



MASTER THESIS | MASTER'S THESIS

Titel | Title

Global Strategies in Addressing the Emerging Challenge of Artificial
Intelligence Crime within the Financial Sector

verfasst von | submitted by

Ornella Reinert

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Laws (LL.M.)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears
on the student record sheet:

UA 999 082

Universitätslehrgang lt. Studienblatt |
Postgraduate programme as it appears
on the student record sheet:

Europäisches und Internationales
Wirtschaftsrecht [Vollzeit Englisch]

Betreut von | Supervisor:

ao. Univ.-Prof. Mag. Dr. Siegfried Fina

Abstract

The study examines the integration of AI in the financial sector, highlighting both its transformative potential and inherent risks. It explores various AI-driven innovations such as algorithmic trading, fraud detection, and credit scoring, which have significantly reshaped financial services while presenting risks. These technologies, while beneficial, present significant risks of data breaches, false positives, algorithmic bias, and more. Through a comprehensive analysis of regulatory frameworks across different jurisdictions, including the US, UK, EU, and international institutions, the study assesses the effectiveness of current regulations in mitigating AI-enabled crimes (AIC). It reviews the regulatory challenges posed by AI's autonomous nature and the complexities of assigning liability for AI-driven actions. The study highlights significant cases, such as the Knight Capital Group incident and the Apple Card gender discrimination case, to illustrate the real-world implications of AI-related regulatory issues. The analysis reveals a fragmented regulatory landscape with varying approaches to AI governance. The US adopts a flexible market-driven model focusing on innovation and implementing a sector-specific approach to AI governance, while the UK employs a principles-based approach leveraging existing sector-specific regulators. The EU, on the other hand, is moving towards comprehensive AI regulation with the AI Act, emphasizing stringent requirements for high-risk AI systems.

The thesis concludes that a unified, strong legal framework and comprehensive definition of AI-enabled crimes should be established to effectively identify and address them. States should intensify efforts to develop regulations targeting AIC, as current gaps make it challenging for courts to pursue AIC cases. This proactive approach will help prevent legislative deficiencies and close existing legal gaps in this area.

Keywords: AI, financial sector, AI-driven innovations, AI-enabled crime (AIC), regulatory frameworks, AI governance, US AI regulation, UK AI regulation, EU AI regulation, liability in AI, AI compliance, financial crimes.

Die Studie untersucht die Integration von KI im Finanzsektor und hebt sowohl ihr transformatives Potenzial als auch die inhärenten Risiken hervor. Sie erforscht verschiedene KI-gesteuerte Innovationen wie den algorithmischen Handel, die Betrugserkennung und die Kreditbewertung, die die Finanzdienstleistungen erheblich umgestaltet haben, aber auch Risiken mit sich bringen. Diese Technologien sind zwar vorteilhaft, bergen jedoch erhebliche Risiken wie Datenverletzungen, falsche Positivmeldungen, algorithmische Voreingenommenheit und mehr. Durch eine umfassende Analyse der Regulierungsrahmen in verschiedenen Rechtsordnungen, einschließlich der USA, Großbritanniens, der EU und internationaler Institutionen, bewertet die Studie die Wirksamkeit der aktuellen Vorschriften zur Minderung von KI-gestützten Verbrechen (AIC). Sie untersucht die regulatorischen Herausforderungen, die sich aus der autonomen Natur der KI und den Komplexitäten bei der Zuweisung von Haftung für KI-gesteuerte Handlungen ergeben. Die Studie hebt bedeutende Fälle hervor, wie den Vorfall bei der Knight Capital Group und den Fall der geschlechtsspezifischen Diskriminierung bei der Apple Card, um die realen Auswirkungen von KI-bezogenen Regulierungsfragen zu veranschaulichen. Die Analyse zeigt eine fragmentierte regulatorische Landschaft mit unterschiedlichen Ansätzen zur KI-Governance. Die USA verfolgen ein flexibles, marktorientiertes Modell mit Fokus auf Innovation und setzen einen sektorspezifischen Ansatz zur KI-Governance um, während Großbritannien einen prinzipienbasierten Ansatz verfolgt, der bestehende sektorspezifische Regulierungsbehörden nutzt. Die EU hingegen bewegt sich mit dem KI-Gesetz in Richtung einer umfassenden KI-Regulierung, die strenge Anforderungen für Hochrisiko-KI-Systeme betont.

Die Arbeit kommt zu dem Schluss, dass ein einheitlicher, starker Rechtsrahmen und eine umfassende Definition von KI-gestützten Verbrechen geschaffen werden sollten, um diese effektiv zu identifizieren und zu bekämpfen. Die Staaten sollten ihre Bemühungen verstärken, um Vorschriften zur Bekämpfung von AIC zu entwickeln, da die derzeitigen Lücken es den Gerichten erschweren, AIC-Fälle zu verfolgen. Dieser proaktive Ansatz wird dazu beitragen, gesetzgeberische Defizite zu verhindern und bestehende rechtliche Lücken in diesem Bereich zu schließen.

Schlagwörter: Künstliche Intelligenz (KI), Finanzsektor, KI-gesteuerte Innovationen, KI-gestützte Kriminalität (AIC), KI-Governance, US-KI-Regulierung, UK-KI-Regulierung, EU-KI-Regulierung, Haftung in der KI, KI-Compliance, Finanzkriminalität.

Table of contents:

<i>Abstract</i>	<i>I</i>
<i>List of abbreviations</i>	<i>V</i>
Chapter I. Introduction	1
Chapter II. Background of AI	3
1. History of AI	3
2. AI integration in the financial sector – The fourth Industrial Revolution	5
2.1. AI-Driven Innovations in Financial Institutions: Efficiency, Risk Management, and Customer Engagement	5
2.2. Addressing Challenges in AI Implementation within Financial Institutions	9
3. The Dark Side of AI: Emergence of AIC	11
3.1. Perpetrators of AIC: Human Agents, Artificial Agents, or Both?	13
3.2. Financial crimes in the age of Artificial Intelligence	15
a. Algorithmic Market manipulation	15
b. Collusion	16
c. Digital fraud	18
Chapter III. Regulatory challenges of AIC	19
1. Legal challenge: attribution and liability	20
2. Balancing Innovation and Protection	22
3. The AI Heist: A Real-World Case of Financial Crime	24
3.1. The Knight Capital Group case	24
3.2. The apple gender discrimination case	26
Chapter IV. Comparative Analysis of the Current Regulatory Landscape: The race for AI supremacy	28
1. Regulatory Approaches in Combating AI Challenges: A Comparative Analysis between UK and US	29
1.1. Balancing Innovation and Ethics: A Market-Driven Approach to AI Regulation for a Good AI Society in the United States	29
1.2. The pro-innovation approach in the UK	33
2. The European Commission: Pioneer of AIC Regulations	38
Chapter V. International Organizations' Plans for a Unified Global Response	42
1. The OECD's Recommendation on Artificial intelligence	43
2. United nations Resolution on AI Governance	45
Chapter VI . Comparative Analysis of Regulatory Principles	46
Chapter VII. Conclusion	50
Bibliography	52

(A) Articles	52
(B) Books	52
(C) Reports	53
(D) Master Theses	53
(E) Papers	53

List of abbreviations

Abbreviation	Definition
AA	Artificial Agent
AGI	Artificial General Intelligence
AI	Artificial Intelligence
AI HLEG	High-Level Expert Group on Artificial Intelligence
AIA	Artificial Intelligence Act
AIC	Artificial Intelligence Crime
AILD	Artificial Intelligence Liability Directive
AML	Anti-Money Laundering
ANI	Artificial Narrow Intelligence
ANI	Artificial Narrow Intelligence
ANN	Artificial Neural Network
CFPB	Consumer Financial Protection Bureau
DLT	Distributed Ledger Technology
DORA	Digital Operational Resilience Act
ECB	European Central Bank
ESA	European Supervisory Authorities
EU	European Union
FCA	Financial Conduct Authority
FOSTA	Fight Online Sex Trafficking Act of 2017
G20	Group of Twenty
G7	Group of Seven
GDPR	General Data Protection Regulation
GenAI	Generative Artificial Intelligence
HL	House of Lords

HLEG	High-Level Expert Group
IoT	Internet of Things
KYC	Know Your Customer
ML	Machine Learning
NYDFS	New York Department of Financial Services
OECD	Organization for Economic Co-operation and Development
PRA	Prudential Regulation Authority
SEC	Securities and Exchange Commission
SI	Superintelligence
SSRN	Social Science Research Network
TFEU	Treaty on the Functioning of the European Union
UK	United Kingdom
UNGA	United Nations General Assembly
US	United States

Chapter I. Introduction

The global financial sector has faced a significant surge in sophisticated cybercrimes in recent years, with AI playing a pivotal role in these malicious activities. The rapid advancement of AI technology has brought about unprecedented opportunities for financial institutions to enhance their operations and services. However, this same technology has also been exploited by cybercriminals to orchestrate complex and devastating attacks. In 2023, AI-driven fraud alone resulted in billions of dollars in losses, underscoring the pressing challenge that financial institutions worldwide must confront to ensure the stability and security of the global financial system¹. As AI continues to evolve and become increasingly integrated into financial systems, it is crucial for both governments, regulatory bodies and financial institutions to understand the impact of these technologies on financial stability and to develop effective industry standards and regulations for mitigating the risks associated with AI-driven cybercrimes.² This thesis, titled '*Global Strategies in Addressing the Emerging Challenge of Artificial Intelligence Crime within the Financial Sector*', aims to explore the intersection of AI and financial stability. It examines the benefits and risks of AI adoption in the financial sector and highlights various approaches developed to enhance cybersecurity, prevent the misuse of AI in financial crimes, and protect consumers. The analysis includes discussions from the UK, US, EU, and international bodies.

The motivation for this thesis stems from a deep interest developed during an internship, where participation in an ad hoc committee focused on drafting a Comprehensive International Convention on Countering the Use of Information and Communications Technologies for Criminal Purposes. This experience sparked a keen interest in exploring the intersection of AI and regulatory frameworks aimed at combating cybercrimes. This subject became particularly compelling as the European AI Act was actively being debated and gaining increasing significance. The intersection of AI regulation and cybersecurity

¹ According to Deloitte, the losses from generative AI-enabled fraud in the United States alone were projected to reach \$12.3 billion in 2023, with expectations of increasing to \$40 billion by 2027 due to the rapid advancement and adoption of these technologies. The Federal Trade Commission (FTC) also reported that total fraud losses exceeded \$10 billion in 2023, marking a significant increase from previous years, with a substantial portion attributed to AI-related scams. Lalchand, S., Srinivas, V., Maggiore, B., & Henderson, J. (2024). Deepfake banking fraud risk on the rise. Deloitte. <https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html>

²Klein, A., & Lee, S. (2024, July 5). AI and financial stability: Mitigating risks, harnessing benefits. Brookings. <https://www.brookings.edu/articles/ai-and-financial-stability-mitigating-risks-harnessing-benefits/> accessed the 10 July 2024.

challenges presents a fascinating and timely area of study, especially given the rapid advancements and widespread integration of AI technologies.

Since AI is a topic that is currently undergoing active development and discussion, this master's thesis does not aim to provide an in-depth analysis of the regulatory approaches of the UK, US, and European Union. Instead, it offers a brief overview of the rationale behind these approaches. In this thesis, the term 'AI system'³ will be used in a broad sense, encompassing various forms of AI without delving into the specifics of each individual system, as there has still not been a commonly agreed definition for what AI systems are.⁴ The focus will be on the overarching implications and legal considerations of AI, rather than differentiating between specific types of AI technologies.⁵ This approach allows for a comprehensive analysis of AI in a global context, acknowledging the diverse applications and impacts of AI without getting bogged down in technical details. For the purposes of this thesis, AI systems will refer to any product or application utilizing AI to perform tasks typically requiring human intelligence. The primary methodology for this thesis will involve an extensive review and analysis of existing research on AI and its legal implications. The focus will be on examining the regulatory approaches to AI-enabled crimes and AI regulation in general, specifically within the contexts of the UK, US, international organizations, and the EU. By leveraging existing studies and data, the thesis aims to project potential future developments in AI regulation and the challenges they may pose. This analysis will provide insights into the evolving landscape of AI and its regulation, highlighting critical issues and possible solutions.

