K2 Integrity Response to the Department of the Treasury's Request for Information on Uses, Opportunities, and Risks of Artificial Intelligence in the Financial Services Sector

August 12, 2024

K2 Integrity appreciates the opportunity to respond to this Request for Information regarding the uses, opportunities and risks presented by developments and applications of artificial intelligence (AI) within the financial sector. We would welcome further discussion on this matter, whether directly or as part of any industry-wide engagement.

K2 Integrity is a premier financial crimes, risk, and regulatory advisory firm, headquartered in the US and operating across various jurisdictions. With a focus on financial crimes risk management, investigations, monitoring, cybersecurity, and virtual asset advisory services, we bring together deep subject-matter expertise with proprietary technology offerings to help clients creatively solve for today while preparing for tomorrow.

K2 Integrity helps clients manage risk—whether that risk be investment, financial, regulatory, acquisition, new market entry, cyber, or reputational in nature—gathering intelligence to enhance critical decisions. We advise governments, companies, and high-net-worth individuals. Our clients frequently include financial institutions, law firms, hedge funds and private equity firms, and private and sovereign clients seeking to recover assets as well as public entities and private companies in a variety of sectors including energy, mining, real estate and construction, education, and technology.

Part A: General Use of AI in Financial Services

Question 1: Is the definition of AI used in this RFI appropriate for financial institutions? Should the definition be broader or narrower, given the uses of AI by financial institutions in different contexts? To the extent possible, please provide specific suggestions on the definitions of AI used in this RFI.

K2 Integrity Response:

Currently there is no single, globally accepted definition of artificial intelligence (AI), with variations existing within and across countries, supranational unions[1], associations[2], businesses, and individual

[1] For instance, the EU AI Act of 2024 uses the following definition for AI systems: "A machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments."

[2] For example, the Association of Southeast Asian Nations' 2024 guidance defines AI systems as: "...a machine-based system that is capable of influencing the environment by producing an output (predictions, recommendations or decisions) for a given set of objectives. It uses machine and/or human-based data and inputs to (i) perceive real and/or virtual environments; (ii) abstract these perceptions into models

academics and practitioners. Despite a lack of universal agreement, K2 Integrity identified common elements that often appear across several definitions of AI including the definition provided in the Department of the Treasury's request for information (RFI) and the definitions listed in the "International Definitions of Artificial Intelligence" published by *IAPP Research and Insights* (2023). The document compiles 51 definitions of AI articulated by legislation and legal instruments; guidance, standards, and voluntary frameworks; civil society and academia; and industry players across different countries[3]. The **common elements across AI definitions** identified by K2 Integrity are the following:

- **Machine-Based Systems:** The majority of definitions explicitly mention that AI involves systems that are operated by machines or computers, as opposed to human or biological intelligence.
- **Human-Defined Objectives:** Many definitions highlight that AI systems are designed to operate within the framework of goals or objectives set by humans. This emphasizes the role of human intention and control in AI development and deployment.
- **Data-Driven Nature:** The ability to process and learn from data is a core characteristic of AI systems mentioned in many definitions. This includes various techniques like machine learning, deep learning, and neural networks.
- **Decision-Making and Problem-Solving:** The capacity to make decisions, recommendations, or predictions, often in complex and uncertain situations, is a key feature highlighted in many definitions.
- **Autonomy and Adaptability:** The ability to operate with varying levels of autonomy and to adapt or learn from experience is another common theme in AI definitions.

In summary, AI can be described as a field of computer science focused on the design and development of machine-based systems using algorithms to perform tasks that—if performed by a human—would require intelligence, such as cognitive learning, problem-solving, and pattern recognition. According to existing definitions, for a system to demonstrate AI, that system should be capable not only of directly perceiving its operating environment and evaluating data collected within that environment, but also adapting to the data it receives (IAPP Research and Insights, 2023). These qualities—AI's perception of and ability to influence the observed environment—make a comprehensive definition of AI fundamental for ensuring the safe, responsible, and effective use of AI technology. Even if not universally accepted, a more robust definition would assist with:

through analysis in an automated manner (e.g., with machine learning), or manually; and (iii) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy."

[3] Definitions include the United States National Artificial Intelligence Initiative Act of 2020, the National Defense Authorization Act for Fiscal Year 2024, the "Blueprint for an AI Bill of Rights" issued by the White House Office of Science and Technology in 2022, the glossary of terms of the U.S. National Institute of Standards and Technology (NIST), the U.S. Patent and Trademark Office, the Organization for Economic Co-operation and Development (OCDE), the United Nations Educational, Scientific and Cultural Organization (UNESCO), etc.; industry players such as Amazon Web Services, Cisco, Google, IBM, Microsoft and Samsung. Definition from different countries including Brazil, Canada, U.S., E.U., U.K., United Arab Emirates, Japan, China, Colombia, India, U.N., Singapore, Spain, Australia, Germany, Norway, South Korea.

- **Creating clarity and consistency** across various sectors and disciplines in order to help mitigate the confusion and misinterpretation that can interfere with effective policymaking, regulatory compliance, and technological development.
- **Developing policies and regulation** to govern the use of AI by helping lawmakers, policy officials, and regulatory bodies draft laws and regulations that adequately address the full range of elements, complexities, and risks associated with AI technologies.
- **Identifying and addressing ethical considerations,** including privacy, bias, and accountability, that might be obscured in a more narrow or generalized definition.
- **Guiding the direction of research and development (R&D)** by providing a framework for innovation and facilitating alignment with recognized standards and objectives. This also helps governments and organizations with economic planning and resource allocation related to AI R&D, thereby ensuring investments are made in areas likely to yield the most positive economic and social impact, account for public expectations and concerns, and help promote the real-world application and broader acceptance of AI technologies.

Categorization of AI

When considering definitions of AI, it is important to note that AI itself can be sorted into three categories:

- **Artificial narrow intelligence (ANI)**—also known as "weak AI"—which is designed to perform a specific task or a limited range of tasks, such a language translation or facial recognition, and has been in use for many years.
- **Artificial general intelligence (AGI)**—also known as "strong AI"—which refers to machines possessing the ability to understand, learn, and apply knowledge across a wide range and variety of tasks at a level comparable to human intelligence.
- **Artificial super intelligence (ASI),** which goes well beyond AGI and represents a level of intelligence that surpasses human capabilities in all aspects.

It is important to account for AGI characteristics in defining the full scope of AI. Unlike ANI, which has established regulatory frameworks and applications, AGI's advanced capabilities and potential to have transformative impacts on society and security demand enhanced foresight and attention to ensure effective regulation that also aligns with human values and the well-being of society.

Considerations for a wider AI definition

Given these considerations, K2 Integrity recommends that Treasury consider widening the aperture of its definition of AI and proposes the following additions (noted in *italics*):

> A machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to perceive real and virtual environments;

abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action. *AI systems possess the following characteristics:*

- ***Adaptability and learning:*** *AI systems can learn from data and experiences, adapt to new circumstances, and improve their performance over time without being explicitly programmed for every specific task;*
- ***Cognitive capabilities:*** *AI systems can perform tasks that typically require human cognitive functions such as understanding natural language, recognizing patterns, solving problems, and making decisions;*
- ***Scalability and autonomy:*** *AI can operate at scale, handling large volumes of data and complex processes autonomously while also maintaining the ability to collaborate with human operators;*
- ***Robustness and reliability:*** *AI systems are designed to be robust and reliable, ensuring consistent performance and resilience in varying conditions and environments; and*
- ***Ethical and responsible design:*** *AI systems are developed with considerations for ethical implications, including privacy, fairness, transparency, and accountability, to ensure they are used responsibly and for the benefit of society.*

K2 Integrity also suggests that Treasury could consider other forms of Artificial Intelligence already in use within any amended definition or guidance. For example, Generative Artificial Intelligence (GenAi) is referenced throughout various sections of our response below. GenAi creates original content (including text, images, video, audio or software code) in response to a user's prompt or request. It leverages deep learning models by identifying patterns in huge amounts of data and using that to understand users' natural language requests to respond with relevant new content (Stryker, C. et al, 2024).

Part A: General Use of AI in Financial Services

Question 2: What types of AI models and tools are financial institutions using? To what extent and how do financial institutions expect to use AI in the provision of products and services, risk management, capital markets, internal operations, customer services, regulatory compliance, and marketing?

K2 Integrity Response: Financial institutions across the globe are adopting and implementing exponential technologies, including AI, in their operations but are doing so at varying levels of transformation, particularly when it comes to the development and management of AI. There are four broad levels of maturity when it comes to the use of AI in financial institutions:

- **Level 1: Desktop automation.** Less technologically mature financial institutions primarily rely on desktop automation strategies. This typically involves creating macros[4] to handle repetitive tasks, using readily available, basic technologies, and implementing toolkits without a cohesive digital platform.
- **Level 2: Robotic Process Automation (RPA).** RPA involves the use of strategic platforms for tactical change, broad application of AI and technology where use cases might not be function-specific, such as rule-based automation, and the automation of non-intelligent, binary decisions.
- **Level 3: Intelligent Automation (IA).** IA involves the development of technology that enhances the user experience, and supports activities such as process mining, functionality of chatbots, document digitization, natural language processing, and knowledge representation.
- **Level 4: AI-driven Decision-making.** A small number of financial institutions that have invested considerably in the use of exponential technologies for purposes of digital transformation can use AI for non-routine tasks that require judgment, cognitive capabilities, dynamic rules, and artificial machine learning. Financial institutions at this level of digitalization use AI to increase value in terms of revenue and customer experience, rather than to improve efficiency and reduce costs.

Below we provide some select examples of established AI application that financial institutions have been using for a long time across various banking and non-banking activities, products, and services. Please see Table 1 below.

In contrast, Table 2 highlights select examples of emerging AI use cases for financial institutions. While some Financial Institutions have already adopted these technologies, their use remains limited.

Table 1. Examples of Established Use Cases of the Application of AI in Financial Institutions

Support Answer to Question 2		
No.	**Activities, Products & Services *(Select)***	**Select examples of Current Use Cases of the Application of AI in Financial Institutions:**
1	**Fraud prevention**	• AI-enabled fraud prevention systems predict the probability of card transactions, payments and customer applications involving fraud. • A leading financial services trade association in the UK noted in its Annual Fraud Report that there is significant cross-sectoral effort to use Generative Artificial Intelligence ("GenAi") to share data and intelligence across financial services firms, telecoms, tech companies mode, and regulatory bodies to mitigate "live-scam" and large scale social engineering attacks, leading to 3,700 unauthorized sender IDs being blocked to prevent them being used to send scam text messages mimicking trusted organizations (UK Finance, 2024)

[4] In this context, a **macro** refers to a sequence of pre-recorded actions or commands that can be automatically executed to perform repetitive tasks on a desktop computer.

