are generally preferred, tailoring rules to the level of risk of different AI applications. Also, there is often a focus on the role of standards under both regulatory models.

In parallel to government consideration, public authorities are acting in their own domains. Regulatory consultations are being issued on specific risk areas, e.g. cyber security. Some of these result in fresh guidelines and may later lead into more formal rules in some sectors.

Many policy and regulatory processes are considering AI holistically, though some contain specific elements focused on Gen-AI. In other cases, certain types of generic guidance may be particularly relevant to Gen-AI, such as third party risk management regulation.

Spotlight: The EU AI Act

The EU AI Act's model is to prohibit AI that violates EU fundamental rights and values, control AI that impacts on health, safety and fundamental rights, and require transparency for AI that creates a risk of impersonation, manipulation or deception.

The June 2024 AI Act is the first binding worldwide horizontal (cross-economy) regulation on AI for the use and supply of AI systems in the EU. It introduces potential fines of up to 7% of global turnover for use of AI for prohibited purposes and requires that AI systems in financial services and elsewhere adhere to specific standards of safety and ethics.

The majority of obligations under the act fall on providers (developers) of high-risk AI and on providers of 'general-purpose AI' systems (GPAI) but there are more limited obligations on deployers of AI systems. Organisations may act as both provider and deployer; notably, a firm that buys in an external model but then retrains it may be subject to the full set of provider obligations.

To identify obligations under the act, consider three key aspects for each Gen-AI use case:

- Whether the AI is GPAI;
- The organisation's role in the AI use (typically deployer and/or provider); and
- The four risk levels identified in the act:

Prohibited Risk	• AI Systems posing unacceptable risks are prohibited. • E.g. social scoring systems, manipulative AI, emotional interference in the workplace, certain crime prediction use cases.
High Risk	• Adoption of AI in high risk processes is subject to obligations under the Act, including standards for transparency, accountability and robustness. • E.g. recruitment and performance evaluation, creditworthiness assessment.
Limited Risk	• Adoption of AI in limited risk processes in subject to lighter 'transparency' obligations, i.e. ensure users are aware they are interacting with AI, e.g. use of chatbots.
Minimal Risk	• Systems with minimal risk are unregulated under the Act, e.g. AI enabled video games and spam filters.

There are outstanding areas of uncertainty, including the exact definition of 'AI', scope of 'creditworthiness' under the high-risk category, and the scope of prohibited use cases. Guidance has been issued but it is complex.

Firms need to take immediate action, with requirements as follows:

- 2 February 2025 for prohibited uses of AI.
- 2 August 2025 for 'general purpose' AI (covers many Gen-AI models).
- 2 August 2026 for high-risk uses of AI (includes creditworthiness).

Key Takeaways – Firms Should:

- Assess the immediate and ongoing impacts of emerging regulatory compliance timelines, notably at this stage the EU AI Act.
- Map out the Gen-AI regulatory regimes that impact their organisation, noting any tensions and inconsistencies between requirements. This can be done by using horizon consultancy scanning tools, external counsel and in-house legal expertise.

- Participate actively in sector engagement to inform and support the development of emerging regulation and help manage the broader risk to industry, for example through relevant collective action initiatives.

Further Reading:

UK AI Whitepaper: Policy paper: A pro-innovation approach to AI regulation March 2023

- https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper#executive-summary
- The UK AI Whitepaper emphasises a pro-innovation approach, aiming to create a flexible regulatory framework and introduces no new penalties. It relies on existing regulators to prepare any necessary guidance to manage AI risks in their domains.

Removing Barriers to American Leadership in Artificial Intelligence – The White House January 2025

- Removing Barriers to American Leadership in Artificial Intelligence – The White House
- Effective replacement for the executive order.

European Artificial Intelligence Act | March 2024 – also see spotlight text.

- https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html - the full legal text
- https://artificialintelligenceact.eu/high-level-summary/ - a summary of all obligations under the Act

The Impact of AI in Financial Services: Opportunities, Risks and Policy Considerations | UK Finance / Oliver Wyman November 2023

- The impact of AI in financial services.pdf
- Section 7 provides an overview of current state of AI regulation, with a specific focus on financial services considerations.

ISO/IEC Information Technology – Artificial Intelligence – Management system

- ISO/IEC 42001:2023(en), Information technology — Artificial intelligence — Management system
- This document intends to help firms responsibly use, develop, monitor or provide products or services that utilise AI.

US Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence | The White House October 2023

- https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/
- The US Executive Order seeks to advance a coordinated, federal government-wide approach to governing safe and responsible development and implementation of AI. It requires U.S. government agencies and departments to take certain actions on AI risks and sets out reporting requirements. Use of industry standards is endorsed, such as NIST AI 100-1. Developers of powerful AI systems are required to share safety test results and other critical information with the U.S. government.
- *This Executive Order has now been revoked by the US administration.*

2.2 RISK MANAGEMENT PRINCIPLES AND FRAMEWORK

Summary:

- Gen-AI can impact existing risks and also introduces novel considerations to firms, both need to be addressed.
- Frameworks are emerging to enable structured approaches to managing Gen-AI risk and building public trust.
- Risks in play include operational, reputational and compliance.