The thesis will begin with an introduction to AI (II), including its history and integration into the financial market, and will highlight both its benefits and challenges. Following this, the concept of AI-enabled crimes will be presented (III), identifying the perpetrators of these crimes and detailing the various forms these crimes can take, supported by specific case studies. The core of the thesis (IV) will then examine the different regulatory approaches of the US, UK, international organizations (V), and the EU, outlining the solutions currently being discussed or implemented to address this growing challenge. Then will follow up an

³ AI systems can be described as those that display intelligent behavior by analyzing their environment and taking actions autonomously to achieve specific goals. Sheikh, H., Prins, C., Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In: Mission AI. Research for Policy. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_2 accessed the 10 July 2024

⁴ Boden, Margaret A., 'What is artificial intelligence?', *Artificial Intelligence: A Very Short Introduction*, Very Short Introductions (Oxford, 2018; online edn, Oxford Academic, 23 Aug. 2018), <https://doi.org/10.1093/actrade/9780199602919.003.0001>, accessed 11 July 2024.

⁵ *ibid* 3.

analysis of the different regulatory principles (VI). The conclusion (VII) will offer a final assessment of the analysis.

Chapter II. Background of AI

Artificial Intelligence encompasses a broad range of technologies and applications designed to mimic human cognitive abilities. While lacking a universally accepted definition, AI generally refers to developing computer systems capable of tasks traditionally requiring human intelligence. These tasks include learning from data, reasoning through complex problems, and adapting to new situations.

Over the years, AI has significantly evolved, driven largely by advancements in artificial neural networks (ANNs). These increasingly complex networks, trained on massive datasets (Big Data), have facilitated breakthroughs in areas like image and speech recognition, natural language processing, and autonomous decision-making. The financial services industry has witnessed a surge in AI adoption. However, this integration raises concerns about potential risks to consumers. For instance, the Singapore Summit 2020 highlighted customer protection as a key concern. Dr. Marcin Detyniecki, Group Chief Data Scientist and Global Head of R&D at AXA, emphasized the need to prioritize customer safety considering AI's novel challenges.⁶

Despite the clear need for safeguards, regulating AI presents a unique challenge due to its vast and sometimes elusive nature. The broad definition of AI makes it difficult to define and pinpoint exactly what needs regulation.

1. History of AI

The history of autonomous machines began in the late 1930s and 1940s with Isaac Asimov's science fiction book 'Runaround,' which inspired scientists to work on AI. Alan Turing developed the Theory of Computation and proposed the Turing Test to measure a machine's intelligence. The Dartmouth Summer Research Project on AI in 1956 marked the beginning

⁶ Detyniecki, M. (2018, september 15). *2018 Singapore Summit*.

of AI research, and the term ‘computational rationality’ was coined. The ELIZA Computer, introduced in the 1960s, was one of the first to attempt the Turing Test, leading to substantial government funding. However, this optimism was short-lived, and AI research faced criticism and funding freezes in the 1970s, known as the first AI winter. AI research recovered in the 1980s with the introduction of expert systems, which allowed AI programs to solve problems within specific domains. This innovation led to further research initiatives and the development of chess-playing programs that could defeat human professionals. Despite this progress, expert systems proved costly and difficult to upgrade, leading to a slowdown in public funding.⁷

The rise of machine learning in the late 1980s and 1990s was a game-changer. Inspired by the brain, neural networks learned from data, leading to breakthroughs in speech recognition and computer vision. The 2000s saw deep learning, a more complex form of neural networks, further revolutionize AI. These networks tackled massive datasets, enabling advancements in natural language processing and accelerating self-driving car development. In this era, the classification of AI expanded to encompass different stages of capability. The first stage is the Artificial Narrow Intelligence (ANI) which is related to machine learning, this system specializes in one area and solves one problem, known illustrations for ANI are SIRI and ALEXA. This form of AI is faster and more accurate than a human would be, but it lacks general intelligence and does not surpass humans across all cognitive abilities. The second stage is called the artificial general intelligence (AGI) which refers to machine intelligence, this is according to some scientists the actual stage which technology currently stands at. AGI encompasses machines capable of executing any intellectual task achievable by humans. These systems possess the capacity to plan, make decisions amidst uncertainty, and mimic human behavior to such a degree that they could successfully deceive or manipulate individuals. Finally, we are currently heading to the final stage, Artificial Superintelligence (SI) which refers to the machine consciousness. Futurists like Nick Bostrom warn of the potential rapid emergence of superintelligence, which could lead to unpredictable advancements in technology beyond human control.⁸

⁷ Harvard University. (2017, August 28). *The history of artificial intelligence*. Science in the News. <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/> accessed the 15 May 2025

⁸ Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Today, AI is omnipresent, from virtual assistants to fraud detection and medical diagnosis. The future of AI is one where the benefits of this technology become more integrated into our daily lives, with opportunities to impact every possible area.

2. AI integration in the financial sector – The fourth Industrial Revolution

AI has permeated numerous sectors, fundamentally reshaping diverse aspects of our lives. In healthcare, AI-powered image recognition software assists professionals in early disease detection through medical scan analysis. In manufacturing, AI optimizes production lines by automating processes, predicting equipment failures, and streamlining logistics. The transportation industry showcases AI integration with the development of self-driving cars. While AI has a long history of development, recent technological advances and digitization have underpinned rapid and unprecedented evolution. Algorithmic trading, an early and leading AI use case, has, according to reports from the European Central Bank, it has 'been growing steadily since the early 2000s and, in some markets, is already used for around 70% of total orders'.⁹

The financial sector is another domain experiencing significant transformation due to AI integration. For the purpose of our research, we will focus on this specific field, where AI's impact is revolutionizing fraud detection, algorithmic trading, credit scoring, portfolio management, regulatory compliance, and many other aspects of financial services (i). But while AI offers considerable benefits, it also presents many risks for financial institutions and potentially the wider financial system (ii).

2.1. AI-Driven Innovations in Financial Institutions: Efficiency, Risk Management, and Customer Engagement

The digital transformation within the financial ecosystem, often referred to as 'Industry 4.0', began in 1982. This initiative aimed to enhance business processes and decision-making capabilities within banks and financial institutions through the integration of advanced

⁹ Buckley, R. P., Zetzschek, D. A., Arner, D. W., & Tang, B. W. (2021). Regulating Artificial Intelligence in Finance: Putting the Human in the Loop. *Sydney Law Review*, 43(1), 43+. <https://link-gale-com.uaccess.univie.ac.at/apps/doc/A663776414/ITOF?u=43wien&sid=bookmark-ITOF&xid=c190ccf9> accessed on the 10 June 2024

technologies. By leveraging new technologies, financial institutions sought to improve efficiency, reduce costs, and offer better services to their customers. By the late 1990s, the concept of the Fintech industry had emerged, driven by significant advancements in computer systems and the expansion of electronic banking during the 1970s and 1980s¹⁰. Fintech, short for ‘financial technology’ encompasses a wide range of innovative business models that leverage technology to revolutionize the financial services industry. These models include online banking, mobile payments, digital lending platforms, and automated investment services, all of which utilize the internet and automation to provide more accessible, efficient, and personalized financial products and services¹¹. Nevertheless, fintech also has the potential to disrupt traditional financial services by offering lower costs, greater convenience, and enhanced user experiences. It simultaneously promotes financial inclusion by providing services to underserved populations who may not have access to traditional banking¹². Indeed, the application of AI within the financial sector solves a huge issue: *financial exclusion*¹³. As previously mentioned previously, AI has the ability to solve a huge issue better than what any past and present regulations have been able to achieve. According to author Clare Chambers-Jones, ‘*AI and other technologies could only be a solution for some people who are financially excluded if the regulation and policy were drafted to ensure they are safeguarded and have at their heart the good of the people they are being drafted to serve.*’¹⁴ Throughout the years the usage of AI systems within the financial sector has grown as much as AI applications are present in most activities.¹⁵ Initially the earlier adopters of fintech offered fundamental financial services such as online stock trading and electronic banking,

¹⁰ RegTechOne. (2023, February 7). *The history and evolution of the fintech industry*. Retrieved from <https://regtechone.co/en/the-history-and-evolution-of-the-fintech-industry/#:~:text=The%20fintech%20industry%20as%20we,in%20the%201970s%20and%201980s> Consulted the 7 July 2024.

¹¹ Paul, Sandeep. 2019. Use of Blockchain and Artificial Intelligence to Promote Financial Inclusion in India Smita Miglani Indian Council for Research on International Economic Relations.

¹² According to the Global Findex Survey, 79.9% of the global population aged 15 and above had accounts with financial institutions in 2017. This represents a significant increase from 53.1% in 2014 and 35.2% in 2011. Demircuc-Kunt, Asli, Leora Klapper, Saniya Ansar, and Aditya Jagati. 2017. *Making It Easier to Apply for a Bank Account: A Study of the Indian Market*. <https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-8205> .Accessed the 4 July 2024

¹³ “*Financial exclusion*” is defined as the inability for particular societal groups to access mainstream financial services. Carbó, S., Gardener, E. P. M., & Molyneux, P. (2005). *Financial exclusion*. Palgrave Macmillan. <https://doi.org/10.1057/9780230508743> Accessed the 7 July 2024

¹⁴ Lui, A., & Ryder, N. (Eds.). (2023). *FinTech, Artificial Intelligence and the Law: Regulation and Crime Prevention* (p. 193). Routledge. ISBN 9781032012469.

¹⁵ AI applications in the digital finance include robo-advisors, AI-based personalized asset management systems, statistical credit underwriting and risk assessment applications, automated and intelligent KYC (Know Your Customer) applications, fraud detection and anti-money laundering (AML), and personalized finance applications for citizens and businesses, as well as a variety of front-office applications such as chatbots. This is a non-exhaustive list of applications as it’s estimated that the more data becomes accessible the more fintech will improve and grow. *Ibid* 16.

then new products appeared which included payments, loans and insurance. These opportunities are said to have been created by four distinct developments: the increase in the computational and data processing power of IT systems. Secondly, data has become available on an unprecedented scale. Third, the costs associated with the storage of data have fallen. Finally, fourth, increasingly sophisticated software services have come onto the market.¹⁶ AI is being used across a range of functions within financial services firms, including:

- Anti-money laundering activities;
- Credit and regulatory capital modeling;
- Insurance claims management;
- Product pricing;
- Order routing and execution of trades;
- and Cybersecurity.

Fintech's surge has shaken up the financial world for traditional institutions. Tech titans like Google, Apple, Amazon, and PayPal are now offering financial services, fueled by customer desires for lower costs, faster adaptation, and better experiences. This flood of new competitors has forced traditional banks and financial institutions to constantly innovate to stay relevant and win the competition.

A key branch of Fintech, Regtech, emerged around 2015. It uses digital tools like anti-money laundering (AML) systems, know-your-customer (KYC) processes, and risk management systems to make following regulations easier. Regtech's goal is to manage risks (financial and non-financial), streamline regulatory compliance reporting, and create compliant business practices and data handling. By pinpointing how regulations affect business models, Regtech helps companies navigate and meet compliance requirements. Despite the advantages Regtech brings, such as improved efficiency and enhanced compliance, it also poses several risks. These technology systems, while offering powerful compliance tools, can permit computer programmers to interpret legal requirements, potentially introducing biases and errors. They can obscure the very risks policy makers aim to mitigate, leading to an 'automation bias' that may privilege personal self-interest over sound judgment. Moreover, the lack of transparency in these systems can hinder oversight and accountability, making it challenging to ensure that they operate as intended and comply with regulatory standards¹⁷. The integration of

¹⁶ Baker, D.J., & Robinson, P.H. (Eds.). (2020). *Artificial Intelligence and the Law: Cybercrime and Criminal Liability* (1st ed.) (p.34). Routledge. <https://doi-org.uaccess.univie.ac.at/10.4324/9780429344015>.

¹⁷ Bamberger, K. A. (2009). Technologies of compliance: Risk and regulation in a digital age. *Texas Law Review*, 88, 669–739.