2	**Automated customer service**	AI automates customer identification, determines customer needs, and creates automated responses to customers.
3	**Cybersecurity**	• Improves system resilience (the capacity to respond to and recover from a cyber-attack) via threat and anomaly detection. • The Bank for International Settlements has highlighted a significant trend among central banks, whereby c.33% have adopted GenAi to strengthen cybersecurity measures (Fatima, 2024).
4	**Asset management**	AI-enabled robo-advisors assess investors' risk tolerance and investment aims.
5	**Loans**	Machine learning models are using pattern analysis and predictions to forecast delinquency and the impact on loan impairment charges.
6	**Credit scoring**	AI is being employed to assess credit risk in client applications and mitigate potential losses due to delinquency. While scoring models have long been used in this context, the growing complexity and sophistication of AI-driven models will make the implementation of clear explanation mechanisms crucial and mandatory in the near future.
7	**Anti-money laundering (AML), countering the financing of terrorism (CFT), and combating sanctions evasion**	• K2 Integrity observe: a) interest from both domestic & international FIs in using 3rd-party generative AI systems or models to comply with BSA/AML obligations; and b) firms deploying AI tools in transaction monitoring & suspicious activity report (SAR) programs to augment existing rules-based monitoring systems – all to strengthen effectiveness & efficiency of internal controls. • Furthermore, a U.S. branch of an international bank has followed its enterprise-wide adoption of a third-party AI system to supplement its existing customer risk rating methodologies with the use of a dedicated customer risk scoring module from the same third-party AI system vendor. In this instance, the AI tool is used to calculate risk scores for certain customers that the existing methodology cannot provide a reasonable risk rating
8	**Global markets**	Algorithmic trading is commonly used to execute trades at high speed and scale. These algorithms can analyze market data and make trading decisions based on pre-defined rules, often outperforming human traders.
9	**Cash management**	AI is being used to adjust to seasonal and company-specific operational activities with the use of a machine learning (ML) model to select project cash flow in accounts.
10	**Customer experience**	Anomaly detection tools highlight overpriced spending and analyze spending patterns and credit score changes. AI also helps with payment reminders and smart analytics to help clients manage their monthly budgets.
11	**Audit and assurance**	Machine learning is evaluating and "scoring" the effectiveness of controls.
12	**Customer relationship management (CRM)**	Systems store vast amounts of customer data, such as demographics, transaction history, and product usage. This data is used to segment customers into groups with similar needs and preferences and facilitate cross-selling of products and services.

Table 2. Examples of Emerging Use Cases of the Application of AI in Financial Institutions

Support Answer to Question 2		
No.	**Activity**	**Select examples of Future Use Cases of the Application of AI in Financial Institutions:**
1	**Financial crime compliance**	• **Emerging trends discovery:** The identification of financial crime emerging trends will be facilitated by the use of unstructured data coming from the news, social media, dark web forums, academic publications, regulatory filings, legal documents, and other publicly available sources. • **Gan analysis of policies:** By using large language models (LLMs), financial institutions can review their compliance policies to identify gaps by comparing them against the regulations. • **Perpetual KyC:** Continuously monitoring customer data and transactions in real-time, using machine learning algorithms to identify suspicious patterns or changes in risk profile, and triggering alerts for further investigation if needed. • **AML/compliance investigations:** Digitization can streamline AML or compliance investigations by automating data collection from diverse sources, analyzing transactions for suspicious patterns using AI, and generating comprehensive reports for compliance teams with a lower error rate than humans. • **Transaction Monitoring / Suspicious Activity Report (SAR) programs:** More robust models will continue the reduction of false positive alerts, leading to more efficient alert reviews and better SAR reporting. Improved AI's ability to learn from historical data will continue to streamline the allocation of resources. • **Collective intelligence (federated learning):** There are innovations like the one offered by Consilient, which makes it possible to leverage the knowledge from multiple institutions by training AI models from diverse data sources while not compromising data security as the data never leaves individual sites, it's just the models that are shared.[5] • **Financial Crime Knowledge Navigator:** AskFIN aids financial institutions in combating financial crime by providing quick answers to questions related to AML/CFT, sanctions, and more. It leverages AI and a vast library of financial integrity resources to enhance the efficiency of risk and compliance teams. Its multilingual capabilities and broad range of topics make it valuable for global institutions.[5]
2	**Hypertargeting**	AI tools can analyze unstructured data—such as emails, pictures, voice notes, and social media posts—to extract valuable insights that can be used for customer sentiment analysis and market research.
3	**Explainable credit decisions**	AI can assist with a growing need for transparency and explainability in how these decisions are made, especially in cases of loan denials or adverse actions. This area is still under development to ensure fairness and compliance with regulations.
6	**Improved stress testing**	Creating realistic, dynamic scenarios that capture various economic and geopolitical factors remains a challenge for the Credit Risk departments. While frameworks like Basel III offer guidelines to calculate ratios for capital adequacy, their credibility is questioned during times of crisis. AI could gather enough information to develop stress tests and

[5] Please refer to K2 Integrity's answer to question 13 *'How do financial institutions, technology companies, or third-party service providers expect to use AI to address and mitigate illicit finance risks?'*

		scenarios that are more representative of real-world events and interdependencies, thus providing a more accurate assessment of a bank's resilience in adverse conditions.
7	**Real-time fraud prevention across multiple channels**	While AI is effective in detecting fraud within specific channels like online transactions, integrating real-time fraud prevention across various channels such as mobile banking, ATMs, and in-person interactions is a complex problem that could be solved through AI.
9	**Faster loan underwriting**	Automates the end-to-end credit assessment process, improving speed and accuracy in loan approvals. Banks are striving to make near-real-time decisions on customer applications for banking products such as loans, credit cards, and accounts. While customer onboarding processes can be automated with digital customer journeys, the final decision on an application is often delayed due to necessary fraud and credit risk-related checks. To overcome this challenge, financial institutions have are developing improved employed AI tools to automate the decision-making processes, enabling straight-through processing of applications. This use of AI significantly enhances customer experience by reducing waiting times and providing faster access to banking products, while still maintaining rigorous standards for fraud prevention and credit risk assessment.

Part A: General Use of AI in Financial Services

Question 3: To what extent does the type of AI, the development of AI, or AI applied use cases differ within a financial institution? Please describe the various types of AI and their applied use cases within a financial institution. Are there additional use cases for which financial institutions are applying AI or for which financial institutions are exploring the use of AI? Are there any related reputation risk concerns about using AI? If so, please provide specific examples.

K2 Integrity Response: Financial institutions leverage various types of AI — including artificial narrow intelligence (ANI), artificial general intelligence (AGI), and machine learning (ML) — in order to strengthen operations, enhance the customer experience, and more effectively identify, assess, and manage risks.

Often described as "weak AI," ANI is designed to perform a specific task or a limited range of tasks, as noted in the response to Question 1. Within financial institutions, uses cases across such narrow tasks include fraud detection, customer service chatbots, credit scoring, and transaction monitoring.

AGI, on the other hand, is known as "strong AI" and is imbued with the ability to understand, learn, and apply knowledge across a wide variety of tasks at a level comparable to human intelligence. The financial services sector is closely monitoring advancements in AGI with an eye to leveraging its potential in future applications.

Further responses outlined below refer to developments and challenges within the financial services industry in light of GenAi.

Machine Learning (ML) is a subset of AI that involves training algorithms to make predictions or decisions on large datasets without being programmed explicitly for each task. ML use cases within financial institutions include algorithmic trading where ML algorithms are used to analyze market data and execute trades at optimal times; risk management, where ML predicts potential risks by analyzing

historical data and identifying trends; and the delivery of personalized financial advice where ML can identify and offer tailored financial products and advice based on individual customer data.

AI Sources of Reputational Risks. Despite the benefits that come from applying AI to strengthen operations, enhance the customer experience, and more effectively identify, assess, and manage risks, financial institutions that utilize AI in their operations confront a range of risks that can adversely affect their reputations with both customers and regulators (Robins-Early, 2024). Informed, in part, by a recent open letter from AI industry experts titled "A Right to Warn about Advanced Artificial Intelligence" (Hilton J. et al., 2024), K2 Integrity has identified the following AI sources of reputational risk facing the financial services sector. We also provide high-level examples that, while not necessarily based on cutting-edge technology, help to illustrate reputational consequences that can accompany the misuse of AI:

- **Bias and discrimination.** AI algorithms can perpetuate existing biases present in training data, leading to discriminatory outcomes in lending, credit scoring, and other financial services. Bias, discrimination, and fairness stand as critical challenges in AI governance due to their significant potential impacts on individuals and communities. These challenges can result in discriminatory outcomes and worsen existing inequalities on a large scale. AI governance should therefore consider legal and ethical standards, including human rights, professional responsibility, human-centered design and technology control, community development, and non-discrimination. While AI's automation of human tasks brings benefits like scalability, efficiency, and accuracy, it also presents the issue of algorithmic bias. This bias manifests as systematic errors where algorithms consistently overlook certain groups more than others. Credit decisioning in financial institutions is an example where algorithmic bias can potentially be prevalent. AI systems can also inadvertently perpetuate biases present in the training data, leading to unfair treatment of certain groups. ***Examples***: Apple Card's algorithm was accused of gender bias in credit limit assignments (Duffy, 2019); and Facebook's AI-powered advertising system was criticized for enabling discriminatory targeting practices (Angwin & Parris Jr., 2016).
- **Data security and privacy.** Financial institutions hold vast quantities of sensitive personal and financial data, making them attractive targets for cyberattacks. AI systems can be vulnerable to these attacks, potentially compromising customer information. Extensive data collection and analysis for AI applications can pose significant privacy risks. AI systems often require vast datasets, which include sensitive personal information, raising concerns about data usage and protection among customers, especially considering frequent data breaches and misuse. Ensuring compliance with privacy regulations is therefore crucial to prevent unauthorized access and misuse of personal data. Privacy laws provide an ethical framework for the use of new technologies, emphasizing the importance of transparency in data collection, processing, and usage. This helps establish trust between customers and institutions and prevents discriminatory practices that could exacerbate existing inequalities. As AI technology continues to advance, safeguarding privacy will be essential to maintaining public confidence and protecting individual

rights. This topic is explored further under our response to question 11. ***Example***: A notable example of a data security and privacy issue impacting financial institutions is the 2017 Equifax data breach. Hackers exploited a vulnerability in Equifax's web application software, gaining access to the personal and financial data of 147 million consumers (Federal Trade Commission, 2019).