Overview:

As set out in section 2.1, governments and authorities are recognising the need to address risks effectively, in order to enable safe Gen-AI adoption and engender public trust in its usage.

While many of the risks associated with Gen-AI are common to existing technologies, such as cloud and other types of artificial intelligence / machine learning, there are some new considerations specific to, or emphasised by, the adoption of Gen-AI, which are likely to require adjustments to firms' risk management policies and processes.

Authorities, industry bodies and vendors are beginning to develop useful AI risk management frameworks to help structure risk management approaches. These frameworks offer structured templates and guidance for identifying, evaluating, and addressing the various risks linked to AI systems. They take a methodical approach to tackling these challenges, ensuring comprehensive and consistent treatment. There is no one-size-fits-all solution, but guidance is emerging which can be adapted to the organisation's risk appetite or used in conjunction with other frameworks.

Frameworks and principles are also available to assist firms in ensuring responsible and ethical adoption of Gen-AI, which is an important consideration for building consumer trust.

Firms are encouraged to consider Gen-AI through a range of risk lenses, including operational, reputational and regulatory compliance. Specific areas of focus include Data Protection and Privacy, Cyber and Information Security, Model Risk, and Third Party and Legal Risk, all of which are treated in further detail in later sections of this document.

Key Takeaways – Firms Should:

- Agree their position on Gen-AI adoption in consultation with their board, reflecting risk appetite, current and expected usage, desired benefits, compliance and regulatory requirements.
- Update governance and risk frameworks to include adequate coverage for Gen-AI related risks, aligned to the firm's risk position.
- Reference industry frameworks and principles to ensure comprehensive coverage of risks and a responsible and ethical approach to Gen-AI adoption.

Spotlight: The National Institution of Standards and Technology (NIST) Gen-AI Risk Profile

NIST's Gen-AI Risk Profile is a companion resource to the NIST AI Risk Management Framework, intended to improve the ability of organisations to incorporate trustworthiness considerations into the design, development, use and evaluation of AI products, services and systems. The profile defines 12 risks that are novel to, or exacerbated by, the use of Gen-AI:

CBRN Information or Capabilities: Eased access to or synthesis of materially nefarious information or design capabilities related to chemical, biological, radiological, or nuclear (CBRN) weapons or other dangerous materials or agents.

Confabulation: The production of confidently stated but erroneous or false content by which users may be misled or deceived.

Dangerous, Violent, or Hateful Content: Eased production of and access to violent, inciting, radicalising, or threatening content.

Data Privacy: Impacts due to leakage and unauthorised use, disclosure, or de-anonymisation of personally identifiable information or sensitive data.

Environmental Impacts: Impacts due to high compute resource utilisation in training or operating Gen-AI models, and related outcomes that may adversely impact ecosystems.

Harmful Bias or Homogenisation: Amplification of biases; performance disparities that result in discrimination, amplification of biases, or incorrect presumptions; undesired homogeneity leading to ill-founded decision-making or amplifying harmful biases.

Human-AI Configuration: Interactions between a human and AI system which can result in the human inappropriately anthropomorphising Gen-AI systems or experiencing algorithmic aversion, automation bias, over-reliance, or emotional entanglement.

Information Integrity: Lowered barrier to entry to generate and support the exchange and consumption of content which may not distinguish fact from opinion or fiction or could be leveraged for large-scale dis- and mis-information campaigns.

Information Security: Lowered barriers for offensive cyber capabilities, including via automated discovery and exploitation of vulnerabilities; increased attack surface for targeted attacks, which may compromise a system's availability or the confidentiality or integrity of training data, code etc.

Intellectual Property: Eased production or replication of alleged copyrighted, trademarked, or licensed content without authorisation; eased exposure of trade secrets; or plagiarism.

Obscene, Degrading, and/or Abusive Content: Eased production of and access to obscene, degrading, and/or abusive imagery which can cause harm.

Value Chain and Component Integration: Non-transparent or untraceable integration of upstream third party components; improper supplier vetting across the AI lifecycle; or other issues that diminish transparency or accountability for downstream users.

Further Reading:

Artificial Intelligence Risk Framework (AI RMF 1.0) | NIST January 2023

- Artificial Intelligence Risk Management Framework (AI RMF 1.0) (nist.gov)
- The goal of the AI RMF is to offer a resource to organisations to help manage the risks of AI and promote trustworthy and responsible development and usage of AI systems.

AI Risk Management Framework: Generative Artificial Intelligence Profile | NIST July 2024

- Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile | NIST
- The Gen-AI Risk Profile is designed to help organisations to decide how to best manage risks which are novel to or exacerbated by the use of Gen-AI, including suggested mitigations.

Generative AI in Finance: Risk Considerations | International Monetary Fund (IMF) August 2023

- Generative Artificial Intelligence in Finance: Risk Considerations (imf.org)
- Provides insights into Gen-AI's inherent risks and their potential impact on the financial sector and recommends close human supervision as regulatory policy evolves.