Distributed Ledger Technology (DLT) into financial systems has further propelled the efficiency of these services. DLT enhances transparency and security by distributing information across multiple nodes, making financial transactions more resistant to fraud and cyber-attacks. It supports not just cryptocurrency transactions but also helps in managing complex multi-party contracts and ensuring compliance through transparent and immutable records. Indeed, AI has significantly impacted digital financial inclusion by enhancing risk detection, measurement, and management; addressing information asymmetry; providing customer support through chatbots; and improving fraud detection and cybersecurity. Therefore, it is recommended that financial institutions continue to expand the use of AI tools and applications, given the substantial benefits they offer.¹⁸

These advancements, however, also raise new challenges and ethical concerns, particularly in relation to data privacy and manipulation of decision-making. As Sundblad¹⁹ noted, some challenges relate to data quality and the responsibility requirements in the rollout of AI technology. To ensure the benefits of big data, which fuels effective AI, companies must find efficient ways to manage and analyze large volumes of data²⁰. They need to be able to manage both streaming data and data at rest this process is known as ‘big data analytics’.²¹ Ensuring high data quality is critical, as poor data can lead to incorrect AI predictions and decisions, potentially causing financial losses and damaging customer trust. Moreover, the implementation of AI systems requires clear responsibility frameworks to manage accountability, especially when AI decisions impact customer outcomes. Financial institutions must address these challenges by establishing robust data governance practices and ensuring transparency in AI operations. This has introduced a race to technological leadership among nations, primarily China and the US who are running to dominate this branch (*see Figure 1*).²²

¹⁸ Mhlanga, D. (2020). *Industry 4.0 in Finance: The Impact of AI on Digital Financial Inclusion*. School of Accounting, University of Johannesburg, Johannesburg, South Africa; Published: 28 July 2020.

¹⁹ Sundblad, W. (2018, October 18). *Data is the foundation for artificial intelligence and machine learning*. Forbes. Retrieved from <https://www.forbes.com/sites/willemsundbladeurope/2018/10/18/data-is-the-foundation-for-artificial-intelligence-and-machine-learning/#12324a7751b4> accessed on 10 June 2024.

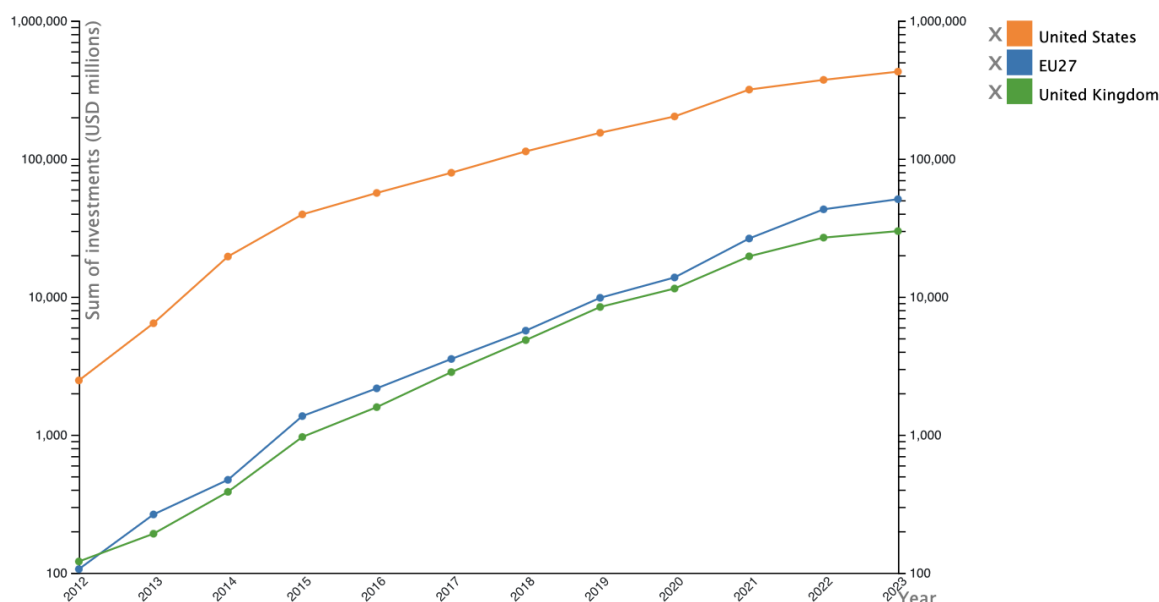
²⁰ This involves not only datasets from legacy banking systems (e.g., customer accounts, customer transactions, investment portfolio data), but also other data sources such as financial market data, regulatory datasets, social media data, real-time retail transactions, and more.

²¹ Zhu, J. (2022). *Big Data and AI in Digital Finance*. Springer. <https://doi.org/10.1007/978-3-030-94590-9> accessed the 11 July 2024.

²² Shabsigh, G., & Boukherouaa, E. B. (2023). *Generative artificial intelligence in finance: Risk considerations*. International Monetary Fund, pp. 4. <https://www.imf.org/en/Publications/fintech-notes/Issues/2023/08/18/Generative-Artificial-Intelligence-in-Finance-Risk-Considerations-537570> accessed the 12 June 2024

This could lead to many countries seeing investments in AI as strategic according to writers Agrawal, J. Gans, and Goldfarb ²³ , which is what is currently happening.

Figure 1 : OECD investment in AI by Country



Source : *OECD.AI (2024)*, visualizations powered by JSI using data from Preqin, accessed on 21/7/2024, www.oecd.ai

2.2. Addressing Challenges in AI Implementation within Financial Institutions

Despite AI models being adaptable, flexible, and scalable, they face significant challenges that impact their robustness and reliability. AI systems can develop unexpected and potentially harmful capabilities when applied to new use cases. Furthermore, AI could be misused in ways that increase their threat potential. The autonomous and self-learning aspects of AI underpin both beneficial and malicious uses.

²³ Agrawal, Ajay, Joshua Gans, and Avi Goldfarb (2019). “Economic Policy for Artificial Intelligence”. In: *Innovation Policy and the Economy*. Ed. by Josh Lerner and Scott Stern. Vol. 19. The University of Chicago Press, pp. 139–159.

A May 2024 review by the European Central Bank (ECB) highlights two primary concerns related to AI use by financial institutions: algorithmic bias²⁴ and a lack of transparency, which leads to a lack of accountability. Algorithmic bias arises when AI systems produce unfair results, leading to unequal outcomes. A prime example of algorithmic bias is illustrated in the book ‘Artificial Intelligence in Finance’ (2023) by Nydia Remolina and Aurelio Gurrea-Martinez²⁵. The scenario involves identical twins, Ms. Beep and Mr. Beep, who applied for the same credit card from a tech company that employs an AI model for approval. Despite having nearly identical backgrounds and credit histories, Ms. Beep received a significantly lower credit limit than her brother. This puzzling difference left the twins frustrated and seeking answers about the AI's decision-making process, ultimately questioning who is responsible for the seemingly unfair outcome. Indeed, as the use of AI technology increases in the financial sector it's become more evident that AI decision-making is prone to bias and, consequently, to introduce discriminatory outcomes they become equally capable of presenting bias and perpetuating discrimination as a human agent would²⁶. This bias is closely associated with the ‘black box’ nature of AI models, where the decision-making processes are not easily understood, even by their developers. Indeed, research²⁷ was able to introduce four key mechanisms of bias arising from AI systems: developer biases, historical biases in training data, biases resulting from lack of appropriate training data, stability bias. Numerous examples of these can be found in the assess creditworthiness which cannot be resolved through the exclusion of the race or any correlated factors to race which would only have a slight impact on the disparity. In a recent case, a multinational technology company was accused of developing a credit card which would determine customers’ credit limits was discriminating against customers on the basis of their gender. This was brought to light through a twitter post made by one of the victims²⁸. The operators of the credit card defended their algorithm stating that credit decisions were ‘*based on a customer’s creditworthiness and not on factors like gender, race, age, sexual orientation or any other basis prohibited by law*’. Investigations made by the New York State department of financial services found there to be a lack of transparency in the approval and appeals process of the models and algorithms as the

²⁴ Bias is the fact of a particular subject or thing or, alternatively, an ‘inclination or prejudice for or against one person or group, especially in a way considered to be unfair’. (...) Bias can have both beneficial and adverse effects. Remolina, N., & Gurrea-Martinez, A. (2023). *Artificial Intelligence in Finance*. Edward Elgar Publishing.

²⁵ *Ibid* 24.

²⁶ *ibid.* p. 320.

²⁷ *ibid.* p. 324.

²⁸ (BBC, 11 November 2019) <<https://www.bbc.com/news/business-50365609>> accessed 25 June 2024.

filter used were considered protected characteristics²⁹. Additionally, AI models can produce false or misleading information, known as ‘hallucinations,’ which are especially problematic in deep learning systems due to their inherent complexity.³⁰ This lack of transparency and explainability makes it difficult to identify and correct errors or unintended consequences, raising significant concerns about accountability and reliability. As AI continues to evolve, regulations are being developed to ensure that AI development aligns with market safety, consumer protection, and market integrity objectives. As of today, there is little AI-specific regulation in force. Even less so that address the issues of bias and discrimination in the context of financial services and other sectors.

3. The Dark Side of AI: Emergence of AIC

Recent advancements in AI have heightened cybersecurity concerns, given the diverse methods criminals can employ AI for disruptive and lucrative purposes. There is a growing recognition that AI could become increasingly pivotal in criminal endeavors.³¹ While AI presents opportunities for financial institutions to boost cybersecurity and combat illegal activities, it also poses risks. AI can be leveraged to destabilize markets and assist criminals in evading detection and accountability, underscoring its dual-edged nature. AI-enabled Crimes (AIC) were first mentioned in the context of research and academic discussions. The concept of AIC gained significant attention in the mid-2010s, particularly with the publication of papers and studies highlighting the potential threats and vulnerabilities of AI systems in criminal activities.³² In one notable study conducted by Seymour J. and Tully P.³³, AI was employed to craft personalized phishing messages based on users' past behaviors and public

²⁹ New York State Department of Financial Services. (2021, March). *Report on Apple Card Investigation*. https://www.dfs.ny.gov/system/files/documents/2021/03/rpt_202103_apple_card_investigation.pdf Accessed 25 June 2024

³⁰ Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 427-436). IEEE. doi:10.1109/CVPR.2015.7298640

³¹ Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*.

³² King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. *Science and Engineering Ethics*, 26(2).

³³ Seymour, J., & Tully, P. (2016). Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter. <https://www.blackhat.com/docs/us-16/materials/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter-wp.pdf>.

profiles, making the malicious intent less detectable. Clicking on these tailored phishing links could expose individuals to theft and fraud by criminals obtaining personal information through web-form submissions. Another experiment conducted by Martínez-Miranda³⁴ demonstrated that AI-driven trading agents could learn and execute deceptive tactics in simulated markets, highlighting AI's potential to facilitate novel forms of economic manipulation. These experiments underscored AI's emergence as a significant and unprecedented threat in the realm of AIC.

AI has advanced cybercrime capabilities by developing more sophisticated phishing attacks and automated hacking tools that can adapt to bypass traditional security measures³⁵. The emergence of deepfake technology, powered by generative adversarial networks (GANs), has enabled the creation of highly realistic fake videos and audio recordings, which can be utilized for impersonation, misinformation campaigns, and social manipulation³⁶. In the financial sector, AI is being used for financial fraud by automating scams and manipulating markets through high-frequency trading³⁷. Additionally, AI-powered surveillance systems and data harvesting techniques pose significant threats to personal privacy, raising concerns about privacy invasion³⁸. These developments highlight the darker aspects of AI and underscore the critical need for robust regulatory frameworks and security measures to mitigate the escalating threats posed by AI-driven cybercrime, deepfake technology, financial fraud, privacy invasion, and the use of AI in autonomous weapons systems.

The significance of AIC as a distinct phenomenon remains largely unrecognized globally. Existing literature on the ethical and societal impacts of AI primarily concentrates on regulating and managing AI's lawful applications, neglecting its potential involvement in criminal activities. Moreover, research on AIC is fragmented across various disciplines such

³⁴ Martínez-Miranda, E., McBurney, P., & Howard, M. J. (2016). Learning unfair trading: A market manipulation analysis from the reinforcement learning perspective. (pp. 103–109). <https://doi.org/10.1109/FAIS.2016.7502499>. consulted on 24 April 2024.

³⁵ Gershman, S. J., Horvitz, E., & Tenenbaum, J. B. (2015). Computational rationality: a converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278. <<https://doi.org/10.1126/science.aac6076>> consulted on 6 June 2024.