- **Lack of transparency and explainability**. Complex AI models used in financial decision-making can be difficult to interpret, raising concerns about fairness and accountability. It often can be challenging to understand the rationale behind AI-driven decisions, especially when they have significant financial consequences. Many AI models, particularly deep learning models, function as "black boxes" that make their decision-making processes difficult to explain. This lack of transparency can undermine trust as customers and regulators struggle to understand how these decisions are made. A significant reason for the mistrust in AI systems is that users, and often even the creators, lack a clear understanding of their inner workings. This problem arises because AI models are either too complex for human comprehension or their details are protected by intellectual property rights. The criticism and concerns stem from the fact that people are usually only informed of the final decisions made by AI such as loan approvals or product pricing but without any insight into how or why these decisions were made. This issue has garnered increasing public interest, as AI systems are making decisions that directly affect human well-being. As a result, financial institutions often make trade-offs to simplify certain system aspects at the expense of efficiency and customer experience. Further comments on explainability are included within question 7. ***Example***: Zillow's AI-powered home valuation tool, Zestimate, faced backlash for inaccurate valuations and potential market manipulation (Harney K., 2017).

- **Exploitation of AI for financial crime**. Malicious actors can exploit AI systems and capabilities both to commit financial crimes—including money laundering, market manipulation, and fraud—and develop sophisticated schemes that help them evade detection. ***Examples:*** Robo-advisors have been scrutinized for their lack of transparency and structural conflicts of interests, difficulties presented by the supervision of algorithms, and potential threats to the stability of the financial system (Maume P., 2021). Deepfake technology has also been used in fraud schemes against financial institutions. In 2023, a UK energy company was tricked into transferring $243,000 to a Hungarian supplier after a fraudster used deepfake audio technology to impersonate the chief executive's voice, instructing an employee to authorize the payment (Damiani J., 2023).

- **Damage from "AI washing."** Sometimes reputational risks related to the use of AI in financial institutions stem not from the actual use of AI but rather from in accurate, misleading, or entirely false claims that an institution is employing AI when it really is using less sophisticated digital technical or technologies—if at all. This has the potential to erode public trust and confidence. ***Example***: In March 2024, the U.S. Securities and Exchange Commission (SEC) fined two investment firms for the practice of "AI washing". In its statement, the SEC advised that the firms

made “false and misleading statements” about using AI and machine learning in their service offerings. While neither firm admitted nor denied the SEC’s findings, the firms agreed to pay civil penalties totaling $400,000 (SEC, 2024).

- **Job displacement**. The automation of financial tasks through AI and the associated operational efficiencies and cost savings may contribute to job losses in the financial services industry, raising concerns about unemployment and social impact regarding the use of AI. ***Example***: Digitization led to c.3,000 branches in the United States in 2020, due to the need for increased online services during the Covid-19 pandemic. (Maio, 2022).

- **Security vulnerabilities**: AI systems are vulnerable to cyber-attacks, which can result in severe financial and reputational consequences. When AI security is compromised, it can lead to the manipulation of outputs[6], theft of sensitive data, or disruption of system operations. Such breaches not only can cause financial harm and tarnish reputations, but they also can potentially lead to physical dangers[7]. While AI security shares similarities with traditional cybersecurity, there are distinct differences. Cybersecurity generally focuses on protecting computer systems and networks from attacks, whereas AI security involves safeguarding the AI system's components—data, models, and outputs. Malicious actors can exploit the unique vulnerabilities of AI algorithms to conduct adversarial attacks, taking advantage of inherent limitations in these systems. ***Example***: “Tesla’s keyless entry system in its latest Model 3 remains vulnerable to relay attacks despite its upgrade to ultra-wideband (UWB) radio which had been touted as a solution to relay attacks. A relay attack tricks a car into unlocking by relaying signals from an owner’s key fob or smartphone, often from a distance. This technique has been used to steal numerous car models for years as it tricks cars entry systems to respond as if the real owner was nearby.”(Alan J, 2024)

- **Systemic risk**. The widespread use of interconnected AI systems in the financial sector can create systemic risks, where failures in one system could trigger a cascade effect across the entire industry. In a potential scenario, the adoption of highly advanced AGI systems by financial institutions for trading and risk management could initially lead to significant profits due to their superior analytical and decision-making capabilities. As these interconnected AGI systems continuously learn and adapt from each other and market conditions, however, their behavior becomes increasingly complex and unpredictable. This could result in unforeseen market volatility, potentially culminating in a flash crash triggered by a minor error or miscalculation. The interconnected nature of financial markets amplifies the impact, leading to a cascade of margin calls, forced liquidations, and a loss of confidence in the financial system, potentially triggering a global financial crisis. ***Example***: A recent IT glitch triggered a widespread technological disruption,

[6] In this context, **outputs** refer to the results or predictions generated by the AI system

[7] Here **physical dangers** refer to the potential for AI security breaches to cause real-world harm or damage, such as compromised critical infrastructure, malfunctioning autonomous systems, or weaponized AI leading to accidents, injuries, or even loss of life

highlighting the vulnerability of interconnected global networks. The July 19th incident caused major outages across various sectors, impacting flights, healthcare, banking, and more. Millions were affected across numerous countries, resulting in an estimated total cost exceeding $5 billion. (Schneider, 2024)

Part A: General Use of AI in Financial Services

Question 4: Are there challenges or barriers to access for small financial institutions seeking to use AI? If so, why are these barriers present? Do these barriers introduce risks for small financial institutions? If so, how do financial institutions expect to mitigate those risks?

K2 Integrity Response: In July 2024, AI thought leaders from K2 Integrity participated in the 23rd Pan-American Risk Congress, held in Cartagena, Colombia, to speak about the applications of AI in financial crime compliance. During the interactive workshop delivered by K2 Integrity, 88 participants from various banking institutions provided insights (via a survey) regarding the specific challenges faced by small financial institutions (SFIs) when implementing AI. This included challenges that adversely impact their risk management operations and ultimately their competitiveness in the financial services sector. Survey responses highlighted the following challenges and barriers:

- **High costs:** Significant financial investments are needed for AI infrastructure, software, talent/expertise acquisition, and training, and limited budgets further hinder the navigation of complex regulations related to data privacy, security, and fairness when implementing AI systems.
- **Data limitations**: SFIs often have smaller and less diverse datasets compared to larger institutions, constraining the overall effectiveness of AI applications.
- **Outdated legacy systems**: Integrating AI into outdated legacy systems is not only time-consuming, it also introduces additional costs to an already expensive and costly endeavor.

These challenges and barriers present several risks for SFIs:

- **Competitive disadvantages**: SFIs may fall behind larger institutions that leverage AI for efficiency, personalization, and risk management and will lag in their ability to enhance customer experiences, streamline operations, and develop innovative products enabled by AI.
- **Increased operating costs:** Without the automation benefits that accompany AI technologies, SFIs will continue to rely on manual processes and inefficient workflows that often lead to higher operational costs.
- **Expose to fraud:** Without advanced AI-driven fraud detection tools, SFIs may face higher fraud risk exposure.

- **Constrained credit risk decisioning:** Similarly, credit risk decision making may remain sub-optimal without AI technologies, contributing to poor underwriting practices and higher credit risks.

Part B: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

Sub-Focus Area: Actual and Potential Opportunities and Benefits

Question 5: What are the actual and expected benefits from the use of AI to any of the following stakeholders: financial institutions, financial regulators, consumers, researchers, advocacy groups, or others? Please describe specific benefits with supporting data and examples.

How has AI been used in financial services to improve fair lending and consumer protection, including substantiating information? To what extent does AI improve the ability of financial institutions to comply with fair lending or other consumer protection laws and regulations? Please be as specific as possible, including details about cost savings, increased customer reach, expanded access to financial services, time horizon of savings, or other benefits after deploying AI.

K2 Integrity Response: In its responses to Questions 2 and 3, K2 Integrity has outlined various use cases and attendant benefits for the application of AI in financial institutions. This section also considered some consumer benefits at a high-level, including increasing the speed and transparency of the lending process. However, we are unable to materially comment in terms of cost savings, increased customer reach, expanded access to financial services and time horizon of savings.

These use cases can, and have the potential to, help strengthen and streamline business operations, enhance customer experiences, and effectively identify, assess, and manage risks for financial institutions. Interest in these actual and expected benefits of AI is borne out by a recent survey. Stakeholder responses to the Annual Financial Institution's Financial Sentiment Survey (FISS) from Lloyd's Bank for 2024 — based on inputs from more than 100 senior decision-makers across banks, insurer, financial sponsors, and asset and wealth managers — revealed that 63% of United Kingdom (UK) financial institutions are currently investing in AI solutions, almost doubling 2023's results of 32%. Moreover, the benefits of AI are tangible: 32% of respondents reported enhanced productivity, while 21% cited a competitive edge as a key benefit. Enhanced customer relations were also evident, with 17% of respondents reporting that AI provided greater insights on customers and 13% stating that AI provided a better overall client experience. UK stakeholders are eager to capitalize on the capabilities and benefits associated with the application of AI technologies, with nearly half (46%) of financial institutions having dedicated AI teams in place to explore AI use cases; 39% considering partnerships with AI firms; and 15% already having partnerships in place. Total volume of respondents is unknown (Lloyds, 2024).

Part B: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

Sub-Focus Area: Actual and Potential Risks and Risk Management: Oversight of AI – Explainability and Bias

Question 6: To what extent are the AI models and tools used by financial institutions developed in-house, by third-parties, or based on open-source code? What are the benefits and risks of using AI models and tools developed in-house, by third-parties, or based on open-source code? To what extent are a particular financial institution's AI models and tools connected to other financial institutions' models and tools? What are the benefits and risks to financial institutions and consumers when the AI models and tools are interconnected among financial institutions?