Emerging Risk and Opportunities of Generative AI for Banks | Project Mindforge consortium. Coordinated by the Monetary Authority of Singapore November 2023

- https://www.mas.gov.sg/schemes-and-initiatives/project-mindforge
- This whitepaper proposes a risk framework for Gen-AI, including high level illustrative case studies of a Gen-AI risk assessment.

Financial Services and AI – Leveraging the Advantages and Managing the Risks | FS-ISAC February 2024

- FSISAC_FinancialServicesAndAILeveragingTheAdvantagesManagingTheRisks.pdf
- Overarching paper introducing six detailed papers produced by FS-ISAC's AI Risk Working group to assist financial services firms to safely manage AI adoption.

Responsible AI Principles | FS-ISAC February 2024

- FSISAC_ResponsibleAI-Principles.pdf
- Explores foundational areas of responsible AI deployment, empowering financial institutions to align AI practices with the highest levels of ethics and trustworthiness.

Generative AI in action – opportunities and risk management in financial services | UK Finance and Accenture January 2025

- https://www.ukfinance.org.uk/policy-and-guidance/gen-ai-report
- Examines the emerging applications of Gen-AI in the financial sector and the state of the art in mitigating key generative AI risks, including three real case studies.

2.3 TECHNICAL IMPLEMENTATION GUIDANCE

Summary:

- As firms adopt and implement Gen-AI solutions, there are significant detailed technical considerations to take into account to ensure the risks are appropriately addressed.
- Authorities, industry bodies and vendors are developing detailed guidance and standards to assist organisations in designing and deploying effective controls.
- Specific areas of technical focus include Data Protection and Privacy, Cyber and Information Security, and Model Risk (Embedded Bias, Robustness / Accuracy, Explainability).
- This section sets out key considerations across each area.

Key Takeaways – Firms Should:

- Introduce mechanisms to identify new and changing Gen-AI adoption across the firm.
- Implement robust controls aligned to firm risk appetite and with reference to industry good practice guidance and standards for Gen-AI adoption, including:
    - Data Protection & Privacy (p10).
    - Cyber and Information Security (p12).
    - Model Risk – covering Embedded Bias, Robustness/Accuracy, and Explainability (p15).
- Continuously review and update controls in response to rapidly evolving regulatory and technology landscapes.
- The following sections provide further detail on each of the categories above.

2.3.1 DATA PROTECTION AND PRIVACY

Overview:

Gen-AI models are typically trained using high volumes of data and extensive datasets, which often include personal data. The increasing popularity and adoption of Gen-AI for a variety of use cases and the rise in availability of data introduces new considerations and risks to privacy, and to the rights and freedoms of individuals.

Naturally, there is overlap in the documented guidance in key areas of consideration for the collection, use and disclosure of personal data within the Gen-AI lifecycle, and guidance largely builds on existing principles and requirements. As Gen-AI is an emerging field, existing guidance is expected to change over time, and firms must remain alert to new and evolving advice.

The Information Commissioner's Office (ICO) has published five chapters on Gen-AI *(in draft)* to outline emerging thinking on how they interpret specific requirements of UK GDPR and part 2 of the Data Protection Act 2018. They have also stated plans to update and consult further on their AI guidance later in 2025.

The chapters on Gen-AI are in addition to a more general AI and data protection toolkit by the ICO, which is designed to provide practical 'best practice' support to organisations to reduce the risks to individuals' rights and freedoms caused by their own AI systems.

In Canada, the Office of the Privacy Commissioner (OPC) has also published guidance identifying considerations for the application of key privacy principles to Gen-AI technologies. The document is intended to help organisations developing, providing, or using Gen-AI, and

focuses on nine key principles. At time of writing, this was the only finalised Gen-AI guidance from a privacy regulator.

Spotlight: OPC principles for responsible, trustworthy and privacy-protective generative AI technologies

1. **Legal Authority and Consent:** Ensure legal authority for collecting and using personal information; when consent is the legal authority, it should be valid and meaningful.
2. **Appropriate Purpose:** Collection, use and disclosure of personal information should only be for appropriate purposes.
3. **Necessity and Proportionality:** Establish the necessity and proportionality of using Gen-AI, and personal information within Gen-AI systems, to achieve intended purposes.
4. **Openness:** Be open and transparent about the collection, use and disclosure of personal information and the potential risks to individuals' privacy.
5. **Accountability:** Establish accountability for compliance with privacy legislation and principles and make AI tools explainable.
6. **Individual Access:** Facilitate individuals' right to access their personal information by developing procedures that enable it to be meaningfully exercised.
7. **Limiting Collection, Use, and Disclosure:** Limit the collection, use, and disclosure of personal information to only what is needed to fulfil the explicitly specified, appropriate identified purpose.
8. **Accuracy:** Personal information must be as accurate, complete, and up to date as is necessary for purposes for which it is to be used.
9. **Safeguard:** Establish safeguards to protect personal information and mitigate potential privacy risks.