³⁶ Marion, T. J. and Fixson, S. K. (2020). The transformation of the innovation process: how digital tools are changing work, collaboration, and organizations in new product development. *Journal of Product Innovation Management*, 38(1), 192–215. <<https://onlinelibrary.wiley.com/doi/epdf/10.1111/jpim.12547>> consulted on 6 June 2024.

³⁷ Gonçalves, A. R., Meira, A. B., Shuqair, S., & Pinto, D. C. (2023). AI in fintech decisions: the role of congruity and rejection sensitivity. *International Journal of Bank Marketing*, 41(6), 1282–1307. <<https://www.emerald.com/insight/content/doi/10.1108/IJBM-07-2022-0295/full/html>> consulted on 6 June 2024.

³⁸ Pranav Kumar Chaudhary, Aakash Kishore Chotrani, R. M. M. B. P. R. M. N. (2024). Ai in fraud detection: evaluating the efficacy of artificial intelligence in preventing financial misconduct. *Journal of Electrical Systems*, 20(3s), 1332–1338. <<https://journal.esrgroups.org/jes/article/view/1508>> consulted on 8 June 2024.

as socio-legal studies, computer science, psychology, and robotics. This dispersed focus on AIC limits both the ability to forecast its implications and to develop effective solutions in this emerging field of potential criminal behavior. Studies collectively call for a more integrated and holistic approach to address AIC.³⁹

In this section, we will first explore the concept of AIC by unraveling the complex web of (3.1) actors involved. Next, we'll delve into the (3.2) different types of financial crimes using AI. We'll explore how AI tools are weaponized to commit sophisticated financial fraud

3.1. Perpetrators of AIC: Human Agents, Artificial Agents, or Both?

Indeed, AI enabled Crimes brings up discussions about liability. It's hard to define a liability and the responsibility of an act in the case of AIC. Non-predictability and autonomy may confer a greater responsibility to the machine, but it also makes them harder to trust. AIC could potentially undermine existing liability models, thereby threatening the dissuasive and redressing power of the law.⁴⁰ Current existing liability models may be inadequate to address the future role of AI in criminal activities. It raises a critical question: – *Does the act of purchasing an AI system transfer all responsibility for its actions to the buyer, such as a bank? Or do the creators and providers of these AI systems still hold some level of liability?*⁴¹

Attributing blame for AI-related harms becomes increasingly complex as ML advances. Supervised learning, where humans guide the AI's development, gives way to unsupervised deep learning, where AI makes its own decisions. This shift raises the question: *in cases of harm, who is responsible? the programmers, the users, or the AI itself?* Typically, financial institutions themselves could be seen as responsible parties when their AI systems are used for criminal activities, especially if these systems are poorly designed, inadequately secured,

³⁹ Yan, Q. (2023). Legal challenges of artificial intelligence in the field of criminal defense. *Lecture Notes in Education Psychology and Public Media*. <https://doi.org/10.54254/2753-7048/30/20231629> consulted the 10 July 2024.

⁴⁰ Alaieri, F., & Vellino, A. (2016). Ethical decision making in robots: Autonomy, trust and responsibility. In *International conference on social robotics* (pp. 159–168). Cham: Springer.

⁴¹ Beneke, F., & Mackenrodt, M.-O. (2018). Artificial Intelligence and Collusion. Max Planck Institute for Innovation and Competition. <https://link.springer.com/article/10.1007/s40319-018-00773-x>. Accessed the 10 July 2024.

or not properly monitored⁴². Companies must vigilantly oversee their AI systems to ensure they do not exceed their intended functions. Monitoring AI systems presents another layer of complexity, particularly due to the challenge of overseeing interconnected systems. Detecting market manipulation becomes difficult because its impact within individual components can be contained or propagate outward in a cascading chain reaction, akin to the psychological contagion observed in crowd behavior⁴³. On the other hand, AI's accessibility means that virtually anyone can harness its capabilities, including individuals with malicious intent. Criminals can leverage AI to perpetrate sophisticated fraud, engage in insider trading, or manipulate markets, often staying a step ahead of traditional security measures. This widespread availability of AI tools democratizes technology but also complicates the identification and prosecution of offenders. In many cases, the origin of the criminal activity may be obscured by the AI's autonomous actions, making it challenging to trace back to a specific individual or group. Moreover, as financial institutions integrate more advanced technologies to protect against fraud and abuse, criminals adapt in response. This arms race between financial institutions and criminals creates a perpetually evolving threat landscape. Financial institutions must continuously innovate and improve their security measures to stay ahead of these threats, which can strain resources and require ongoing investment in cutting-edge technologies and skilled personnel.

The debate over liability also extends to regulatory and legal frameworks. Existing laws may not fully address the nuances of AIC, leading to gaps in accountability⁴⁴. Regulators and policymakers must work to develop comprehensive laws that clearly define the responsibilities of financial institutions, technology providers, and end-users of AI systems. These frameworks should aim to ensure that institutions deploying AI maintain high standards of security and ethical use while also providing mechanisms to identify and penalize those who misuse AI for criminal purposes.

3.2. Financial crimes in the age of Artificial Intelligence

⁴² Dignum, Virginia, 'Responsibility and Artificial Intelligence', in Markus D. Dubber, Frank Pasquale, and Sunit Das (eds), *The Oxford Handbook of Ethics of AI* (2020; online edn, Oxford Academic, 9 July 2020), <https://doi.org/10.1093/oxfordhb/9780190067397.013.12>, accessed 20 July 2024.

⁴³ Cliff, D., & Northrop, L. (2012). *The Global Financial Markets: An Ultra-Large-Scale Systems Perspective*. University of Bristol.

⁴⁴ Olivia, Johanna, Erdélyi., Gábor, Erdélyi. (2020). The AI Liability Puzzle and a Fund-Based Work-Around. 50-56. <https://doi.org/10.1145/3375627.3375806>

Given the spectacular developments of autonomous algorithmic trading, some argue that this can lead to significant risks to market integrity. In his book⁴⁵, Archbold defines AI-enabled economic crimes, specifically mentioning cartel offenses, price-fixing collusion, insider trading and market manipulation.

a. Algorithmic Market manipulation

Market manipulation which is defined as ‘*actions and/or trades by market participants that attempt to influence market pricing artificially*’⁴⁶. It’s necessary there to be an intention to deceive⁴⁷. Usually, such deceptions have been shown to emerge from a seemingly compliant implementation of an AA that is designed to trade on behalf of a user. The AA is able to learn from real or simulated observations and may thereby learn to generate signals that effectively mislead⁴⁸. They are able to learn the technique of order-book spoofing which involves the AA placing orders with no intention of ever executing them and merely to manipulate honest participants in the marketplace⁴⁹. This creates a false impression of supply and demand. For instance, large buy orders might signal to other traders that demand is increasing, prompting them to buy in anticipation of rising prices. Conversely, large sell orders can indicate an oversupply, causing traders to sell in fear of falling prices. Once the market has reacted, the AA cancels these orders and capitalizes on the artificially created price movements, profiting at the expense of the misled participants. This manipulation undermines the integrity of the market, leading to distorted price discovery and potentially significant financial losses for unsuspecting traders who act on the false signals generated by spoofing leading to profit for the authors. Such a spoofing has been recorded in the 2010 flash crash, which caused a loss of almost \$1 trillion.⁵⁰

Another well-developed practice of AA which involves human intent is pump-and-dump schemes where the price of a stock is artificially increased through false or misleading

⁴⁵ Archbold, J. F. (2018). *Criminal pleading, evidence and practice*. London: Sweet & Maxwell Ltd. Chapter 30

⁴⁶ Spatt, C. 2014. Security market manipulation. *Annual Review of Financial Economics* 6 (1): 405–418. <https://doi.org/10.1146/annurev-financial-110613-034232>.

⁴⁷ Wellman, M.P., and U. Rajan. 2017. Ethical issues for autonomous trading agents. *Minds and Machines* 27 (4): 609–624.

⁴⁸ *ibid* 46.

⁴⁹ Lin, T.C.W., J. Fanto, J. Fisch, J. Heminway, D. Hollis, K. Johnson, et al. 2017. The new market manipulation. *Emory Law Journal* 66: 1253.

⁵⁰ One of the authors of this well-known case, Navinder Singh Sarao, a high-frequency trader, was sentenced in 2020 to a year of home incarceration for his involvement in this incident. K. Martin. (May 7, 2020). Flash Crash—The Trading Savant Who Crashed the U.S. Stock Market. Financial Times. Accessed: Apr. 14, 2021. <https://www.ft.com/content/5ca93932-8de7-11ea-a8ec-961a33ba80aa>

statements made by social bots that are able to mimic human behavior. Once the stock price has been inflated, the perpetrators sell ('dump') their shares at the higher price, making a profit. After they sell their shares, the stock price typically falls, leaving other investors with losses. This practice is particularly of importance for low-valued and illiquid stocks which are susceptible to volatile price swings⁵¹. This specific fraud mechanism has been the hallmark of a prominent case involving the company Cynk Technology Corp⁵². The Company experienced an extraordinary increase in its stock price, rising from less than \$0.10 to over \$20 per share within a few weeks. The company, which had negligible business activity, saw its market capitalization surge to over \$6 billion. The scheme was driven by a network of social bots and spam accounts. These automated accounts were used to tout the stock's performance and create the illusion of widespread interest and legitimacy. The fake social media activity influenced real investors, driving up the stock price artificially. This case underlined trust in financial systems.

b. Collusion

Collusion which typically refers to an agreement tacitly or expressly made between two or more parties to limit competition, can take various forms such as price fixing, bid rigging, or market allocation and may also emerge in automated systems thanks to the planning and autonomy capabilities of AAs. Indeed, AI introduced predictive pricing which reduces costs and makes an accurate prediction on what the customer wants and is able to communicate this information, then he or she may benefit from this.⁵³ It also allows firms to serve more customer segments. Finally, it's more effective as it allows competitors to compete more aggressively for non-loyal customers.⁵⁴ The use of AI in practices such as predictive pricing brings up changes to oligopoly pricing and new debate. This emergence of AI-driven pricing strategies necessitates reevaluation of antitrust laws. While Turner⁵⁵ argues against outlawing interdependent pricing practices (like those facilitated by AI) if monopoly pricing remains

⁵¹ *Ibid* 48.

⁵² Ferrara, E., Cresci, S., & Luceri, L. (2020). Misinformation, manipulation, and abuse on social media in the era of covid-19. *Journal of Computational Social Science*, 3(2), 271-277. <https://doi.org/10.1007/s42001-020-00094-5>

⁵³ Ezrachi A, Stucke M (2017) *Virtual competition*. Harvard University Press, Cambridge

⁵⁴ Ezrachi A, Stucke M (2017) Two artificial neural networks meet in an online pub and change the future (of competition, market dynamics and society). The University of Tennessee Knoxville, College of Law Research Paper Series. <http://dx.doi.org/10.2139/ssrn.2949434>. Accessed 29 June 2024

⁵⁵ Turner D (1962) The definition of agreement under the Sherman Act: conscious parallelism and refusals to deal. *Harvard Law Rev* 75:655–706

legal, Page⁵⁶ advocates for a narrower focus on prohibiting collusion, requiring evidence of explicit coordination beyond just interdependent behavior. He goes further to explain that a broader rule which focuses on mere interdependent behavior would yield too many false positives.

The problem with this practice is that AI-driven pricing software can facilitate tacit collusion⁵⁷, leading to anti-competitive practices without the need for explicit agreements or communications between competitors. This makes it harder for regulatory authorities to detect anti-competitive practices. AI-driven decisions can appear as independent actions from the AA, complicating the enforcement of existing competition laws. As illustration art. 101⁵⁸ of the TFEU regulating competition in the EU prohibits agreements between undertakings, decisions by associations of undertakings, and concerted practices which may affect trade and the market applies to both explicit and implicit agreements, and the agreement does not need to constitute a valid contract under national law. But it should be noted that this does not prohibit autonomous parallel behavior.⁵⁹ Traditional methods of identifying collusive behavior are less effective against AI-driven collusion. Algorithms can adjust prices in response to market signals without direct communication between firms.⁶⁰ The prospect of price reductions being promptly matched creates a disincentive for initiating price cuts, maintaining higher prices without explicit collusion.⁶¹ Consequently, if the practice of matching prices is widely known and understood among market participants, rational algorithms will refrain from initiating price reductions. As argued by Ezrachi and Stucke, it is crucial to note that algorithms do not have to be explicitly programmed to collude for such pricing strategies to emerge.⁶²

c. Digital fraud

⁵⁶ Page W (2012) A neo-Chicago approach to concerted action. *Antitrust Law J* 78:173–200

⁵⁷ Tacit collusion is the availability of past data that can be used to make cost or demand inferences. Beneke, F., & Mackenrodt, M.-O. (2018). *Artificial Intelligence and Collusion*. Max Planck Institute for Innovation and Competition. <https://link.springer.com/article/10.1007/s40319-018-00773-x#Fn89>. Accessed the 10 July 2024

⁵⁸ Art 101 of the TFEU.