K2 Integrity Response: In focusing on the latter part of this question — the benefits and risks to financial institutions and consumers when AI models and tools are interconnected among financial institutions — K2 Integrity observes the following benefits:

- **Enhanced decision-making.** Financial institutions theoretically can leverage interconnected AI models to make more informed decisions. By sharing data and insights, they collectively can improve risk assessment, fraud detection, and investment strategies. Additionally, when AI interconnected is correctly deployed, consumer decision-making also benefits from streamlined processes, personalized financial advice, and tailored services.
- **Risk mitigation:** Cross-institutional collaboration leads to shared data that can identify systemic risks, market trends, and potential threats.
- **Efficiency and cost reduction.** Interconnected AI tools present an opportunity to streamline processes, automate routine tasks, and reduce operational costs for institutions. These benefits redound to consumers, who should experience faster transactions, lower fees, and improved customer service.

Despite these benefits, institutions must also contend with the risks that flow from interconnected AI models and tools. These risks include:

- **Data privacy and security breaches.** Sharing data across institutions increases the risk of data breaches, unauthorized access, and identity theft. Consumers may face privacy concerns if their financial information is shared with third parties providing services to a financial institution.
- **Bias and lack of fairness.** Interconnected AI models may inherit biases present in shared data that leads inadvertently to unfair practices, such discrimination based on race, gender, or socioeconomic status. Such biases in algorithms contribute to unfair treatment of consumers and reputational risks for the financial institutions involved.
- **Cascading systemic risks.** Dependency on interconnected AI systems creates systemic risks, with a failure in one institution's model cascading to others and consumers suffering from larger market-wide disruptions.

- **Lack of transparency.** Complex interconnected models can be opaque, obscuring the bases of decision-making for consumers, who may not know why certain financial decisions are made that directly affect their financial goals and stability.
- **Regulatory challenges.** Consumers need clear regulations that protect their interests, but regulating interconnected AI currently is fragmented across global markets, and regulators will be further hampered by challenges in harmonizing rules across institutions. For example, we refer to data privacy issues facing EU regulators under our response to section 11.

Overall, interconnectivity offers significant benefits to both institutions and consumers, but participating financial institutions have a responsibility to identify, assess, and manage the risks associated with cross-institutional, interconnected AI tools and models. Such risk management is crucial to ensuring a fair, secure, efficient, and resilient financial ecosystem for both institutions and consumers. Please also see K2 Integrity's response to Question 3, which includes more details on AI sources of reputational risks to financial institutions that complement the interconnected risks outlined above.

Part B: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

Sub-Focus Area: Actual and Potential Risks and Risk Management: Oversight of AI – Explainability and Bias

Question 7: How do financial institutions expect to apply risk management or other frameworks and guidance to the use of AI, and in particular, emerging AI technologies? Please describe the governance structure and risk management frameworks financial institutions expect to apply in connection with the development and deployment of AI. Please provide examples of policies and/or practices, to the extent applicable.

What types of testing methods are financial institutions utilizing in connection with the development and deployment of AI models and tools? Please describe the testing purpose and the specific testing methods utilized, to the extent applicable.

To what extent are financial institutions evaluating and addressing potential gaps in human capital to ensure that staff can effectively manage the development and validation practices of AI models and tools?

What challenges exist for addressing risks related to AI explainability? What methodologies are being deployed to enhance explainability and protect against potential bias risk?

K2 Integrity Response: K2 Integrity observes the following AI explainability challenges that typically arise from the complex nature of models and decision-making processes:

- **Black box problem: complexity and lack of transparency**. Many AI models, especially deep learning ones, operate as "black boxes" where the inputs and outputs may be known to some extent, but the internal workings remain opaque. This lack of insight and understanding

regarding how an AI system arrives at its decisions limits explainability and thus poses safety, ethical, and legal concerns. The overall complexity and lack of transparency make it increasingly difficult for financial institutions to fully understand and evaluate the associated risks, and the lack of transparency erodes public trust (Stewart, A. 2024).

- **Divergent objectives.** Stakeholders—including developers, users, and regulators—have varying needs for AI explanations. Engineering goals, however, are prioritized over these other stakeholder objectives, overshadowing other considerations and contributing to inadequate explainability.
- **Data privacy risks.** Providing real-time, high-quality explanations to end-users can often be at odds with the imperative to ensure data privacy. Moreover, third-party vendors often manage sensitive customer data, which expands the risk vectors for data breaches and privacy violations.

Managing the risks associated with AI are significant. For the AI system vendor, the need to accurately and fully document the theoretical design and practical implementation of an AI system — including all relevant components (third party, in-house, and open-source tools) — forms the basis of a comprehensive model risk management framework. As part of the vendor model governance framework, the vendor should ensure the statistical analyses used to develop, tune, and implement the AI model are adequately documented and made available to AI model end-users, such as financial institutions, to ensure model explainability and ongoing oversight.

For their part, financial institutions must continually evaluate and adapt their risk management practices in order to keep pace with the rapidly evolving nature of AI technologies and their associated risks. This includes customizing their third-party risk management practices to fit the specific context, taking into account factors such as a third-party's reliance on AI, the regulatory environment, and internal capabilities. Third-party risk management also involves continual monitoring of the third-party's performance and compliance through regular audits, reviews, and assessments, as well as developing and maintaining incident response plans specific to AI-related breaches or failures.

Ultimately, achieving meaningful AI explainability requires addressing these challenges while balancing technical, ethical, and practical considerations. Please also see K2 Integrity's response to Question 3, which includes more details on AI sources of reputational risks to financial institutions that complement some of the challenges outlined above.

Part B: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

Sub-Focus Area: Actual and Potential Risks and Risk Management: Oversight of AI – Explainability and Bias

Question 8: What types of input data are financial institutions using for development of AI models and tools, particularly models and tools relying on emerging AI technologies? Please describe the data governance structure financial institutions expect to apply in confirming the quality and integrity of data.

Are financial institutions using "non-traditional" forms of data? If so, what forms of "non-traditional" data are being used? Are financial institutions using alternative forms of data? If so, what forms of alternative data are being used?

K2 Integrity Response: K2 Integrity observes that financial institutions, and banks in particular, employ a wide range of input data in the development of AI models and tools, especially those relying on emerging AI technologies.

Traditional datasets used by banks for a general range of operations and activities include:

- **Customer data:** Demographic information (collected during Know Your Customer (KYC) collection); transaction activity; credit history; risk profiles; financial statements; consumer preferences such as card purchases; relationship with other customers; customer counterparty information, such as beneficiaries; and computer IP addresses.
- **Product data:** Interest rates and pricing information; costs by segment and country; expected revenue associated with banking products; target markets; and sales channels.
- **Institutional financial data:** Profit and loss reports; balance sheets; delinquency reports; income statements; cash flow statements; regulatory filings; and financial forecasts.
- **Market data:** Stock prices; trading volumes; interest rates; and macroeconomic indicators.
- **International risk indicators:** Financial Action Task Force (FATF) risk countries; Office of Foreign Assets Control (OFAC) sanctioned entities; and Basel III guidelines for capital adequacy include parameters that could be inputted into ML models.

More specifically, K2 Integrity observes that AI models used in BSA/AML internal control frameworks as part of a financial crime compliance program will draw from customer and transactional data (including those identified above), as well as from screening list data such as sanctions lists, country lists, political exposed persons (PEP) lists, and adverse media sources. Additionally, integral parts of a sound data governance framework will include clear and complete data architecture documentation and diagrams detailing the requisite data source systems; all relevant key data elements, data lineage or APIs between source systems and the AI model environment; and descriptions of any Extract Transform and Load (ETL) processes. As part of data governance, K2 Integrity recommends data quality and data lineage testing by AI model owners should be performed on a periodic basis — with clearly defined protocols for both root cause and impact analysis of identified data gaps — in order to ensure accuracy and completeness of requisite data.

In addition to traditional datasets currently being leveraged by banks, there are also non-traditional datasets — stemming, in part, from emerging AI potential and evolution in unstructured data analysis — that could be ingested or utilized by AI-powered tools in the future. While adoption of these data sources is growing, only few institutions are harnessing AI's power to process and extract insights at scale.

These **non-traditional datasets** include:

- **Local and international jurisdictional regulations** — such as Basel Accords (Basel III framework), the Financial Action Task Force (FATF) 40 recommendations on combating money laundering and terrorist financing, the International Organization of Securities Commissions (IOSCO) principles, Bank Secrecy Act (BSA), etc— could be analyzed as part of the AI-tools such as LLMs to analyze bulk text and produce a significative and targeted response based on a clear objective.
- **Institutional regulatory information** such as internal policies and governance-related documentation lend themselves to relatively easy analysis.
- **Emails, voice notes and other types of communications** from different channels.
- **Social media information**, mostly by humans during investigations processes, is currently being leveraged by AI-based digital workers. In the near future, posts, trends and pictures shared on social media could be used for more targeted marketing as part of the new capabilities of AI-tools to read and analyze unstructured data.
- **Information from the Internet of Things (IoT)** could be easily extracted from wearable devices or smart home devices.
- **Geospatial location data** from mobile devices rather than just IP addresses.

Part B: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

Focus Area: Actual and Potential Risks and Risk Management: Fair Lending, Data Privacy, Fraud, Illicit Finance, and Insurance

Question 11: How are financial institutions addressing any increase in data privacy risk related to the use of AI models, particularly emerging AI technologies? Please provide examples of how financial institutions have assessed data privacy risk in their use of AI.

In what ways could existing data privacy protections (such as those in the Gramm-Leach-Bliley Act (Pub. L. 106-102)) be strengthened for impacted entities, given the rapid development of emerging AI technologies, and what examples can you provide of the impact of AI usage on data privacy protections?

How have technology companies or third-party providers of AI assessed the categories of data used in AI models and tools within the context of data privacy protections?