Key Takeaways – Firms Should:

- Ensure that data protection and privacy risks are mitigated by an appropriate set of controls, aligned to industry good practice.
- Identify and understand all data storage, retention, third party transfers and international transfers involved in Gen-AI processes.
- Assess third party risk management practices, clearly defining appropriate roles and responsibilities between developer and deployer, considering the degree of shared liability in relation to Gen-AI systems.
- Set and enforce appropriate use case guidelines for Gen-AI models, ensuring use cases are rigorously assessed and clearly defined during the development phase.
- Understand, limit and protect any personal data used to train or finetune LLMs, setting appropriate controls to restrict access to datasets based on user roles and responsibilities.
- Implement assurance mechanisms including regular audits and model performance testing and carry out exercises to identify unintended outcomes, such as bias or discrimination.
- Ensure human involvement during the Gen-AI lifecycle to achieve safe, ethical and responsible use, and assess and address any potential skills gaps.
- Continually monitor, review and update internal governance frameworks and policies to adapt to evolving regulatory and technological landscapes.

Further Reading:

Principles for responsible, trustworthy and privacy-protective generative AI technologies | Office of the Privacy Commissioner of Canada December 2023

- Principles for responsible, trustworthy and privacy-protective generative AI technologies - Office of the Privacy Commissioner of Canada
- Designed to help organisations developing, providing or using Gen-AI to apply key privacy principles.

Guidance on AI and Data Protection, including AI and Data Protection Risk Toolkit | ICO March 2023

- https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/
- Detailed guidance on applying UK GDPR to AI. Not limited to Gen-AI considerations.

ICO consultation series on generative AI and data protection | ICO September 2024 (currently in draft, consultation closed)

- ICO consultation series on generative AI and data protection | ICO
- Information Commissioner's Office response to the consultation series on generative AI | ICO
- Sets out draft initial thinking on five elements of UK GDPR that are key or challenging to apply to Gen-AI: 1) legal basis for web scraping and training, 2) purpose limitation, 3) accuracy, 4) individual rights, and 5) controllership (roles and responsibilities in the supply chain).

Generative AI for Organisational Use: Internal Policy Checklist | Future of Privacy Forum July 2023

- Generative-AI-Checklist.pdf
- Intended as a starting point for the development of organisational Gen-AI policies.

2.3.2 CYBER INFORMATION SECURITY

Summary:

- Gen-AI poses new types of cyber and information security threat and risk for firms, which must be recognised and addressed.
- Industry bodies are responding with standards and approaches to assist firms in managing this fast-evolving area, including the UK National Cyber Security Centre (NCSC), FS-ISAC, and The Open Worldwide Application Security Project (OWASP).
- New specialist tooling is emerging which may assist with Gen-AI risk mitigation, but this area is nascent and there is no silver bullet.
- The power of AI can also be leveraged to improve security posture. AI-enabled tooling has a role to play in defending against both traditional and Gen-AI enabled threats.

Overview:

CISOs and security practitioners need to ensure that a balanced approach is adopted to allow firms to leverage the advantages of Gen-AI without exposing themselves to unacceptable cyber and information security risk.

Gen-AI poses new types of cyber and information security threat and risk for firms, which need to be recognised and addressed in parallel to the broader implementation of Gen-AI technology. These include novel threats relating to threat actors leveraging Gen-AI to create attacks or exploiting lack of effective controls in firm developed Gen-AI capability.

Additionally, it is important that firms take steps to secure their own Gen-AI deployment effectively to avoid inadvertent data loss or other security exposures.

The high pace of Gen-AI development introduces the risk that security can be a secondary consideration. Firms should continue to ensure security is seen as a core requirement and built in from the start. Firms can implement 'Security by Design and Default' by following guidance through the Software Development Life Cycle (SDLC), such as that provided by NCSC, complemented by structured consideration of threats using resources from trusted industry sources, such as OWASP, Mitre and FS-ISAC.

As well as protecting against novel cyber-attacks such as jailbreaking or prompt injections, a structured secure by design and default approach should protect against inappropriate disclosure of sensitive information e.g. due to excessive access rights or inadvertent use of data to train third party LLMs.

AI and machine learning have been used in security tooling for many years and make a significant contribution to the effectiveness of defences in areas such as endpoint protection, intrusion detection and user behaviour analytics. There is optimism that Gen-AI has the potential to further augment defensive efforts, including through processes such as the ability to identify polymorphic malware, identification of inappropriate access permissions, and acceleration of incident response tasks. However, appropriate due diligence is required, and a watching brief is advised given the rapidly evolving vendor marketplace.

Spotlight: Jailbreaking and Prompt Injection Attacks

Jailbreaking is an attack in which a malicious actor, using specially crafted inputs (such as prompt injections), makes an AI capability perform actions, respond to, or release information beyond that which is intended by the organisation. For example, a consumer-facing chatbot might be made to release information about other consumers or respond with inappropriate comments.

Reducing the possibility of successful jailbreaking and prompt injection attacks can be achieved by ensuring that:

- The AI capability / agent inherits the access rights of the user / consumer and does not itself hold access rights to all data as part of the session.
- Background prompts and grounding processes are precisely written to limit the data the AI capability may use.
- Background prompts and grounding processes are designed to prevent the AI capability responding inappropriately or outside defined boundaries.
- Filtering of inputs and outputs is undertaken to reduce the risk that certain prompt injection words or characters can be used, or that certain output words are returned.
- Appropriate authentication of the user occurs before permitting interaction with the AI capability.
- As well as standard technical penetration testing, specific AI testing is completed by a skilled party to help identify weaknesses that could be exploited by maliciously formed prompts. Identified matters must be resolved prior to releasing capability for use.
- Monitoring of AI systems and capabilities is conducted while in use to specifically look for signs of unintended use, high-volume / rapid use and other indicators that an attempted attack may be in progress.