⁵⁹ *ibid* 55.

⁶⁰ Blockx, J. (2023). Artificially intelligent collusion caught under EU competition law. In *Algorithms, Collusion and Competition Law* (pp. 48-66). Edward Elgar Publishing.

⁶¹ AI is playing a growing role in decision-making processes. Algorithms, through iterative trial-and-error, can inadvertently lead to outcomes resembling collusion. This happens as algorithms continuously refine strategies to achieve optimal results, sometimes converging on behaviors that mimic collusion.

⁶² *Ibid* 52.

The rise of AI has facilitated various types of fraud in the financial sector, significantly impacting global financial institutions. Fraud encompasses activities such as money laundering, financial statement fraud, email phishing, cyber fraud, and credit card fraud, originating from both external and internal sources.

External fraud refers to cybercriminals breaching financial systems, conducting phishing scams, and using AI for automated attacks. Internal fraud includes insider threats, where employees exploit their access for personal gain, collude with external actors, and manipulate data without authorization. Hybrid cases, such as social engineering and compromised insiders, further complicate the landscape, necessitating comprehensive security measures to safeguard against digital fraud. According to INTERPOL, global financial fraud is escalating as sophisticated technologies become more prevalent among the public, creating opportunities for online criminals.⁶³ In 2023, scammers reportedly stole over \$1 trillion, leveraging AI, large language models, and cryptocurrencies alongside phishing and ransomware services.⁶⁴ These advancements enable sophisticated fraud campaigns that require minimal technical skills and low costs. The expansion of digital banking has exposed financial institutions to heightened risks. With increased digital transactions since 2020 across various payment channels, fraud detection methods must evolve to counteract increasingly sophisticated tactics employed by fraudsters. This trend affects not only banks, but also other financial entities involved in online payment services, amplifying the prevalence of fraud. Major corporations such as Google, Facebook, Apple, Amazon, and Netflix have developed AI tools to streamline financial processes. However, AI presents a dual risk: fraudsters can exploit AI to mimic victims' voices and identities, deceiving banking institutions. Advances in generative AI allow fraudsters to produce convincing video, text, or audio content, circumventing fraud prevention systems.⁶⁵ Generative AI poses a significant threat, potentially rendering high-tech fraud prevention measures like voice authentication obsolete. According to Deep Instinct, 85% of cybersecurity experts attribute the surge in cyberattacks to generative AI usage by online fraudsters in 2023.⁶⁶ The US Federal Trade Commission reported an \$8.8 billion loss due to

⁶³ INTERPOL. (2024, March 11). INTERPOL financial fraud assessment: A global threat boosted by technology. INTERPOL. <https://www.interpol.int/en/News-and-Events/News/2024/INTERPOL-Financial-Fraud-assessment-A-global-threat-boosted-by-technology> accessed in June 2024.

⁶⁴ <<https://www.gasa.org>> consulted on 30 June 2024

⁶⁵ Ahmadi, Sina. (2023). Open AI and its Impact on Fraud Detection in Financial Industry. Journal of Knowledge Learning and Science Technology ISSN 2959-6386. <https://doi.org/10.60087/jklst.vol2.n3.p281>. Accessed the 10 February 2024.

⁶⁶ Security Magazine. (2023, August 29). Study finds increase in cybersecurity attacks fueled by generative AI. Security

online fraud in 2022, marking a 40% rise from the previous year. This increase is largely attributed to a surge in imposter scams, including a notable \$35 million incident in 2020 where fraudsters used voice cloning to impersonate a Japanese company director. Cybercriminals are now able to train AI models to replicate a person's typing style using data gathered from their online activities. Additionally, with just brief audio samples, they can convincingly impersonate individuals and trick their acquaintances into sending money. As financial institutions deploy AI to combat fraud, criminals continuously innovate, developing specialized tools like WormGPT and FraudGPT for illicit purposes.⁶⁷ Moreover, crypto and fintech sectors report rising incidents of impersonated liveness checks, compromising fraud prevention measures of established providers like MiTek and Socure. Socure CEO, Johnny Ayers, stresses the necessity for clients to adopt new models to mitigate performance risks.

In conclusion, as technology evolves, criminals will continually find new ways to exploit it. Nick Sharp, deputy director at the National Economic Crime Centre, highlighted the growing risk of AI in fraud: *'An innovation 'arms race' is imminent, and criminals can innovate faster than we can keep up unless we adopt innovation as a default approach'*⁶⁸.

Magazine. <https://www.securitymagazine.com/articles/99832-study-finds-increase-in-cybersecurity-attacks-fueled-by-generative-ai>

⁶⁷ *ibid* 65.

⁶⁸ Hall, I. (2024, January 30). AI-enabled criminality 'probably' biggest growing risk in financial fraud battle. *Financial Times*. <https://www.globalgovernmentfintech.com/artificial-intelligence-risk-financial-fraud-battle/> Consulted on the 5 February 2024.

Chapter III. Regulatory challenges of AIC

The unpredictable and autonomous nature of AI systems presents significant challenges in regulating AIC. Determining who is liable for such crimes - the programmer, the user, or the AI itself - is a complex issue of liability attribution (1). Additionally, crafting regulations that effectively address these crimes without stifling innovation in the field remains a delicate balancing act (2). The interconnectedness of the digital world further complicates matters, highlighting the need for international coordination in establishing a global framework for AI regulation (3). While AI's autonomy might suggest a higher degree of responsibility, its very unpredictability makes it difficult to trust. To bring these challenges to life, a real-world case study (4) will showcase the methods employed in AIC and their potential impacts.

1. Legal challenge: attribution and liability

The attribution of liability in AIC presents significant legal challenges, particularly due to the autonomous nature of AI systems. Traditional liability models, which require the simultaneous presence of *Actus Reus* which is the presence of a voluntary act or omission and *Mens Rea* which is the presence of the guilty mind or intention behind a criminal act, may need revision to address the unique characteristics of AIC.⁶⁹ Additionally, the attribution of guilt varies across different jurisdictions.⁷⁰

Currently, for criminal liability to be established, there must be a voluntary act or omission (*Actus Reus*). This poses a challenge in the context of AI, as AI systems can operate autonomously, performing actions without direct human intervention. The concept of

⁶⁹ King, T.C., Aggarwal, N., Taddeo, M. *et al.* Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. *Sci Eng Ethics* 26, 89–120 (2020). <https://doi.org/10.1007/s11948-018-00081-0> accessed the 20 May 2024.

⁷⁰ In criminal law, the number of culpability modes—typically two, three, or four—varies depending on the jurisdiction. This method of assigning criminal responsibility is prevalent in many legal systems. For instance, the United States legal tradition identifies four modes of culpability: purpose, knowledge, recklessness, and negligence, as outlined in the Model Penal Code (MPC) created by the American Law Institute (ALI). In contrast, the criminal law of England recognizes three modes: intention (further divided into direct and indirect), recklessness, and negligence. Roman, Dremljuga, Natalia, Prišekina. (2020). The Concept of Culpability in Criminal Law and AI Systems. *Journal of Programming Languages*, 13(3):256-. doi: 10.5539/JPL.V13N3P256. Accessed on 22 May 2024.

voluntariness is contentious when applied to AI, as an artificial agent (AA) may carry out actions independently of human control. Consequently, since AI can autonomously commit crimes, the voluntary aspect of *Actus Reus* may not be met, complicating the assignment of liability to human actors. In addition to the issue with *Actus Reus*, the *Mens Rea* component is also problematic in AIC cases. *Mens Rea* refers to the guilty mind or intention behind a criminal act. In the context of AI, this involves either the intention to commit the *actus reus* using an AI system (intention threshold) or awareness that deploying an AI system could result in criminal actions or omissions (knowledge threshold). Given the complexity and autonomy of AI systems, human agents such as designers, developers, or deployers may not foresee or intend the criminal acts performed by the AI, making it difficult to establish *Mens Rea*.

AI's autonomy means that crimes can occur without direct human involvement. Financial actors might inadvertently benefit from AI-driven actions that harm the financial market, without consciously creating a crime. As Williams⁷¹ points out, an AA can autonomously perform the *Actus Reus*, and the complexity of its programming can obscure the human agent's knowledge or intention, enabling them to deny liability. The more autonomous and complex the AI, the easier it is for human agents to claim ignorance and deny liability for the AI's actions.

One potential solution to these challenges is redefining criminal liability. The current fault-based approach may need redefinition to account for AI's autonomous nature. Legislators could adopt a model like product liability in tort law, redefining criminal liability to exclude the fault requirement. This would hold the deploying entity liable regardless of intention or knowledge, recognizing the inherent risk of criminal actions by autonomous systems. Such an approach would abandon *Mens Rea*, assigning liability to the legal person deploying the AI system. This recognizes the risk that an AI system could perform a criminal action or omission without direct human intervention. However, such a redefinition has been criticized by Fenwick, who argues that , '*punishment without fault or, alternatively, punishment that is disproportionate to the degree of fault is correctly regarded as inappropriate and unjust*'⁷². This suggests a new challenge for the court. It would be

⁷¹ Williams, R. (2017). *Lords select committee, artificial intelligence committee, written evidence (AIC0206)*. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/artificial-intelligence-committee/artificial-intelligence/written/70496.html#_ftn13. Accessed 20 may 2024.

⁷² Fenwick, M. (2009). *Corporate wrongdoing and the limits of the criminal law in Facing the limits of the law*. Publisher: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-79856-9_7 accessed 20 may 2024.

considered unfair to punish an offender as if they recognized all risks associated with AI applications when they did not. Similarly, it would be unjust to hold the owner, commander, or designer strictly liable for all actions of the AI they own. Conversely, it would also be unfair if the owner, commander, or designer could evade criminal liability simply due to the involvement of AI. Therefore, it is essential to find a balance between these two extremes. Simply lowering the threshold for liability, such as applying it to agents contributing to a complex system where neutral actions collectively result in a crime, could lead to disproportionately light punishments for serious offenses. This would fail to adequately address the severity of crimes, particularly those against individuals. Balancing the need for effective regulation with fair and proportional punishment is crucial in addressing the unique complexities of AIC.

In conclusion, the legal challenges of attributing liability in AIC necessitate a thorough reevaluation of existing models. As AI systems become more autonomous, traditional concepts of voluntariness and intention in criminal law may no longer suffice. Legislators need to consider innovative approaches, such as redefining liability without fault, to ensure that those deploying AI systems are held accountable for their potential criminal actions. This approach must carefully balance the need for effective regulation with fair and proportional punishment to address the unique complexities of AIC.