K2 Integrity Response: In focusing on the ways existing data privacy protections (such as those in the in the Gramm-Leach-Bliley Act) can be strengthened for impacted entities, K2 Integrity offers the following background and insights:

In May 2018, the European Union (EU) enacted the General Data Protection Regulation (EU GDPR), and following its exit from the EU, the United Kingdom (UK) combined these rules with the 2018 Data Protection Act (referred to as the UK GDPR), which came into force in January 2021. Both sets of

protections have extraterritorial implications for U.S. firms operating in the UK or the EU. For instance, these rules apply to any third parties engaged by an organization — even if the activity in outsourced — due to various laws that note organizations remain accountable for regulatory compliance, regardless of the activity outsourced (this is also pertinent to section 16, in reference to the management of third-party risk). Both GDPR regimes differ from the Gramm-Leach-Billey Act ("GLBA") in the United States in that they apply not only to financial institutions, but they also apply to any organization that processes the data of UK and EU citizens (Gupta, V., 2024). As U.S. financial institutions operating in the UK and EU must comply with local requirements, it may be prudent for U.S. policymakers to consider strengthening data privacy laws in the United States to ensure a more consistent approach is applied.

In light of examples of AI's impact on data privacy protections, the company Meta announced it will not be launching its "multi-modal" AI models (which operate across multiple devices) in the EU yet due to various inquiries from EU-based regulatory bodies around EU GDPR compliance (Fried, 2024). The company also had to suspend GenAi operations in Brazil, following concerns raised by that country's National Data Protection Authority (ANPD). In July 2024, Meta released a new privacy policy that granted them access to users' personal data to train its GenAI systems, which is currently under discussion with the ANPD due to data privacy concerns (Lakshmanan, Ravie 2024).

Regarding technology companies assessing the category of data used in AI models and tools, K2 Integrity understands that there has been extensive engagement by U.S. technology firms with the Republic of Ireland's Data Protection Commission (DPC), which acts as the "leading EU regulator" due to many U.S. firms having their EU headquarters in the region. It is the DPC's view that AI creates a number of potential data privacy issues; regulators need to decide if AI firms can trawl the internet for public data to train AI models and on what legal basis this data can be used. AI firms also need to understand and recognize individuals' data rights, including the right to erase their data, within the EU. The DPC also highlighted that the risk of AI models providing incorrect personal data about individuals must also be addressed. The European Data Protection Board is also currently designing guidance on how AI should operate under the EU GDPR, as well as the new EU AI Act (please refer to section 19 for further information on the EU AI Act) (Humphries, C., 2024).

Part B: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

Focus Area: Actual and Potential Risks and Risk Management: Fair Lending, Data Privacy, Fraud, Illicit Finance, and Insurance

Question 12: How are financial institutions, technology companies, or third-party service providers addressing and mitigating potential fraud risks caused by AI technologies? What challenges do organizations face in countering these fraud risks? Given AI's ability to mimic biometrics (such as a photos/video of a customer or the customer's voice) what methods do financial institutions plan to use to protect against this type of fraud (e.g., multifactor authentication)?

K2 Integrity Response: In the UK, it has been acknowledged that the full extent of GenAi use by fraudsters is nearly impossible to determine as governments, law enforcement and the financial services industry are still coming to grips with the threat (UK Public Sector Fraud Authority, 2024).

The leading financial services trade association in the UK (UK Finance) reported that 76% of authorized push payment (APP) fraud in the UK originated from online sources in 2023, but the extent of GenAi's impact on these figures is unknown (UK Finance, 2024). The UK Government has identified that GenAi is impacting romance fraud due to specially trained chatbots. This has significant financial implications. For example, a UK citizen withdrew £350k from her pension fund after a romance fraudster had used the latest deepfake technology to trick her during video calls, even simulating a fake marriage proposal. In addition, ChatGPT has also created LoveGPT, which is meant to support users with online dating skills. In reality, however, fraudsters are using this AI-enabled capability to create multiple fake profiles on several dating services, simultaneously scraping data from interactions with the dating services' users, including their profile pictures and profile text (UK Public Sector Fraud Authority, 2024). UK Finance identified a 12% increase in romance fraud cases in 2023 but noted that this increase was not directly attributed to GenAi (UK Finance, 2024).

There is significant cross-sectoral effort in the UK to use GenAi to share data and intelligence across financial services, the telecommunications sector, technology companies, and regulatory bodies in order to mitigate "live-scam" and large-scale social engineering attacks, leading to 3,700 unauthorized sender IDs being blocked to prevent them being used to send scam text messages mimicking trusted organizations (UK Finance 2024).

Further GenAi developments in the financial services industry to counteract fraud include Visa's new Visa Account Attack Intelligence ("VAAI") scoring system in the United States, launched in May 2024. Each transaction will have a VAAI risk score in "real-time" to help firms prevent fraudulent Card-Not-Present transactions (Visa, 2024).

Other types of fraud that GenAi will likely impact include persistent account takeover attempts, impersonation scams, CEO fraud and pig butchering (a form of investment fraud) (, Experian, 2024). An example of an impersonation scam is demonstrated by a Japanese company, who lost $35m after deepfake technology was used to clone a company director's voice in 2020. A branch manager was duped into believing that an acquisition was to be made by the company via a fraudulent phone call, and subsequently transferred the funds (UK Public Sector Fraud Authority, 2024). An unnamed company was reported to be the victim of CEO fraud in 2024, resulting in a loss of $25m – fraudsters used GenAi to pose as the company's CEO and other senior officers within the firm (Robson, K., 2024).

Financial institutions confront challenges in identifying and combating these fraud risks. A lack of awareness by members of the public is one such challenge. In 2022, the FBI counted 21,832 instances of business email fraud with losses estimated at US$2.7 billion (Lalchand et al, 2024).

A further risk area for financial services firms is criminals using GenAi to create "deepfakes" to circumvent biometric data security measures, generally used for identification and verification purposes

(Muckleroy, J., 2024). Fraud GPT, a product mimicking the legitimate ChatGPT platform, is now available on the dark web and can deploy machine-learning algorithms to generate malicious content for cybercriminals, such as persuasive phishing emails, fraudulent websites and malware (Lawler, E., 2024). This product, and others like it, will undoubtedly accelerate existing levels of AI-facilitated fraud, increasing the hurdles financial institutions and other organizations face in identifying and combating these risks.

In terms of methods used by financial institutions to prevent fraud in general outside the US, the European Union introduced "Strong Customer Authentication" (SCA) via the second Payment Services Directive (PSD) in 2018, which applied from 14 September 2019 (European Central Bank, 2019). The UK adopted this regime ahead of exiting the European Union on 1 January 2020 and had until 2022 to adopt the associated regulatory changes enacted by several statutory instruments (Financial Conduct Authority, 2021). In practice, SCA involved multi-factor authentication methods, such as biometrics, passcodes sent to email addresses/mobile devices, as well as multiple stages or questions to confirm customer consent before payments were processed.

In addition, the UK's Payment Systems Regulator (PSR) introduced "Confirmation of Payee" (COP) for online and mobile payments in 2020 to the six largest banking groups, which was gradually extended to all "Clearing House Automated Payment System" (CHAPs) payments. COP was designed to reduce certain types of APP scams and accidentally misdirected payments, by checking the name of the payee's account against the other details provided by the payer (UK Payment Services Regulator, 2022). These requirements were expanded in 2022 to all Payment Service Providers (PSPs), with a final deadline of 31 October 2024, on a phased approach (depending on their role within the payment chain) (UK Payment Services Regulator, 2022). As COP is a form of Artificial Intelligence and machine learning provided by various suppliers, it provides a further layer of multi-factor authentication outside those listed above in the preceding paragraph. UK PSPs are incentivized to deploy sophisticated technologies in light of the PSR's upcoming changes to enhance fraud prevention measures – from 7 October 2024, PSPs will have to reimburse all in-scope customers who fall victim to APP fraud in most cases, capped at £415k. This may involve splitting the cost as the "sending" or "receiving" PSP (UK Payment Services Regulator, 2023).

K2 Integrity understands that neither Strong Customer Authentication, nor reimbursement of fraudulent authorized payments, are mandatory in the United States at a federal level as yet – however, legislation was introduced to the House and Senate for mandatory reimbursement by the Democrats in early August 2024 (American Bankers Journal, 2024). In addition, SCA applies to US PSPs operating in the UK or EU – therefore it may be prudent for US payment regulators (such as the Federal Reserve Board) to consider a similar regime to increase consumer protection within the US.

<u>Part B</u>: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

<u>Focus Area</u>: Actual and Potential Risks and Risk Management: Fair Lending, Data Privacy, Fraud, Illicit Finance, and Insurance

Question 13: How do financial institutions, technology companies, or third-party service providers expect to use AI to address and mitigate illicit finance risks? How do financial institutions use AI to comply with applicable AML/CFT requirements?

K2 Integrity Response:

a) Federated Learning

One of the most promising AI approaches that financial institutions and technology companies expect to leverage in order to mitigate illicit financing risks is Federated Learning. Federated Learning is emerging as a powerful tool for financial institutions to combine efforts in identifying and combating illicit financial activity while simultaneously addressing data privacy concerns. "Federated learning is a way to develop and validate AI models from diverse data sources while mitigating the risk of compromising data security or privacy, as the data never leaves individual sites." (Rieke, 2019)

How does Federated Learning work? Federated machine learning can be used to develop AI models using the six following steps (McMahan, 2017):

1) **Initial model development:** An organization—the central hub which could be a central bank or a private institution— develops a learning algorithm that is designed to identify activities and patterns that point to potential illicit financial activity. This algorithm is trained on an initial dataset in order to create a preliminary model that will detect trends and anomalies.
2) **Model shared to nodes:** That preliminary model or learning algorithm is then shared with institutions—such as banks and other financial institutions, law enforcement entities, and regulatory bodies—that are participating as nodes for the hub.
3) **Model training in nodes:** Each participating nodal institution will then train a copy of the model on their own institutional transaction data. Each model copy is re-trained across these participants, reflecting new parameters and weights based on the training data available at each participating node.
4) **Re-trained models returned to hub:** Each participating nodal institution then transmits back to the hub either a version of the retrained model or detailed information on the updated parameters and weights. It does this without sharing any of their data, thereby preserving data privacy. This transmission back to the bub could occur periodically or on a set schedule independent of other participants.
5) **Aggregation by the hub:** Upon receiving the transmission from the participating nodes, the hub server aggregates and analyzes the revised model parameters and updates the central model based on this new information.
6) **Updated model shared back to nodes:** The hub then shares the revised model back to participating nodal institutions. This model now reflects insights derived from analysis across all

the participants' data. Alternatively, the hub could instead share revised weights and parameters for each participant to use in their own individual risk identification models.