Key Takeaways – Firms Should:

- Review and enhance cyber and information security policies, standards and controls to ensure the related risk exposure remains aligned to risk appetite.
- Expand established SDLC good practices to adopt an AI specific threat modelling approach and mandate its usage during the design stage to combat novel threats.
- Choose Gen-AI third party providers based on their security and policy customisability to effectively address threats identified.
- Ensure that Gen-AI capabilities inherit the access rights of the user's session rather than any broader access rights.
- Implement procedures to report any inappropriate access rights surfaced by Gen-AI usage.
- Maintain a watching brief on emerging AI tooling. This will help address novel Gen-AI risks and augment defensive teams in addressing both traditional and Gen-AI related threats.
- Continually monitor, review and update internal governance frameworks, policies and controls to adapt to evolving technology and the changing threat environment.

Further Reading:

Cyber Security risks to AI | UK Department for Science, Innovation and Technology (DSIT) & Mindgard February 2024

- https://assets.publishing.service.gov.uk/media/663cf205bd01f5ed32793891/Cyber_Security_for_AI_recommendations_-_Mindgard_Report.pdf
- Makes both technical and general recommendations to address cyber security risks across different phases of the AI development life cycle based on comprehensive review of existing guidance documentation and reported cyber-attacks against AI.

CISO's guide to Generative AI and ChatGPT Enterprise Risks | Team8 April 2023

- Team8-Generative-AI-and-ChatGPT-Enterprise-Risks.pdf
- Provides CISOs, Security practitioners and others with actionable tools to understand and communicate the security implications of Gen-AI for firms and supports creation of organisational policies to enable safe and secure adoption.

Secure Design, Secure Development, Secure Deployment, Secure Operation & Testing |NCSC November 2023

- Guidelines for secure AI system development - NCSC.GOV.UK
- Recommends guidelines for providers of any systems that use AI, to help them build systems that function as intended, are available when needed and work without revealing sensitive data to unauthorised parties.

Combatting Threats and Reducing Risks posed by AI | FS-ISAC February 2024

- FSISAC_CombatingThreatsAndReducingRisksPosedByAI.pdf
- Provides recommendations and best practices to the financial service industry to combat cybersecurity threats and reduce the risks posed by AI. Designed to arm cybersecurity experts with information and defence techniques pertinent to the current threat landscape.

AI Security Risk assessment framework | Microsoft December 2021

- Best practices for AI security risk management | Microsoft Security Blog
- Designed as a first step for organisations to assess the security posture of their AI systems, building AI specific considerations onto existing traditional security risk assessment frameworks. Includes comprehensive overview of AI system security, outline of threats to critical AI assets and guidance to secure them, and framework for AI security risk assessments.

A Deeper Dive:

Adversarial AI Framework - Taxonomy, Threat Landscape and Control Frameworks | FS-ISAC February 2024

- FSISAC_Adversarial-AI-Framework-TaxonomyThreatLandscapeAndControlFrameworks.pdf
- Provides an approach to tracking and assessing AI-enabled threats in the financial services sector, specifically focusing on recent developments in Gen-AI. It captures and enhances the existing landscape of AI threat, risk and control frameworks, and attempts to standardise the AI threat taxonomy for the financial sector.

AI Risk Management Framework: Generative Artificial Intelligence Profile | NIST July 2024

- Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile | NIST
- The Gen-AI Risk Profile is designed to help organisations to decide how to best manage risks which are exacerbated by the use of Gen-AI, including suggested mitigations.

ATLAS (Adversarial Threat Landscape for AI Systems) Matrix | MITRE

- ATLAS Matrix | MITRE ATLAS™
- Globally accessible, living knowledge base of adversary tactics and techniques against AI-enabled systems based on real-world attack observations and realistic demonstrations from AI red teams and security groups. Modelled after, and complementary to the MITRE ATT&CK framework.

OWASP Top 10 for Large Language Model Applications | OWASP October 2023

- https://owasp.org/www-project-top-10-for-large-language-model-applications/assets/PDF/OWASP-Top-10-for-LLMs-2023-slides-v1_1.pdf
- A living document that aims to educate developers, designers, architects, managers and firms about the potential security risks when deploying and managing LLMsIt identifies critical security risks and related attack scenarios and recommends secure practices firms should implement to combat them.

2.3.3 MODEL RISK

Summary:

- Model risk refers to the various ways in which the implementation of a given Gen-AI model or system could inadvertently lead to the crystallisation of a firm or sector risk.
- While model risk management is a well-established practice for financial sector firms, Gen-AI introduces novel nuances which require specific treatment.
- Notable areas of model risk requiring focus by firms are Accuracy, Explainability and Bias.
- Industry bodies and authorities have developed frameworks and examples of how to implement effective approaches to ensure ‘trustworthiness’ of Gen-AI models.