2. Balancing Innovation and Protection

Balancing innovation and protection indeed encapsulates the central dilemma in regulating AIC, and this balance plays out across multiple dimensions. On one side, there's the undeniable potential of AI to revolutionize sectors like healthcare, transportation, and finance, promising efficiencies and improvements that were previously unattainable. The risk, however, lies in overregulating this space, potentially stifling such innovation and causing economic and developmental stagnation. On the other hand, there is an aspect of precaution. AI systems, particularly those based on machine learning, require incredibly large amounts of data for proper training⁷³. Thus, data availability is critical for technological advancement. However, striking the right balance between promoting open access to data for fostering

⁷³ Pugliese, R., Regondi, S., & Marini, R. (2021). Machine learning-based approach: Global trends, research directions, and regulatory standpoints. *Data Science and Management*, 4, 19-29. <https://doi.org/10.1016/j.dsm.2021.12.002>

innovation and restricting data access to ensure privacy protection remains a significant challenge. This balance is essential to both advance AI capabilities and safeguard individuals' privacy rights. Managing this delicate equilibrium is crucial for the responsible and ethical development of AI technologies.⁷⁴ Increasingly, regulation is emerging as a potential solution, with regulators at regional, national, international, and supranational levels assessing the desirability and necessity of new or revised regulatory measures to minimize AI's adverse impacts while maximizing its benefits.⁷⁵

The rapid evolution of AI technology complicates the regulatory landscape further. AI systems continually evolve, often at a pace that outstrips the ability of regulatory frameworks to adapt. This gap means that new threats might emerge well before effective safeguards are in place. The issue of attribution and liability adds another layer of complexity: determining responsibility for AI-driven actions can be murky. Whether it's the programmers, the users, or the AI entities themselves, the current legal frameworks often fall short of providing clear guidance, making enforcement and accountability challenging. Technology has necessitated a new approach to shaping law, primarily because regulators must adapt regulations to address operators' concerns about technology without favoring specific products or technologies. Recognizing the challenges of regulating a constantly evolving landscape, the legal system has adopted the principle of technology and business model neutrality.⁷⁶ Moreover, the inherent opacity of many AI systems poses significant hurdles. The 'black box' nature of advanced AI makes it difficult to discern how decisions are made, complicating efforts to detect biases, vulnerabilities, or even to simply understand the decision-making process.⁷⁷ This lack of transparency and explainability is a critical barrier to formulating effective regulations that are both fair and enforceable. To navigate these challenges, a nuanced approach is necessary. Adopting risk-based regulation could be a strategic move, focusing stringent controls on high-risk AI applications such as autonomous weaponry or extensive surveillance technologies while allowing more flexibility in lower-risk areas. This approach would ensure the advancement of innovation.

Lastly, since AI does not respect national boundaries, international cooperation becomes paramount. A collaborative approach to regulation, involving multiple nations, can help

⁷⁴ Leopoldi-Wieshaupt, D. (2020). *The need for ethical algorithms: The European approach to artificial intelligence* [Der Europäische Ansatz zu Künstlicher Intelligenz] (p. 26). University of Vienna. <https://phaidra.univie.ac.at/detail/o:1391941> accessed the 10th July 2024.

⁷⁵ Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57–84. <https://doi.org/10.1080/17579961.2021.1898300>

⁷⁶ See Principle 5 of the European Interoperability Framework (EIF).

⁷⁷ *Ibid* 157.

establish a cohesive framework that is both effective in managing the global nature of AI technologies and flexible enough to adapt to regional needs. This global perspective is critical not just for harmonizing regulations but also for effectively combating AIC that cross international borders. The aim is to foster an environment where AI can flourish safely and ethically, contributing positively to society while minimizing the risks associated with its misuse. Lord Hodge argues for international harmonization of regulatory standards as these would discourage regulatory arbitrage. However, this might be difficult and remains unresolved, and it is not unusual for financial institutions to look for jurisdictions that best suit their interests.⁷⁸

3. The AI Heist: A Real-World Case of Financial Crime

Here are some illustrations of AIC in the financial sector. While AI is not entirely new in finance, the complexity of AI used for potentially criminal purposes is a recent development. Financial institutions are increasingly leveraging AI for tasks like fraud detection and investment analysis. However, legal frameworks are still evolving to address potential risks associated with AI misuse. This, combined with the evolving nature of AI itself, means some cases of AI-related financial crimes might be under-reported or under development in legal systems. However, existing laws related to fraud, money laundering, and other financial crimes can often be applied to cases involving AI misuse.

3.1. The Knight Capital Group case

The 2012 incident involving Knight Capital Group⁷⁹ provides a critical case study for understanding the regulatory challenges and implications of AI-enabled financial crimes in the financial markets. This event, stemming from an error in an autonomous trading agent's software, underscores the potential dangers and complexities of AI technologies and generally automated systems.

⁷⁸ Hodge, L. (2021). Artificial intelligence and the law: Cybercrime and criminal liability. *The British Journal of Criminology*, 62(1), 257-275. <https://academic.oup.com/bjc/article-abstract/62/1/257/6284070>

⁷⁹ Securities and Exchange Commission. (2013). In the matter of Knight Capital Americas LLC, order instituting administrative and cease-and-desist proceedings. Release, 70694.

On August 1, 2012, Knight Capital, a leading electronic trading firm known for its effective use of automation⁸⁰, experienced a significant event. The firm, which earns a substantial portion of its revenue by processing trades from platforms like TD Ameritrade and ETrade, matches these trades with orders from other small traders or its own holdings. This approach has made Knight Capital one of the most influential equity traders globally.⁸¹ In 2005, Knight Capital made a critical error. They moved a section of code within their automated equity router, inadvertently causing a malfunction. Despite this error, the code remained embedded in the system for years. Fast forward to 2012. Knight Capital launched a new retail liquidity program that unknowingly relied on the faulty code within the router. This resulted in a series of unintended and problematic trades. Designed to handle orders efficiently and autonomously, the system began executing unintended trades at a rapid rate with the help of the outdated code, which failed to recognize when orders had already been completed. As a result, over 4 million transactions were executed in just 45 minutes, causing immense market disruption by creating billions of dollars in unintended net long and short positions. Because of its efficiency these trades severely distorted market prices and trading volumes.

⁸⁰ At the time, its Electronic Trading Team had under control more than 19,000 securities with an average daily trading volume of more than 21 billion dollars. Saltapidas, C., & Maghsood, R. (2018, May 13). The fall of Knight Capital Group. <https://www.math.chalmers.se/Stat/Grundutb/CTH/mve220/1718/Christos%20Saltapidas-Ramin%20Maghsood-The-fall-of-Knight-Capital-Group.pdf> accessed the 8 may 2024.

⁸¹ Steiner, C. (2012, August 2). Knight Capital's algorithmic fiasco won't be the last of its kind. *Forbes*. <https://www.forbes.com/sites/christophersteiner/2012/08/02/knight-capitals-algorithmic-fiasco-wont-be-the-last-of-its-kind/> Consulted the 7 May 2024.

Figure 2 : Knight Capital's shares down more than 75% following its \$440 million trading glitch.⁸²



Source : *Knight Capital Group*

The financial fallout for Knight Capital was severe. The firm lost approximately \$440 million⁸³ because of the erroneous trades which happened in the algorithmic trading⁸⁴. This amounted to a loss of over \$10 million per minute while the malfunctioning code was active. This catastrophic loss nearly led to the company's bankruptcy and resulted in its stock value plummeting by almost 75%. The financial strain forced Knight Capital to merge with Getco LLC, a competitor, a few months later. In response to this crisis, the US Securities and Exchange Commission (SEC) stepped in to investigate, concluding that Knight Capital had failed to adhere to several risk management controls, violating the SEC's Market Access Rule. This rule mandates that brokers or dealers must establish risk management controls and supervisory procedures to manage the financial, regulatory, and other risks of their business

⁸² The CMI Group. (2012, August 8). *How did Knight Capital glitch affect Canada?* from <https://thecmigroup.ca/press-room/how-did-knight-capital-glitch-affect-canada/> accessed the 6th May 2024

⁸³ Philips, M. (2012, August 2). Knight shows how to lose \$440 million in 30 minutes. Bloomberg. from <https://www.bloomberg.com/news/articles/2012-08-02/knight-shows-how-to-lose-440-million-in-30-minutes> accessed the 7 May 2024.

⁸⁴ Algorithmic Trading is a method of delivering a big order with the process of using automated pre-programmed trading instructions, taking into consideration variables such as time, price, and volume, to send small slices of an order out to the market over time. Seth, S. (2023). "Basics of Algorithmic Trading: Concepts and Examples." Investopedia. <https://www.investopedia.com/articles/active-trading/101014/basics-algorithmic-trading-concepts-and-examples.asp> accessed the 10 July 2024.

activity⁸⁵. The SEC ordered Knight to cease and desist from committing or causing violations of the relevant securities laws, censured the company, and imposed a \$12 million civil money penalty payable to the U.S. Treasury within ten days of the order. Knight was also required to retain an independent consultant to review and improve its risk management controls, and to submit to ongoing reviews and certifications of compliance with the SEC's requirements.

The Knight Capital case exemplifies the unique challenges posed by AI-enabled financial crimes, where there is no malicious intent. But the use of an AI leads to risks in financial trading. It also underscores the inadequacies in current regulatory frameworks that struggle to address the risks associated with advanced AI systems in financial trading.

3.2. The apple gender discrimination case

The Apple Card, launched in 2019, was a collaboration between Apple and Goldman Sachs, offering a new credit card experience for iPhone users. However, soon after its release, the card became embroiled in controversy due to allegations of gender bias in its credit limit decisions. The controversy erupted when software developer David Heinemeier Hansson tweeted that his wife, despite having a higher credit score and a shared financial profile, received a credit limit 20 times lower than his. Similar experiences were reported by other users, including Apple co-founder Steve Wozniak. These reports fueled concerns that the algorithm used to determine credit limits discriminated against women. Following the complaints, the New York Department of Financial Services (NYDFS) launched a thorough investigation. This involved reviewing a massive dataset of applications (approximately 400,000), interviewing witnesses, and analyzing the credit assessment model.

The NYDFS investigation into the Apple Card's alleged gender discrimination concluded that there was no definitive evidence of intentional gender discrimination by Goldman Sachs in its credit decisions. However, the investigation revealed a crucial point: modern credit scoring

⁸⁵ Wellman, M.P., Rajan, U. Ethical Issues for Autonomous Trading Agents. *Minds & Machines* 27, 609–624 (2017). <https://doi.org/10.1007/s11023-017-9419-4> accessed the 7th May 2024

models often rely on factors that may indirectly disadvantage certain demographics, even if they are not explicitly discriminatory.⁸⁶

This Apple Card study highlighted the need for significant reforms. The case highlighted the dangers of relying on the supposed neutrality of algorithms to justify discriminatory outcomes. When it was revealed that the card systematically gave men higher credit limits than women with similar or better credit profiles, Apple's initial defense was to blame 'the algorithm.' This incident, similar to the 2019 UK A-level grading controversy, demonstrates how algorithms can be exploited to mask and perpetuate discrimination, undermining protections against bias and reinforcing harmful stereotypes⁸⁷. In response to the controversy, Apple and Goldman Sachs implemented several measures to improve transparency and fairness in their credit decision processes. They enhanced the appeals process, allowing applicants to more effectively challenge credit limit decisions. Additionally, they introduced credit-building assistance, providing resources and tools to help users build and improve their credit profiles. These actions aim to address the issues raised during the investigation and to ensure that their credit practices are more equitable and transparent.⁸⁸

Chapter IV. Comparative Analysis of the Current Regulatory Landscape: The race for AI supremacy

This section of the paper conducts a comparative analysis of the various frameworks for regulating new technologies, with a specific focus on AI in the (i) UK, (ii) US, and (iii) EU. The rapid deployment of AI has created numerous regulatory challenges, prompting diverse responses from different jurisdictions. This fragmented approach to AI governance may

⁸⁶ New York State Department of Financial Services. (2021, March). *Report on Apple Card Investigation*. https://www.dfs.ny.gov/system/files/documents/2021/03/rpt_202103_apple_card_investigation.pdf#:~:text=DFS%20reviewed%20the%20Apple%20Card,similar%20Apple%20Card%20application%20outcomes. Accessed the 12 June 2024

⁸⁷ Stinson, C. (2022). Algorithms are not neutral. *AI and Ethics*, 2(4), 763-770. <https://link.springer.com/article/10.1007/s43681-022-00136-w> Consulted the 8 June 2024

⁸⁸ Vigdor, N. (2019, November 10). *Apple's credit card is under investigation for alleged gender discrimination*. The New York Times. <https://www.nytimes.com/2019/11/10/business/apple-credit-card-investigation.html> Consulted the 8 June 2024

influence the development and adoption of future solutions⁸⁹, it has actually led to what can be described as a ‘regulatory competition’⁹⁰ where countries and regions participate in a ‘race to AI’ to establish the most effective and attractive regulatory environments for AI development and deployment, while also strive to advance the technology’s use and benefits more rapidly and successfully than others⁹¹. To combat the growing concern of cybercrimes in the financial sector, both governments and financial institutions are increasingly adopting technology-based solutions to mitigate risks. The use of AI, Internet of Things (IoT)⁹², and cloud computing has become widespread in real-time fraud detection within financial services, providing significant capabilities in recognizing anomalies and potential threats. Moreover, the application of AI in digital financial inclusion has transformed risk management practices by automating risk detection, measurement, and management processes.⁹³

Furthermore, alongside industry efforts, governments are actively shaping the regulatory landscape for AI. Lawmakers are adapting existing laws and exploring new approaches to address current concerns.