Federated Learning is occurring in action, offering pathways now for financial institutions to more effectively and efficiently managing illicit financing risks. Consilient, for example, is a company dedicated to establishing a next-generation system for anti-money laundering and countering the financing of terrorism (AML/CFT). It has launched a secure, federated learning AI platform that aims to prevent financial crime and enable secure collaboration between and among financial institutions while simultaneously helping protect and advance privacy and data security (Intel, 2020). By sharing industry insights, Consilient enables institutions to leverage the most up-to-date and optimal models for AML/CFT and specific financial crime risks. Institutions bring these models in from their secure platform, deploying locally, in accordance with all model risk management approaches (Consilient, 2022).

Federated Learning models provide tangible benefits to participating financial institutions and their stakeholders. These benefits include:

- **Optimizing collective security.** Using a form of collective security, the combined efforts of multiple institutions strengthen their ability to combat illicit financial activity across various stakeholders in the financial services sector. By training on a shared machine learning model, these stakeholders can more effectively and efficiently identify emerging trends in illicit finance by helping institutions identify complex financial crime schemes that span multiple institutions and jurisdictions. By analyzing transaction patterns across a committed network of banks and institutions, the model can detect suspicious activity that might go unnoticed in isolation.
- **Potential reduction in false positives and costs.** By learning from a broader range of legitimate transactions across multiple institutions, the model can more accurately distinguish between normal and suspicious activity. This helps institutions reduce the number of false positive alerts, saving time and resources spent on unnecessary investigations. The collective knowledge pooled from diverse datasets helps the model refine its understanding of what constitutes truly unusual behavior, leading to more precise identification of actual financial crime threats (Shiffman, 2023).
- **Protecting and advancing data privacy and security**. Federated learning models uphold strict regulatory requirements for data privacy and security by ensuring that sensitive financial information never leaves the premises of the participating institutions. This decentralized approach eliminates the need to share raw data, significantly reducing the risk of data breaches or unauthorized access. The model learns from aggregated insights rather than individual transactions, preserving the confidentiality of customer data while still enabling effective collaboration against financial crime.

b) AI-powered Financial Crime Knowledge Navigators

There are AI tools empowering financial institutions to significantly enhance their effectiveness and efficiency in safeguarding against illicit actors. These are AI-powered navigators that provide financial institutions with quick, clear, authoritative answers to questions related to AML/CFT, sanctions, and

financial crimes, and otherwise assists their risk and compliance teams with day-to-day tasks. AskFIN is a good example. Developed by the Institute for Financial Integrity (IFI), AskFIN integrates cutting-edge technology with IFI's proprietary eLearning platform DOLFIN®—the Dedicated Online Financial Integrity Network—which includes the world's largest and most credible online library of financial integrity resources curated and maintained by certified subject matter experts. The DOLFIN library integrates relevant laws, regulations, and guidance from official standard-setting and regulatory bodies, hundreds of training modules with knowledge checks, and an extensive industry glossary, across a broad range of topics including:

- Basic and advanced AML/CFT issues, including risks associated with higher risk customers, products, and services including correspondent banking and trade finance
- Global sanctions issues, including specific programs imposed by the UN, US, EU, UK, and other jurisdictions, and sanctions evasion threats and typologies
- Proliferation finance and export controls, including risks, typologies, and case studies
- Anti-fraud and anti-bribery and corruption (ABC) standards and typologies
- Key lessons drawn from enforcement actions associated with AML, sanctions, and ABC violations
- Resources designed for public sector authorities including supervisors, financial intelligence units, investigators, and prosecutors
- AskFIN was built with multilingual capabilities, meaning users can ask questions and receive answers in any language, which is useful for financial institutions with a global presence.

Part B: Actual and Potential Opportunities and Risks Related to Use of AI in Financial Services

Focus Area: Actual and Potential Risks and Risk Management: Third-Party Risks

Question 15: To the extent financial institutions are relying on third parties to develop, deploy, or test the use of AI, and in particular, emerging AI technologies, how do financial institutions expect to manage third-party risks? How are financial institutions applying third-party risk management frameworks to the use of AI? What challenges exist to mitigating third-party risks related to AI, and in particular, emerging AI technologies, for financial institutions?

K2 Integrity Response: From a model governance perspective, the lack of transparency and explainability often seen in third-party AI models used for BSA/AML programs can limit a financial institution's understanding of the development and operation of the AI model in question. As noted in our response to Question 7, the complexity of AI technologies and their inherent opacity complicate risk assessment and ongoing monitoring, making it challenging for institutions to identify, assess, and manage associated risks effectively. These challenges are compounded by the need for institutions to adapt their risk management practices in order to keep pace with evolving AI technologies and ensure the secure handling of sensitive data. As a result, the ability for a financial institution to satisfy its own model risk

management requirements and comply with relevant regulatory expectations will likely be compromised.

In order to address these challenges, financial institutions should incorporate prescriptive oversight requirements — including periodic performance metrics, model explainability standards, model tuning and testing, and statistical analyses — into third-party vendor agreements. Additionally, continuous monitoring of the third-party's performance and compliance through regular audits, reviews, and assessments is also critical. Finally, financial institutions should develop and maintain incident response plans that enable the organization to respond to, and recover from, AI-related breaches or failures that do occur.

By tailoring risk management practices to their specific context, institutions can ensure better adaptability to emerging AI technologies and mitigating third-party risks. By prioritizing the handling of sensitive data, the need for robust security measures, and prescriptive oversight requirements into third-party vendor agreements, these tailored controls will enhance visibility into AI model validity and performance, thereby mitigating third-party risks and ensuring compliance with regulatory standards and best practices.

Part C: Further Actions

Question 18: What actions are necessary to promote responsible innovation and competition with respect to the use of AI in financial services? What actions do you recommend Treasury take, and what actions do you recommend others take? What, if any, further actions are needed to protect impacted entities, including consumers, from potential risks and harms?

Please provide specific feedback on legislative, regulatory, or supervisory enhancements related to the use of AI that would promote a financial system that delivers inclusive and equitable access to financial services that meet the needs of consumers and businesses, while maintaining stability and integrity, protecting critical financial sector infrastructure, and combating illicit finance and national security threats. What enhancements, if any, do you recommend be made to existing governance structures, oversight requirements, or risk management practices as they relate to the use of AI, and in particular, emerging AI technologies?

K2 Integrity Response: Our response focuses on the actions necessary to promote responsible innovation and competition with respect to the use of AI in financial services.

In the UK, several regulatory bodies and AI providers contributed to the Government's consultation response in light of a "pro-innovation approach to AI regulation," dated February 2024. A pilot scheme has been established as part of the UK's AI and Digital Hub, which is a new advisory service led by several regulatory bodies (including the Competition and Markets Authority (CMA) and the Information Commissioner's Office (ICO) (UK Department of Science, Innovation and Technology, 2024).

In April 2024, the CMA published their strategic update in the context of AI, articulating the following three key risks:

- **Distortion of market outcomes and competitiveness**: AI systems that make recommendations or offer choices to consumers could affect or distort market outcomes and competitiveness, where incorrect options are given "undue prominence";
- **Inadvertent increases in prices**: AI systems and algorithms used to set prices could inadvertently result in higher prices, or the facilitation of collusion between firms;
- **Exclusion of new entrants:** Personalized offers, or selective targeting of consumers who are likely to switch to other providers of the same products, could lead to new entrants being excluded from the market.

The CMA also recognized the risks of "Foundation Models" ("FMs"), which are used in AI development, defined as the "pre-training of large, general models." As several large technology providers are extremely active in this field, this raises the risk of reducing choice and quality for consumers and preventing diversity and choice (UK Competition and Markets Authority, 2024).

K2 Integrity observes that the notes accompanying this RFI outline the Department of the Treasury's impact assessment of new non-bank entrants on competition in consumer finance markets in the context of AI in November 2022 (US Treasury, 2024). Competition in this industry is also being explored by other U.S. agencies, including the Department of Justice (US Department of Justice, 2024). Accordingly, it would be helpful if US cross-agency collaboration produces consolidated guidance for firms to ensure that any regulations or industry best practice is implemented effectively.

The Treasury Department and other US stakeholders play a crucial role in shaping the financial landscape. To encourage responsible innovation and competition in AI, K2 Integrity recommends the following actions:

1) **Establish a Clear Regulatory Framework:**
   - **Develop Comprehensive, Risk-Based Regulations:** Implement a regulatory framework that balances innovation with consumer protection and financial stability. This should include guidelines on ethical AI use, transparency requirements, and accountability measures.
   - **Regular Updates and Reviews:** Ensure the framework is dynamic, with periodic reviews and updates to keep pace with technological advancements and emerging risks.

2) **Promote AI Literacy:**
   - **Education and Training Programs:** Invest in comprehensive education and training initiatives for regulators, financial institutions, and consumers. This includes workshops, certification programs, and online courses focused on AI technologies, their implications, and regulatory requirements.

- **Public Awareness Campaigns:** Launch campaigns to increase public awareness of AI's benefits and risks, empowering consumers to make informed decisions.

3) **Enhance Data Privacy and Security:**
   - **Strengthen Data Protection Regulations:** Implement robust data protection laws to ensure the responsible use of data in AI models, protecting consumer privacy without impeding innovation.
   - **Encourage Best Practices:** Promote industry best practices for data security, including encryption, anonymization, and secure data storage.

4) **Monitor Systemic Risks:**
   - **Regular Risk Assessments:** Conduct regular assessments to identify and mitigate systemic risks posed by AI in the financial system, such as compliance risk, operational risk, credit risk, and market manipulation.
   - **Early Warning Systems:** Develop and implement early warning systems to detect and address potential AI-driven disruptions.

5) **International Cooperation:**
   - **Develop Global Standards:** Collaborate with international regulatory bodies to create global standards for AI in finance, ensuring consistency and a level playing field across jurisdictions.
   - **Share Best Practices:** Facilitate the exchange of best practices and lessons learned with global partners to enhance AI governance.

Part C: Further Actions

Question 19: To what extent do differences in jurisdictional approaches inside and outside the United States pose concerns for the management of AI-related risks on an enterprise-wide basis? To what extent do such differences have an impact on the development of products, competition, or other commercial matters? To what extent do such differences have an impact on consumer protection or availability of services?