Overview:

Model risk refers to the various ways in which the implementation of a given Gen-AI model or system could inadvertently lead to the crystallisation of a firm or sector risk. This includes the generation of information that is inaccurate or incomplete, also known as a model hallucination (accuracy); an inability to explain how a specific piece of information or decision was arrived at (explainability); or the generation of flawed, biased, or unfair information (bias).

- **Accuracy:** The generation of inaccurate information can result from a range of factors that impact the way in which a model operates or behaves. These include the quality and relevance of the data it is trained on, the ability of the model to generalise beyond its training data, and inconsistencies through the model development process. The potential for inaccurate outputs is inherent to the technology; the risk can be mitigated but cannot, for now, be completely removed.

- **Explainability:** The complexity of the LLMs that underpin Gen-AI systems can make it challenging to understand, and therefore to explain, how a particular outcome was reached. This is the so-called 'black box' problem. This challenge is exacerbated if there is a lack of technical understanding more broadly within a firm, which can lead to ownership and accountability not being maintained at an appropriate level of seniority.

- **Bias:** Gen-AI systems will reflect human flaws or biases in the datasets on which they are trained. This can lead to unfair bias in model outputs, or outcomes that conflict with the values and ethics of an organisation. These systems also have the potential to generate or amplify harmful content. The risk is further complicated by how difficult it is to assess for and test for bias in the sets of unstructured data that Gen-AI systems utilise. Similarly, given the diverse nature of outputs, it can be challenging to objectively identify and measure the incidence of bias, depending on the use case. This risk can overlap with accuracy and hallucination.

Model risk management in a Gen-AI context is complementary to, as opposed to distinct from, existing and well-established model risk management good practices. These should continue to be observed as a baseline approach; the nuances in operational risk that Gen-AI systems introduce should be addressed by supplementing and enhancing an organisation's overarching response.

NIST describes 'trustworthiness' as the high-level outcome of the effective implementation of a Gen-AI system. Trustworthiness may require a firm to balance trade-offs between different mitigation approaches based on the context of a firm's requirements and the model risks it faces, for instance where enhanced accuracy may undermine explainability and transparency.

To support strategic decision-making in this context, firms should consider implementing a principles- or outcomes-based approach to inform their thinking around Gen-AI model risk management, which can provide a coherent, overarching framework within which more granular implementation approaches can be configured. Both the US Treasury and the Monetary Authority of Singapore have developed examples of how this principles-based approach can be delivered in practice.

Key Takeaways – Firms Should:

- Expand established model risk management processes to ensure Gen-AI specific risks are addressed, notably: Accuracy, Explainability, and Bias, with reference to industry good practice frameworks, such as NIST's AI Risk Management Framework (see section 2.2 of this paper, and also below).
- Implement measures to ensure the quality of data inputs, including any training data, that underpin any Gen-AI system. Where a Gen-AI model is procured from a third party, this may involve model fine-tuning or the use of tools like retrieval augmented generation.
- Ensure Gen-AI systems are subject to ongoing testing and monitoring against specific model risks of concern, including to ensure that false inferences aren't drawn.

- Prioritise upskilling to ensure all colleagues understand the critical importance of human interaction within the wider Gen-AI process. This includes interactions within the model development process and at the point of usage. Organisations should prioritise developing higher levels of Gen-AI understanding more generally across their workforce to increase the 'surface area' of beneficial interactions throughout the model lifecycle.

Further Reading:

Artificial Intelligence Risk Framework (AI RMF 1.0) | NIST January 2023

- Artificial Intelligence Risk Management Framework (AI RMF 1.0) (nist.gov)
- Section 3 (AI Risks and Trustworthiness) sets out the key characteristics of trustworthy AI and offers guidance for addressing them.

Enabling an Efficient Regulatory Environment for AI – Practical Considerations for Generative AI | ASIFMA January 2024

- *Enabling an Efficient Regulatory Environment for AI - Practical Considerations for Generative AI ASIFMA*
- Section 7B Model-Related Considerations

Navigating Artificial Intelligence in Banking | Bank Policy Institute April 2024

- Navigating Artificial Intelligence in Banking - Bank Policy Institute
- Section IV. B Model Risk Management

International Scientific Report on the Safety of Advanced AI (Interim Report) | AI Seoul Summit May 2024

- International Scientific Report on the Safety of Advanced AI - GOV.UK
- International AI Safety Report 2025 - GOV.UK – full report.

Principles to promote Fairness, Ethics, Accountability and Transparency in the use of AI and Data Analytics | Monetary Authority of Singapore

- FEAT Principles Final.pdf (mas.gov.sg)

A Deeper Dive:

ML Model Access | MITRE May 2021 - case studies through a threat vector lens

2.4 THIRD PARTY AND LEGAL CONSIDERATIONS

Summary:

- The pace of change in third party AI use requires firms to rapidly integrate Gen-AI risk into existing third party risk management processes to ensure a comprehensive assessment approach is adopted for all aspects of third party AI usage.
- Risk is created both by the use of third party developed Gen-AI systems (or elements), and by the introduction of Gen-AI systems by third parties. Both of these angles need to be addressed.
- Roles and responsibilities along the supply chain must be determined, including alignment with role definitions under the EU AI Act and processor / controller roles under the GDPR. Shared responsibility models will assist in addressing complexity in this area.
- It is important to understand third party provider permitted uses of Gen-AI solutions, both for contractual compliance but also due to regulatory implications if tools are used for purposes for which they are not authorised.