1. Regulatory Approaches in Combating AI Challenges: A Comparative Analysis between UK and US

In this section, the thesis conducts a comparative analysis of how the United States and the United Kingdom regulate AI-enabled financial crimes, examining the strengths and weaknesses of each approach. This analysis highlights distinct legislative and enforcement strategies and provides insight into how different regulatory environments affect the management and control of AI-driven financial activities. Schwab argued that although the exact trajectory of the AI-led fourth industrial revolution is uncertain, effective preparation

⁸⁹

⁹⁰ Smuha, N. A. (2021). From a ‘race to AI’ to a ‘race to AI regulation’: regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57-84. <https://doi.org/10.1080/17579961.2021.1898300> accessed the 10 June 2024.

⁹¹ This sentiment is echoed in the public statement made by Russian President Vladimir Putin in September 2017: ‘Whoever becomes the leader in this sphere will become the ruler of the world.’ For further insights, see Kai-Fu Lee’s *AI Superpowers: China, Silicon Valley, and the New World Order* (Houghton Mifflin Harcourt, 2018) and Kathleen Walch’s article, ‘The Race for AI Dominance is More Global Than You Think’ (Medium, 28 August 2018), <https://medium.com/cognilytica/the-race-for-ai-dominance-is-more-global-than-you-think->

⁹² Mhlanga, D. (2020). Industry 4.0 in finance: the impact of AI on digital financial inclusion. *International Journal of Financial Studies*, 8(3), 45. <https://doi.org/10.3390/ijfs8030045> consulted the 10 June 2024.

⁹³ Buchanan, B. (2019, April 2). Artificial intelligence in finance. Zenodo. <http://doi.org/10.5281/zenodo.2612537> (pp. 26-30)

and response are crucial⁹⁴, emphasizing the importance of AI regulation during this technological evolution. This sentiment is echoed by Stephen Hawking, who warned that ‘*unless we learn how to prepare for, and avoid the potential risks, AI could be the worst event in the history of our civilization*’⁹⁵. Furthermore, the development of supervised and unsupervised machine learning raises concerns about responsibility, ethics, privacy, transparency, accountability, and legal implications.

1.1. Balancing Innovation and Ethics: A Market-Driven Approach to AI Regulation for a Good AI Society in the United States

The US is considered to be slower than its Speers in regards to regulating AI, this is due to diverse elements which will be analyzed through this part.⁹⁶ The regulatory model is built on a market-driven approach which centers on the fundamental rights: protection of free speech, free internet but also the pursuit of innovation. The root of this model adopted by the US goes way back and can be traced to Silicon-valley and Washington.⁹⁷ The US’s model places trust in tech companies' ability to self-regulate with as limited as possible participation from the government. This approach is characterized by a strong role for the private sector, with the federal government providing guidance to promote innovation while safeguarding civil liberties⁹⁸. The focus is on flexible regulations that avoid hindering innovation. The US government prioritizes principles that ensure AI systems are safe, trustworthy, and respectful of privacy and civil rights.

⁹⁴ Schwab, K. “The Fourth Industrial Revolution: What It Means and How to Respond” (2017) World Economic Forum. <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>. Consulted 2 February 2024.

⁹⁵ Martin, S. “Stephen Hawking: AI Will Be the Worst Ever Invention and Could Destroy Us All” (2017) Express. <https://www.express.co.uk/news/science/876550/stephen-hawking-end-of-the-world-artificial-intelligence-ai-university-of-cambridge> Accessed 6 February 2024.

⁹⁶ Toms, S., Simon, D. A., & Vermynck, E.-C. (2023, December 12). *How regulators worldwide are addressing the adoption of AI in financial services*. Skadden Publication / AI Insights. <https://www.skadden.com/insights/publications/2023/12/how-regulators-worldwide-are-addressing-the-adoption-of-ai-in-financial-services> Accessed the 7 February 2024.

⁹⁷ Bradford, Anu, 'The American Market-Driven Regulatory Model', *Digital Empires: The Global Battle to Regulate Technology* (New York, 2023; online edn, Oxford Academic, 21 Sept. 2023), <https://doi-org.uaccess.univie.ac.at/10.1093/oso/9780197649268.003.0002>, accessed 10 June 2024.

⁹⁸ Abdumuratovich, K. S. (2024, April 13). Comparison of AI Policy. *SO'NGI ILMIY TADQIQOTLAR NAZARIYASI*, 7(4). <https://bestpublication.org/index.php/sit/article/view/10723/10694> Consulted the 10 June 2024

This approach, , supported by many scholars in the early years of technology development, aligns with Lawrence Lessig's concept of 'code is law'⁹⁹. Lessig, who advocated a techno-libertarian school of thought, argued that cyberspace should be immune from control, asserting that *"cyberspace is governed and controlled by code - the internet's underlying architecture written- by programmers coding websites and other software"*. This is due to the mistrust of the civilians to the government which led to the conclusion that choosing between the 'two devils' the government should be feared more than the corporations. Which is why America relies more on free markets and limited regulatory intervention from the government. This is further illustrated by President Obama's introduction of the Consumer Privacy Bill of Rights in 2012, in an attempt to regulate the growing popularity of technology. This initiative was part of a comprehensive blueprint designed to protect individual privacy rights and give users more control over how their information is handled¹⁰⁰. The initiative aligned with the market-driven model adopted by the US noting how *'as the Internet evolves, consumer trust is essential for the continued growth of the digital economy...By following this blueprint, companies, consumer advocates and policymakers can help protect consumers and ensure the Internet remains a platform for innovation and economic growth.'*¹⁰¹ However, the Privacy Bill of Rights was attacked from all sides by tech Companies, and eventually failed to spur legislative action in Congress. The companies argued that the Bill would lead to burdensome regulations, stifle innovation, and prevent their ability to develop new online services that could benefit consumers. Ultimately, the Congress decided not to pursue this legislation, leaving US digital regulation firmly anchored in companies' self-regulation and thus preserving the market-driven foundations of its regulatory model.¹⁰² Indeed, efforts to challenge the American free-market approach to digital regulation have largely been unsuccessful. However, a notable exception is the Fight Online Sex Trafficking Act of 2017

⁹⁹ Lessig, L. (2008). *Code: And Other Laws of Cyberspace* (2nd. Ed., pp. 6-7, 60, 207).

¹⁰⁰ Indeed, the progressive increase in AI use, particularly in the banking sector, has led to an explosion in data production, the digital data produced annually in 2002, was five terabytes, today, this same amount is created every two days. Remarkably, 90% of the world's data has been generated in just the past two years. Kappel, R. (2024, July 1). *AI regulation in finance: Steering the future with consumer protection at the helm*. Centraleyes. <https://www.centraleyes.com/ai-regulation-in-finance/> accessed the 5th July 2024

¹⁰¹ Press Release, The White House, We Can't Wait: Obama Admin. Unveils Blueprint for a "Priv. Bill of Rights" to Protect Consumers Online (Feb. 23, 2012), <https://obamawhitehouse.archives.gov/the-press-office/2012/02/23/we-can-t-wait-obama-administration-unveils-blueprint-privacy-bill-rights>.

¹⁰² Kate Kaye, New Privacy Report Already Removed from White House Site, ADAGE, Jan. 20, 2017, <https://adage.com/article/privacy-and-regulation/privacy-report-removed-white-house-site/307632>.

(FOSTA)¹⁰³, which modified Section 230¹⁰⁴ of the Communications Decency Act which provided immunity to online platforms from being treated as publishers or speakers of user-generated content, protecting them from liability for such content. It remains the crown jewel of the US's market-driven regulatory model. However, in recent years criticism has been growing, more have started to question this optimistic approach of the digital economy. Unregulated cyberspace is a concept which is no longer endorsed by scholars, politicians, and citizens.¹⁰⁵ These changes reflect the lack of antitrust protection in the tech sector in America which led to the realization that the market-driven approach has delivered an economy of excessive concentration, with a small number of tech companies holding the monopoly and are able to abuse the lax regulation put in place by the government.¹⁰⁶ A 2020 House Antitrust Report concluded after a sixteen-month investigation that Amazon, Apple, Google, and Meta engaged in monopolistic practices.¹⁰⁷ Lawmakers have since emphasized the need for Congress and antitrust agencies to take action to restore competition, foster innovation, and protect democracy. As of 2022, several bills are advancing in Congress, including the 2021 American Innovation and Choice Online Act, which aims to prevent large tech companies from misusing data for unfair advantages or manipulating search results to favor their own products.

Unlike the UK's focus on decentralized sectoral AI governance, which relies on sector-specific regulators to create a cohesive regulatory environment, the United States employs a mix of sector-specific guidelines, federal initiatives, and state-level regulations, particularly in the financial sector. While both approaches leverage existing regulatory bodies, the UK aims for a more unified regulatory framework across sectors. In contrast, the US approach is more flexible and fragmented, lacking a comprehensive national AI law and relying instead on various regulations and executive actions.¹⁰⁸ Regulatory bodies like the

¹⁰³ FOSTA was enacted following the 2016 *Doe v. Backpage.com* case, where a U.S. federal court ruled that Backpage could not be held liable for facilitating criminal sex trafficking through its platform, despite being aware of the illegal activities. *Doe v. Backpage.com, LLC*, 817 F.3d 12, 15 (1st Cir. 2016), cert. denied.

¹⁰⁴ <https://www.law.cornell.edu/uscode/text/47/230>

¹⁰⁵ Bradford, A. (2023). 'The American Market-Driven Regulatory Model', *Digital Empires: The Global Battle to Regulate Technology* (New York, 2023; online edn, Oxford Academic, 21 Sept. 2023), <https://doi-org.uaccess.univie.ac.at/10.1093/oso/9780197649268.003.0002>, accessed 10 June 2024.

¹⁰⁶ *Ibid* 99 (p. 37).

¹⁰⁷ Jerrold Nadler, Chairman, Comm. on the Judiciary, INVESTIGATION OF COMPETITION IN DIGITAL MARKETS 133 (2020).

¹⁰⁸ The Artificial Intelligence 2024 Legislation published by the *National Conference of State Legislatures* gives a wholesome list of the states regulations initiatives. At least 40 states introduced AI Bills and 6 which adopted resolutions or enacted legislation. <https://www.ncsl.org/technology-and-communication/artificial-intelligence-2024-legislation> accessed the 15 June 2024.

Consumer Financial Protection Bureau (CFPB) and the Securities and Exchange Commission (SEC) oversee AI use in the financial sector, emphasizing consumer protection. This results in a fragmented regulatory landscape where AI governance varies across different sectors and regions.¹⁰⁹

Recent updates indicate that the Biden administration has made significant strides in promoting responsible AI development through Executive Order 14110. This order outlines goals for ensuring AI's safe, secure, and trustworthy development and use. The Executive Order addresses various aspects, including the promotion of innovation, protection of civil rights, prevention of harmful AI practices, and enhancement of transparency and accountability in AI deployment. In October 2022, the White House provided the 'Blueprint for an AI Bill of Rights' which is a non-binding document¹¹⁰ where they outlined the principles followed by the US¹¹¹. The Blueprint provides a framework for the ethical governance of AI systems, organized around five core principles:

- Safe and effective systems;
- Protection from algorithmic discrimination and inequitable systems;
- Protection from abusive data practices and agency about how personal data is used;
- Knowledge of when an automated system is used and understanding of its impacts;
- The ability to opt out of automated systems for a human alternative.