K2 Integrity Response:

European Union (EU) jurisdictional approach

- The European Union approved the final text of the AI Act on 21 May 2024. The EU AI Act notably has extra-territorial reach, meaning that U.S., UK, and other non-EU jurisdictional firms will be impacted if their AI services are used by EU customers (Elbashir, M., 2024). Key developments of the Act include:
  - A four-tiered risk matrix for AI providers, from "unacceptable" to "low" risk.
  - AI systems deemed "unacceptable" (e.g. clearly threatening the "safety, livelihoods and rights of people") will be banned.

- High-risk activities (including creditworthiness assessments, health/life insurance, and border control processes) will be subject to stringent obligations before going to market.
- Fines of up to €35 million, or 7% of a firm's annual global revenue (whichever is higher), may apply (McNaul, J. and Kleingunther, K., 2024).

- In May 2024, the European Securities and Markets Authority (ESMA) issued a warning to investment firms using AI, stating that management bodies remain responsible for all of a firms' decisions, whether they are made by humans or AI tools, and must continue to protect customers. Again, this relates to US-headquartered firms with EU-based operations and customers, who must be mindful of consumer protection.
  - ESMA listed further inherent risks in May's warning, namely algorithmic bias, data quality issues, and privacy/security risks of data storage and processing within AI systems.
  - ESMA also outlined the need for effective risk management frameworks, focused on AI implementation and application. These critical frameworks should include robust governance structures, regular AI model testing, and robust monitoring of AI systems to identify and mitigate potential risks and biases. The importance of training and awareness should not be underestimated (McNaul, J. and Kleingunther, K., 2024).

United Kingdom (UK) jurisdictional approach

- In April 2023, the UK Government launched an AI Safety Institute, designed to enable the safe, reliable development and deployment of advanced AI systems. At present, the Government's top priority is understanding the capability and risk of these systems, ahead of implementing a regulatory framework (UK Department of Science, Innovation and Technology, 2024).
- Various public authorities have set out their approach to the UK's AI landscape. In April 2024, the Financial Conduct Authority ("FCA"), Bank of England, and Prudential Regulation Authority set out their response to the UK Government's AI Regulation Policy Paper from July 2022. They all welcome the proposed principles-based approach and none are advocating for further regulation at this point (Bollans S. et al, 2024).

China jurisdictional approach

China has introduced several measures, including the *Generative AI Measures* and the *Deep Synthesis Provisions*, to regulate the use of AI in various online information services. These measures emphasize the balance between promoting innovation and ensuring security, requiring service providers to implement management systems and adhere to content screening and labeling guidelines. Additionally, the *Ethical Review Measures* address the social and ethical challenges of AI development, mandating ethical reviews for certain scientific and technological activities. Non-compliance with these regulations can result in fines and other penalties. The article concludes by offering compliance suggestions for businesses operating in China's evolving AI regulatory landscape (Li at al., 2024).

- The Cyberspace Administration of China (CAC) issued the Deep Synthesis Provisions, which came into force in January 2023.
- China issued the final version of the Generative AI Measures, which took effect in August 2023. These measures were jointly adopted by seven Chinese central governmental agencies.
- In August 2023 China released guidance on labeling for generative AI services, requiring a "Generated by AI" label on AI-generated content.
- The *Ethics Review Measures* were jointly released by China's Ministry of Science and Technology and other government departments, effective from December 1, 2023.
- In June 2024 the CAC released the most recent announcement on algorithm filings. The *Algorithm Recommendation Provisions require* algorithm filing with the CAC for algorithms capable of influencing public opinion or social engagement.

United States of America (U.S.) jurisdictional approach

- Based on the latest version of the congressional bill *Advancing American AI Act*, the Senate aims to foster AI innovation and adoption within the federal government while upholding American values. (U.S. Congress. Senate, 2022)
    - The Act mandates the Department of Homeland Security (DHS) to issue policies and procedures for AI acquisition and use, addressing risks and impacts related to privacy, civil rights, civil liberties, and security.
    - It encourages agencies to modernize their systems and processes through AI applications, enhancing mission effectiveness and business efficiency. It directs the Office of Management and Budget (OMB) to identify and initiate pilot programs for new AI use cases, leveraging commercially available technologies and prioritizing privacy-preserving techniques.
    - The Act also promotes collaboration between agencies and the utilization of commercially available AI technologies. It amends existing laws to increase funding limits for innovative commercial items and extend DHS's authority to carry out prototype projects.
    - It establishes mechanisms for inventorying AI use cases, conducting pilot programs, and ensuring that AI procurement aligns with established guidelines. It requires agencies to prepare and maintain inventories of their AI use cases, share them with other agencies (where appropriate), and make them publicly available.
- In more recent developments, the White House and its subordinate agencies, particularly the Office of Management and Budget (OMB) and the Office of Information and Regulatory Affairs OIRA, have been proactive in addressing AI risks and opportunities. Through policies such as the *OMB M-24-10 AI guidance*, they have established requirements for agencies' use of AI, focusing on risk management practices, and are developing guidance for federal contracts involving AI

procurement. Proposals include issuing new guidance for AI use by recipients of federal funds and incorporating AI risk assessment into the evaluation of applications for federal funding. (Shaw, 2024)

- The White House also aims to update the regulatory review process, requiring agencies to consider the impact of AI on their regulatory actions. Additionally, the administration is considering using the *Federal Property and Administrative Services Act* to impose binding conditions on federal contractors regarding AI use, addressing risks such as discrimination and privacy breaches. (Shaw, 2024)
- The administration is also exploring the potential use of emergency powers, such as the International *Emergency Economic Powers Act*, to respond to AI-related threats to national security and critical infrastructure. The White House is actively preparing for various scenarios where AI might pose a threat and is developing response plans and memoranda outlining the president's potential actions under existing authorities. (Shaw, 2024)

The differences in jurisdictional approaches can have a significant impact on the development of products, competition, and other commercial matters. In some cases, the stricter regulations in one jurisdiction may act as a barrier to entry for companies from other jurisdictions. This can hinder innovation and limit competition, particularly for smaller companies that may not have the resources to comply with multiple sets of regulations. For example, a US-based company offering AI services to customers in the EU or China must comply with the strict regulatory requirements, even if those requirements are more stringent than those in the US. This can create a situation where the company has to develop different versions of its AI products or services for different markets.

On the other hand, some argue that stricter regulations can actually foster innovation by forcing companies to develop more responsible and ethical AI products and services (Leverton, 2024). This can create a competitive advantage for companies that are able to meet these higher standards.

The differences in jurisdictional approaches can also have an impact on consumer protection and the availability of services. Stricter regulations, such as those in the EU and China, are generally designed to protect consumers from the potential harms of AI, such as deepfakes and public opinion influencing. However, these regulations can also lead to the unavailability of certain AI services in some jurisdictions if companies are unable or unwilling to comply with the requirements.

Whilst the extent of these differences in terms of impact on consumer protection or availability of services is in its infancy, K2 Integrity has cited examples of the potential impact to consumers and regulatory limitations on certain AI product offerings in question 11.

References

Alan J, 2024. "Tesla's Ultra-Wideband Still Vulnerable to Relay Attacks Despite Upgrades," *Firewall Daily*, *The Cyber Express by Cyble*. Accessed on August 10, 2024: https://thecyberexpress.com/tesla-ultra-wideband-vulnerable-relay-attacks/

American Bankers Association Banking Journal, 2024. "Democrats introduce bill to require reimbursements for electronic transfer fraud". *American Bankers Association*. August 5, 2024. Accessed on August 6, 2024: https://bankingjournal.aba.com/2024/08/democrats-introduce-bill-to-require-reimbursements-for-electronic-transfer-fraud/

Angwin, J., and Parris Jr., T., 2016. "Facebook Lets Advertisers Exclude Users by Race". *ProPublica*. October 28, 2016. Accessed on July 27, 2024: https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race

(ASEAN) Association of Southeast Asian Nations, 2023. "SEAN Guide on AI Governance and Ethics." Accessed August 10, 2024: https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf

Bollans, S., et al, 2024. "AI update – regulatory approach to AI in financial services". *Stephenson Harwood*. May 2, 2024. Accessed on July 29, 2024: https://www.shlegal.com/insights/ai-update-regulatory-approach-to-ai-in-financial-services

(CMA). UK Competition and Markets Authority, 2024. "CMA AI Strategic Update". CMA Research and Analysis, April 29,2024. Accessed online July 29, 2024: https://www.gov.uk/government/publications/cma-ai-strategic-update/cma-ai-strategic-update

Consilient, 2022. "Our Approach." Accessed on July 29, 2024: https://consilient.com/what-we-do#why-consilient

Damiani, Jesse. 2023. "A Voice Deepfake Was Used to Scam a CEO out of $243,000." *Forbes*. September 3, 2019. Accessed online July 27, 2024: https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/

(DOJ). US Department of Justice, 2024. "Justice Department and Stanford University to Cohost Workshop, Promoting Competition in Artificial Intelligence"". DOJ Press Release, May 21, 2024. Accessed on July 29, 2024: https://www.justice.gov/opa/pr/justice-department-and-stanford-university-cohost-workshop-promoting-competition-artificial

(DSIT) UK Department of Science, Innovation and Technology, 2024. "Introducing the AI Safety Institute". DSIT policy paper, January 17, 2024. Accessed on August 7, 2024: https://www.gov.uk/government/publications/ai-safety-institute-overview/introducing-the-ai-safety-institute

Duffy, C., 2019. "Apple Card Investigation Finds No Evidence of Gender Bias, But More Transparency Needed." *CNN Business.* November 12, 2019. Accessed on July 27, 2024: https://www.cnn.com/2019/11/12/business/apple-card-gender-bias/index.html

(ECB) European Central Bank, 2018. "The revised Payment Services Directive (PSD2) and the transition to stronger payments security". ECB Press Release, March 2018. Accessed on August 7, 2024: https://www.ecb.europa.eu/press/intro/mip-online/2018/html/1803_revisedpsd.en.html

Elbashir, Mohammed, 2024. "EU AI Act sets the stage for global AI governance: Implications for US companies and policymakers". *Atlantic Council*. April 22, 2024. Accessed on July 29, 2024: https://www.atlanticcouncil.org/blogs/geotech-

cues/eu-ai-act-sets-the-stage-for-global-ai-governance-implications-for-us-companies-andpolicymakers/#:~:text=The%20AI%20Act's%20extraterritorial%20reach,of%20their%20primary%20market%20focus

EU AI Act., 2023. "Article 3." Accessed on August 10, 2024: https://artificialintelligenceact.eu/article/3/