- Adoption of Gen-AI introduces risk relating to liability and intellectual property, requiring review by appropriately qualified legal experts.

Overview:

Existing third-party review processes are likely to be insufficient due to the complex technical nature of Gen-AI, with a potential requirement for skill uplift within existing third party risk management teams.

Third party risk reviews will need to address novel AI risks such as those relating to use of data for model training, model poisoning, model flaws, model resilience, data corruption, AI guardrails, logging, access to data for Responsible AI review purposes, etc.

Third party privacy compliance considerations include data retention, international data transfer for model processing (which may take place in jurisdictions other than agreed storage locations), challenges relating to transparency and lawful bases for collection and processing of data, and technical documentation provision by third parties which may be required to evidence assessment of fairness and bias and potential for impact on human rights.

Legal and compliance teams will need to support contractual review, including AI clauses or addendums to address novel AI concerns, including liability and intellectual property.

Liability is a key concern, with new regulatory penalties under the EU AI Act in addition to existing regulations such as the GDPR. New and amended European AI and Product Liability Directives aim to make it easier for victims to prove liability and receive compensation in cases where AI causes harm (though the AI Liability Directive appears to have been cancelled, at least for now).

AI use is affecting intellectual property rights in multiple ways, with risks relating to both copyright infringement for input / training data and potential lack of copyright for AI-generated outputs. Organisations need to be aware of how jurisdictions are approaching concerns around IP, including issues relating to data scraping, use of third party data in AI tools and a lack of clarity on how intellectual property rules apply.

Spotlight: The FS-ISAC Gen-AI Vendor Risk Assessment Guide and Toolkit

This guide and associated workbook are a tool designed to help financial sector firms to assess and select Gen-AI vendors while managing associated risks by supplementing existing third party risk management programs with Gen-AI specific considerations.

It includes a questionnaire to gather vendor information in the following categories that are relevant to Gen-AI:

1) General (discovery)
2) Data privacy, retention and deletion.
3) Model training, validation and maintenance.
4) Information security.
5) Technology integration.
6) Nth party risk / usage.
7) Legal, regulatory and compliance.

The output from the completed questionnaire is an auto-populated worksheet designed to capture a point-in-time due diligence record. While the questionnaire includes a risk tiering to determine a due diligence plan, organisations will need to use their own internal risk rating processes and methodologies to ensure risk tiering of vendors is appropriately aligned to their unique circumstances.

Industry best practices including NIST's AI Risk Management Framework were considered in the development of the questionnaires.

Key Takeaways – Firms Should:

- Ensure that third party risks are identified and mitigated by an appropriate set of controls, aligned to industry good practice, including updating third party risk and control frameworks to mitigate Gen-AI specific risk.
- Ensure that all third party use of AI is identified, both for existing suppliers and for any new or proposed process, with a mandatory notification process to address future integration or amended use of AI in existing processes.
- Consider existing due diligence and contractual processes to determine if skill uplift is needed to address the complexities introduced by Gen-AI.
- Implement processes to ensure permitted uses of third party AI solutions are understood and adhered to.
- Consult legal experts on Gen-AI adoption, notably regarding intellectual property, liability and contractual risks.

Further Reading:

The Generative AI Vendor Evaluation & Qualitative Risk Assessment Guide and Toolset | FS-ISAC February 2024

- FSISAC_GenerativeAI-VendorEvaluation&QualitativeRiskAssessment.pdf
- Aims to describe the purpose and methodology behind the Gen-AI vendor evaluation and risk assessment workbook linked.
- FSISAC_GenerativeAI-VendorEvaluation&QualitativeRiskAssessmentTool.xlsx

Assurance of Third-Party AI Systems for UK National Security | Alan Turing Institute January 2024

- Assurance of Third-Party AI Systems for UK National Security | Centre for Emerging Technology and Security

2023 EY Global Third-Party Risk Management Survey | EY October 2023

- 2023-ey-global-third-party-risk-management-survey

Global Guide to IP Considerations for AI | Norton Rose Fulbright July 2024

- Generative AI | Global law firm | Norton Rose Fulbright

Liability Rules for Artificial Intelligence | European Commission September 2022

- Liability Rules for Artificial Intelligence - European Commission

2.5 EDUCATION AND AWARENESS

Summary:

- Embedding a 'Responsible AI' culture and mitigating Gen-AI risks and threats requires firms to invest in educating and upskilling colleagues.
- Clarity on the 'rules of the road' for Gen-AI adoption and usage can be achieved by setting out a Gen-AI Acceptable Use Policy (AUP) for the firm, and ensuring it is understood by all colleagues.
- Colleagues will require up-skilling to reap the benefits of Gen-AI, e.g. in prompt engineering, as well as education in relation to the risk of AI-augmented social engineering.
- Specialist technical up-skilling for key SME roles is vital to address associated risks.

Overview:

Given the rapid development and adoption of Gen-AI, firms must take proactive steps to make employees aware of the 'rules of the road' they must follow in respect of this emerging technology, clarifying employee accountability to ensure safe, ethical and responsible AI adoption.

Some may argue for a stringent approach to Gen-AI adoption, blocking systems that are still nascent and not fully tested and vetted. However, this removes the opportunity to harness the potential benefits of Gen-AI and may also risk pushing employees into uncontrolled workarounds. As such, enabling adoption within a clear set of guardrails aligned to firm risk appetite is likely to be a more moderate and prudent approach for most firms.

Whilst many of the risks relating to Gen-AI may be covered by existing policies, a specific Gen-AI AUP will provide clarity and help to avoid any wrong assumptions by outlining the acceptable behaviours, practices and procedures related to developing, implementing and using Gen-AI systems. The policy should be sufficiently comprehensive to apply to everyone in the organisation who may encounter / use Gen-AI systems, i.e. not just developers, and to cover both internally developed systems and those consumed from external sources.

Considerations for your Acceptable Use Policy (AUP) should include protection of sensitive data, access controls, compliance with regulatory requirements, user training and awareness, reporting violations and requesting new technology with AI capabilities.

Colleagues will require training in order to take advantage of the benefits and address the risks related to Gen-AI adoption. Areas of focus may include:

- General user awareness in relation to the risk of AI-augmented phishing and deepfake threats.
- Training users in Gen-AI specific skills such as prompt engineering to enable effective and productive adoption.
- Specialist technical upskilling for roles such as technology, security, privacy and legal teams to enable effective support.
- Awareness of the importance of 'human in the loop' in AI development and usage to address model risk.

Spotlight: The Use of AI Tools by Malicious Actors to Improve Targeted Phishing

Using AI tools, malicious actors can conduct research and profiling of people at speed, with broader coverage, increased accuracy and hence with better targeted lures for social engineering attacks such as phishing, vishing and smishing.

For example, the employees of an organisation can be profiled using professional social media sites such as LinkedIn, helping bad actors to identify people in key positions and establish their reporting lines and contacts. Using other social media, they can further understand people's hobbies and interests. All this information can be used to craft a well-targeted phishing email, perhaps also using AI to ensure convincing grammar and spelling, appropriate tone and compelling messaging, increasing the likelihood the subject will be tricked into thinking a communication is genuine.

As with all phishing, a combination of mitigation controls is required, including:

- Email filters regularly updated with phishing signatures;
- Ethical phishing tests providing instant feedback to users who click;
- Easy ways to report suspected phishing emails;
- Procedures to rapidly act on phishing attacks such as remote removal of phishing emails from user inboxes;
- Regular awareness campaigns that reference the increased sophistication of AI-enabled phishing, including focused awareness for higher risk groups, e.g. finance, IT and HR.

An increase in deepfake attacks on organisations is also being reported, driven by Gen-AI-powered tooling for which barriers such as cost, and complexity have reduced significantly in recent months. Such attacks will use publicly available video and audio footage to generate convincing fakes, which can be used to persuade employees to disclose information or make payments by giving the false impression these have been requested by individuals in authority. Technical defences against deepfakes are nascent and limited, so organisations should deploy a combination of strong awareness messaging and segregation of duties (maker-checker) controls, including encouraging employees to seek confirmation via a different channel before taking action.

Key Takeaways – Firms Should:

Embed effective employee awareness and accountability in respect of Gen-AI adoption:

- Determine your firm's position and philosophy - reflecting risk appetite, expected usage, desired benefits, current usage and compliance and regulatory requirements.

- Use this to outline a clear AUP, which clarifies accountability, emphasises transparency, recognises limitations, is clear on improper use and takes into consideration the needs of different roles.
- After consultation to engender support for successful adoption, communicate and implement the AUP, embedding it into broader training and awareness activity.
- Monitor and measure effectiveness, and review periodically to maintain in line with latest developments.
- Ensure that security education and awareness programmes address the risk of Gen-AI-powered social engineering attacks by educating employees on the risks of deepfakes and highly convincing phishing lures (see spotlight above).
- Review training needs across different user groups in relation to Gen-AI, covering the full spectrum from general user awareness of Gen-AI related attacks through to deep technical up-skilling for key SME roles.

Further Reading:

Preparing your organisation for AI use | CIPD June 2023

- *Preparing your organisation for AI use | CIPD*
- Guidance for creating and implementing an Acceptable Use Policy for AI

Framework of an Acceptable Use Policy for External Generative AI | FS-ISAC September 2023

- *FrameworkOfAnAcceptableUsePolicyForExternalGenerativeAI.pdf*
- Offering guidance on both permissive and stringent approaches, this adjustable policy helps organisations to arrive at the right balance in designing a Generative AI Acceptable Use Policy that aligns to their specific circumstances.

The Impact of AI in Financial Services: Opportunities, Risks and Policy Considerations | UK Finance / Oliver Wyman November 2023

- *The impact of AI in financial services.pdf*
- Provides helpful insights into financial sector firm experiences of AI implementation.