Fatima, Rida, 2024. "BIS Report Shows Generative AI as the Top Cybersecurity Choice for Central Banks". *Tech Report.* May 28, 2024. Accessed on August 3, 2024: https://techreport.com/crypto-news/bis-report-shows-generative-ai-as-the-top-cybersecurity-choice-for-central-banks/

(FCA) Financial Conduct Authority, 2021. "Payment Services and Electronic Money – Our Approach. The FCA's role under the Payment Services Regulations 2017 and the Electronic Money Regulations 2011". FCA Guidance, November 2021. Accessed on August 7, 2024: https://www.fca.org.uk/publication/finalised-guidance/fca-approach-payment-services-electronic-money-2017.pdf

(FTC) Federal Trade Commission, 2019. "Equifax Data Breach Settlement". FTC Enforcement, February 2024. Accessed on July 27, 2024: https://www.ftc.gov/enforcement/refunds/equifax-data-breach-settlement

Fried, Ina, 2024. "Scoop: Meta won't offer future multimodal AI models in EU." *Axios*. Accessed on August 10, 2024: https://www.axios.com/2024/07/17/meta-future-multimodal-ai-models-eu

Gupta, Vishal, 2024. "Data Privacy Week: Will the US Adopt a Federal Data Privacy Law in 2024?" *Infosecurity Magazine*. January 25, 2024. Accessed on August 7, 2024: https://www.infosecurity-magazine.com/opinions/us-federal-data-privacy-law/

Harney, K.R. 2017. "Zillow Faces Lawsuit over "Zestimate" Tool that Calculates a House's Worth," *Washington Post*. May 10, 2017. Accessed on July 27, 2024: https://www.washingtonpost.com/realestate/zillow-faces-lawsuit-over-zestimate-tool-that-calculates-a-houses-worth/2017/05/09/b22d0318-3410-11e7-b4ee-434b6d506b37_story.html

Hilton J. et al, 2024. "A Right to Warn about Advanced Artificial Intelligence". Right to Warn, June 4, 2024. Accessed on July 27, 2024: https://righttowarn.ai/

Humphries, Conor, 2024. "Top EU data regulator says tech giants working closely on AI compliance." *Reuters*. May 28, 2024. Accessed on 28 July 2024: https://www.reuters.com/technology/cybersecurity/top-eu-data-regulator-says-tech-giants-working-closely-ai-compliance-2024-05-28/

IAPP Research and Insights, 2023. "International Definitions of Artificial Intelligence," *(IAPP) International Association of Privacy Professionals*. Accessed on August 7, 2024: https://iapp.org/resources/article/international-definitions-of-artificial-intelligence/

Striker, C. et al, 2024. "What is generative AI?". IBM, March 22, 2024. Accessed on August 9, 2024: https://www.ibm.com/topics/generative-ai

Intel Newsroom, 2020. "Intel and Consilient Join Forces to Fight Financial Fraud with AI". *Intel*. Accessed on July 29, 2024: https://www.intel.com/content/www/us/en/newsroom/news/fight-financial-fraud-ai.html#gs.ctftis

Lakshmanan, Ravie. 2024. "Meta Halts AI Use in Brazil Following Data Protection Authority's Ban." *The Hacker News*. Accessed on July 29, 2024: https://thehackernews.com/2024/07/meta-halts-ai-use-in-brazil-following.html

Lalchand, S. et al. 2024. "Generative AI is expected to magnify the risk of deepfakes and other fraud in banking". Deloitte, May 29, 2024. Accessed on August 7, 2024: https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html

Lawler, Edmund. 2024. "Banks face the twin-edged sword of generative AI." *Bank Administration Institute*, March 4, 2024. Accessed on August 7, 2024: https://www.bai.org/banking-strategies/banks-face-the-twin-edged-sword-of-generative-ai/

Leverton, Jaime. 2023. "Well-Done Regulation Can Spur Innovation: How Companies Can Get Involved". *Forbes Business Council.* Accessed on August 10, 2024: https://www.forbes.com/sites/forbesbusinesscouncil/2023/03/27/well-done-regulation-can-spur-innovation-how-companies-can-get-involved/

Li, Barbara et al. 2024. "Navigating the Complexities of AI Regulation in China." *Reed Smith In-depth,* August 7, 2024. Accessed on August 11, 2024: https://www.reedsmith.com/en/perspectives/2024/08/navigating-the-complexities-of-ai-regulation-in-china

Lloyd's Bank. 2024. "Financial Institutions Sentiment Survey". *Lloyd's Bank*, July 16, 2024. Accessed on July 29, 2024: https://www.lloydsbank.com/business/resource-centre/insight/financial-institutions-sentiment-survey.html

McMahan. 2017. "Federated Learning: Collaborative Machine Learning Without Centralized Training Data," *Google* (as cited in FinRegLab, 2020). Accessed on July 29, 2024: https://finreglab.org/research/ai-faqs-federated-machine-learning-in-anti-financial-crime-processes/#kqy-6

Maume, P. 2021. "Robo-Advisors: How Do They Fit in the Existing EU Regulatory Framework, in Particular with Regard to Investor Protection?" Policy Department for Economic, Scientific, and Quality of Life Policies, European Parliament, June 2021. Accessed on July 27, 2024: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662928/IPOL_STU(2021)662928_EN.pdf

McNaul, J., and Kleingunther, K., 2024. "The Ever-Changing Landscape of Artificial Intelligence". *UK Finance*, July 12, 2024. Accessed on July 12, 2024: https://www.ukfinance.org.uk/news-and-insight/blog/ever-changing-landscape-artificial-intelligence

Maio, H. 2022. "U.S. Banks Close Record Number of Retail Branches in 2021, Wells Fargo Shutters Most." *CNBC.* January 21,2022. Accessed on July 27, 2024: https://www.cnbc.com/2022/01/21/banks-close-record-number-of-branches-in-2021-led-by-wells-fargo.html

Muckleroy, Julie. 2024. "AI's impact on fraud: A growing challenge for global banking". *SAS Institute Inc*, February 29, 2024. Accessed on August 7, 2024: https://blogs.sas.com/content/sascom/2024/02/09/ais-impact-on-fraud-a-growing-challenge-for-globalbanking/#:~:text=The%20explosion%20and%20accessibility%20of,fraud%20due%20to%20generative%20AI.

(PFSA) UK Public Sector Fraud Authority, 2024. "Introduction to AI Guide with a focus on Counter Fraud (HTML)". *PFSA Guidance*, March 18,2024. Accessed on July 29, 2024: https://www.gov.uk/government/publications/introduction-to-ai-with-a-focus-on-counter-fraud/introduction-to-ai-guide-with-a-focus-on-counter-fraud-html

(PSR) UK Payment Systems Regulator. 2022. "Extending Confirmation of Payee coverage. Response to consultation CP22/2". *PSR Policy Statement*, October 2022. Accessed on August 7, 2024: https://www.psr.org.uk/media/migeob4s/ps22-3-extending-cop-coverage-oct-2022.pdf

(PSR) UK Payment Systems Regulator. 2023. "Fighting authorised push payment scams: final decision". *PSR Policy Statement*, December 2023. Accessed on August 7, 2024: https://www.psr.org.uk/media/kwlgyzti/ps23-4-app-scams-policy-statement-dec-2023.pdf

Rieke, Nicola. 2019. "What is Federated Learning?" *NVIDIA*. October 13, 2019. Accessed on July 29, 2024: https://blogs.nvidia.com/blog/what-is-federated-learning/

Robins-Early, 2024. "OpenAI and Google DeepMind Workers Warn of AI Industry Risks in Open Letter." *The Guardian*. 4 June, 2024. Accessed on July 27, 2024: https://www.theguardian.com/technology/article/2024/jun/04/openai-google-ai-risks-letter

Robson, Kurt. 2024. "Employee duped into wiring $25m of company funds by video deepfake scam". *Verdict*. February 6, 2024. Accessed on August 7, 2024: https://www.verdict.co.uk/employee-sends-25m-of-company-funds-after-video-call-with-ai-deepfake-ceo/?cf-view

(SEC) U.S. Securities and Exchange Commission, 2024. "SEC Charges Two Investment Advisers with Making False and Misleading Statements About Their Use of Artificial Intelligence." *SEC Press Release*, March 18, 2024. Accessed on July 27, 2024: https://www.sec.gov/news/press-release/2024-36

Schneider, Henrique. 2024. "Meltdown in an over-networked world," *GIS reports Online*. Accessed on August 10, 2024: https://www.gisreportsonline.com/r/crowdstrike-networks/

Shaw, R. et al. 2024. "Taking Further Agency Action on AI. Center for American Progress". *Center for American Progress,* June 17, 2024. Accessed on August 10, 2024: https://www.americanprogress.org/article/taking-further-agency-action-on-ai/

Shiffman, Gary M. et al. 2023. "Artificial Intelligence and the Revolution in Financial Crimes Compliance". *Global Association of Risk Professionals.* Accessed on August 12, 2024: https://www.garp.org/hubfs/Whitepapers/a2r5d000006RYkPAAW_RiskIntell.WP.Artificial%20Intelligence%20and%20the%20Revolution%20in%20Financial%20Crimes%20Compliance.11.22.pdf

Stewart, Adam. 2024. "Achieve Control and Scale In A Privacy-Safe Way With Google's AI-Powered Commerce Solutions." *Forbes*, May 1, 2024. Accessed on July 29, 2024: https://www.forbes.com/sites/think-with-google/2024/05/01/achieve-control-and-scale-in-a-privacy-safe-way-with-googles-ai-powered-commerce-solutions/?

UK Finance. 2024. "Annual Fraud Report 2024". *UK Finance report*, May 22, 2024. Accessed on July 28, 2024: https://www.ukfinance.org.uk/policy-and-guidance/reports-and-publications/annual-fraud-report-2024

U.S. Congress. Senate. 2022. "Advancing American AI Act", *S. 1353, 117th Cong., 2d sess*. Accessed on August 10, 2024: https://www.congress.gov/bill/117th-congress/senate-bill/1353/text

Visa. 2024. "Visa Announces Generative AI-Powered Fraud Solution to Combat Account Attacks". *Visa Press Release, July 5, 2024.* Accessed on July 29, 2024: https://usa.visa.com/about-visa/newsroom/press-releases.releaseId.20661.html