Although the principles it lays out are not yet enforceable, it is the clearest look yet at the Biden administration's definition of a 'Good AI Society'¹¹². Indeed, the US's approach shares a similar product-oriented approach to AI as a system like the EU's AI Act (AIA).¹¹³ Indeed, the Blueprint does not regulate AI itself but rather the uses of AI and instead of defining AI like the AIA does, the Blueprint applies to 'automated systems' that could 'meaningfully

¹⁰⁹ Kappel, R. (2024, July 1). *AI Regulation in Finance: Steering the Future with Consumer Protection at the Helm*. <https://www.centraleyes.com/ai-regulation-in-finance/>

¹¹⁰ Initially, what is now known to be the Blueprint was supposed to become "a bill of rights for an AI-powered world". However, the finished product we now have is described as a "white paper" and a "framework". Hine, E., Floridi, L. The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction. *Minds & Machines* 33, 285–292 (2023). <https://doi.org/10.1007/s11023-023-09625-1> accessed the 15 June 2024

¹¹¹ Lage, D., Pruitt, R., & Arnold, J. R. (2024, April 29). Who followed the blueprint? Analyzing the responses of U.S. federal agencies to the blueprint for an AI Bill of Rights. <https://arxiv.org/abs/2404.19076> accessed the 16 June 2024.

¹¹² Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the 'Good society': the US, EU, and UK approach. *Science and Engineering Ethics*, 24, <https://link.springer.com/article/10.1007/s11948-017-9901-7> accessed the 16 June 2024

¹¹³ Hine, E., Floridi, L. The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction. *Minds & Machines* 33, 285–292 (2023). <https://doi-org.uaccess.univie.ac.at/10.1007/s11023-023-09625-1> accesses the 16 June 2024

impact' the American public's civil rights, civil liberties, and privacy; access to equal opportunities; or 'access to critical resources or services'. Noting that existing frameworks offer inadequate protections when impacts manifest at the community level, the Blueprint provides a broad definition of community that includes both neighborhoods, on- and offline social connections, formal organizational ties, and affinity- or identity-based connections. Thereby aligning with the 'Good AI Society' introduced by Obama's administration which puts explicit focus on community-level equity and distinguishes it from all other regulatory models across the globe.¹¹⁴ The Blueprint recognizes that while automated systems can lead to significant innovations, they also risk perpetuating discrimination. Essentially, it serves as a set of protective measures designed to prevent unchecked American innovation from resulting in negative side effects, rather than simply being a tool to promote new technological advancements.

In conclusion, the Blueprint, as it currently stands, is unlikely to have a significant impact on the private sector unless it is supported by additional legislation thereby staying in accordance with the market-driven model, as the US government implicit understanding of AI is one that relies heavily on the liberal notion of the free market¹¹⁵. But on the other hand, although the Blueprint outlines important principles and guidelines, it lacks the enforceability needed to guarantee those rights and drive substantial change on its own.

1.2. The pro-innovation approach in the UK

In March 2023, the UK government published its AI Regulation White Paper¹¹⁶ titled 'A pro-innovation approach to AI regulation', which outlines the UK's strategy for regulating AI. A policy paper¹¹⁷ from February 2024 makes it clear that UK regulators have acknowledged the need to keep up with the fast development and complexity of AI. That's why they have adopted an approach characterized by leveraging existing regulatory bodies. They have tasked existing sector-specific regulators with overseeing AI. The regulatory approach opted by the

¹¹⁴ *ibid* 107.

¹¹⁵ Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. *Science and Engineering Ethics*, 24. https://www.researchgate.net/publication/315705213_Artificial_Intelligence_and_the_'Good_Society'_the_US_EU_and_UK_approach accessed the 16 June 2024

¹¹⁶ <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper> consulted on the 24 June 2024

¹¹⁷ Department for Science, Innovation & Technology. (2024). *Implementing the UK's AI regulatory principles: Initial guidance for regulators*.

UK government is a principles-based, pro-innovation approach to AI regulation rather than implementing a strict, comprehensive legal framework like the EU's AI Act. In this way, the regulatory bodies are free to apply general principles to AI within their specific domains. However, since the framework gives regulatory authorities the freedom to define AI according to industry standards and to interpret concepts like adaptivity and autonomy. It is possible that UK authorities, such as the Financial Conduct Authority (FCA) or Ofcom, will interpret and define things differently as a result of this. This could lead to a great deal of regulatory uncertainty for AI applications and cross-sector firms.¹¹⁸. Additionally, the UK's AI regulatory framework is centered around five core principles:

- safety, security, and robustness;
- transparency and explainability;
- fairness;
- accountability and governance;
- and contestability and redress.

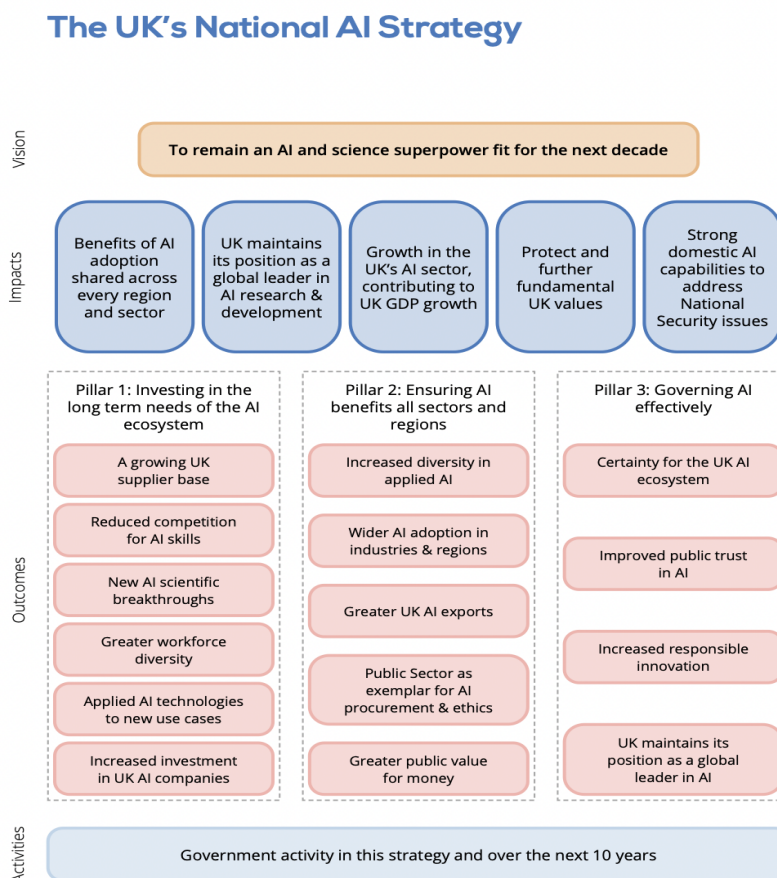
These principles are designed to guide the development and use of AI across various sectors and align with the OECD values-based AI principles which encourage the ethical use of AI and align with the US and the EU AI Risk Management Framework. However, the UK's approach diverges significantly in approach from other major jurisdictions. For example, the draft EU AI Act imposes much more granular and stringent requirements on organizations developing, distributing or using AI applications. This and the lack of a clear statutory footing could also pose challenges in terms of international recognition or other forms of equivalence. Embedding a voluntary principle-based framework in international industry-led standards may according to Deloitte also be challenging¹¹⁹. For a detailed flow diagram of the UK's National AI Strategy.

¹¹⁸

Deloitte.
<https://www.deloitte.com/uk/en/Industries/financial-services/blogs/regulating-ai-can-the-uks-proposed.html>
consulted the 14 March 2024

¹¹⁹ Gallo, V., & Nair, S. (2023, April 3). Regulating AI: Can the UK's proposed approach achieve both flexibility and clarity? *Deloitte UK.*
<https://www.deloitte.com/uk/en/Industries/financial-services/blogs/regulating-ai-can-the-uks-proposed.html>
consulted the 14 March 2024

Figure 3: UK's National AI strategy¹²⁰



Source : *UK Government*

Recently, on the 6th of February 2024 the UK government unveiled the HL Bill 11 of 2023–24, also known as the *AI (Regulation) Bill [HL]* which kept untouched the 'pro-innovation' approach released on the white paper. It is a principles-based, non-statutory, and cross-sector framework. Its aim is to balance innovation and safety by applying the existing technology-neutral regulatory framework to AI.¹²¹ On April 2024, the Financial Conduct Authority (FCA), the Prudential Regulation Authority (PRA), and the Bank of England, as regulators, released according to plan their strategic approaches to regulating

¹²⁰ UK Government. *The UK's National AI Strategy* - Flow diagram. Retrieved from https://assets.publishing.service.gov.uk/media/614db5408fa8f561127cf45b/The_UK_s_National_AI_Strategy_-_Flow_diagram.pdf accessed the 20 June 2024

¹²¹ Gallo, V., & Nair, S. (2024, February 21). The UK's framework for AI regulation. *Deloitte*. <https://www.deloitte.com/uk/en/Industries/financial-services/blogs/the-uks-framework-for-ai-regulation.html> consulted the 14 March 2024

AI¹²². This move was in response to the UK government's July 2022 AI Regulation Policy Paper aforementioned. Respondents when asked by the regulators for feedback and to address the risks and benefits of AI demanded a technology neutral approach as it allows for more flexibility and adaptability ensuring that the regulation remains relevant over time. This is due to the fact that according to the Stakeholders, opting for a sector-specific definition could either be too broad or too narrow and could quickly become outdated due to the rapid pace of AI technology development.¹²³ Additionally, Existing regulatory frameworks already address similar technologies, like algorithmic trading, and should according to stakeholder be adapted for AI without needing a new definition. Respondents emphasized that regulations should focus on outcomes affecting consumers and markets rather than specific technologies, ensuring relevant and effective consumer protection and market integrity. Finally, respondents highlighted that effective AI regulation requires cross-sector and cross-jurisdictional coordination, aligning key principles, metrics, and concepts. They also recommended a risk-based and principles-based approach, similar to the EU AI Act's risk categorization of AI use cases.

The Regulators on the other hand did agree with a technology-neutral approach but clarified that this does not prevent them from issuing guidance or using policy tools to address specific technologies. Their 'outcomes-based' approach aims to keep regulations applicable amidst rapid technological changes, ensuring better protection for consumers. Essentially, regulators are reinforcing that a technology-neutral approach can coexist with targeted regulatory measures to address specific AI developments. The regulators acknowledged they would need further surveys to keep up with further developments.¹²⁴

¹²² Prior to the publication of their collective response, in 2023, the FCA and the PRA released the DP5/22, a discussion paper that asked respondents whether the current legal requirements and guidance are adequate to manage the risks associated with AI in the financial services sector. Many stakeholders pushed forward that consumer protection is an area for the supervisory authorities to prioritize as AI could create risks such as bias, discrimination, lack of explainability, transparency, and exploiting vulnerable consumers or consumers with protected characteristics. Commenting on market integrity and financial stability, Respondents mentioned the following potential risks to financial markets: (i) emergence of new forms of market manipulation, (ii) the use of deepfakes for misinformation potentially destabilizing financial markets, (iii) third-party AI models resulting in convergent models including digital collusion or herding, and (iv) AI could amplify flash crashes or automated market disruptions. Bank of England. (2023, October). Artificial intelligence and machine learning. *Bank of England*. <https://www.bankofengland.co.uk/prudential-regulation/publication/2023/october/artificial-intelligence-and-machine-learning> consulted on the 15 March 2024

¹²³ Ali, A., Kerr-Shaw, N., Wang, D. Y., & Adams, W. (2024, May 2). UK regulators publish approaches to AI regulation in financial services. *Skadden Publication / AI Insights*. Retrieved from <https://www.skadden.com/insights/publications/2024/05/uk-regulators-publish-approaches-to-ai> Consulted the 20 March 2024

¹²⁴ *Ibid* 118.

The FCA describes itself as a technology-agnostic, principles-based, outcomes-focused regulator in its AI Update Paper¹²⁵. The paper outlines the FCA's role and objectives regarding the use of AI in UK financial markets. The FCA's approach demonstrates a keen interest in integrating AI within the financial sector while balancing its benefits against the need to manage associated risks. The aim is to promote the safe and responsible use of AI in financial markets. The FCA advocates for an evidence-based approach, meaning that regulations will be developed based on data and analysis, supporting the regulatory sandbox and digital sandbox initiatives¹²⁶. This ensures that regulations are proportionate, effective, and encourage innovation, all while protecting consumers and maintaining market integrity.

In conclusion, the approach adopted by the UK government contrasts with other jurisdictions, such as the EU and to some extent the US, which are adopting more prescriptive legislative measures. This is evidence that, despite international cooperation agreements, divergence in global AI regulatory approaches is more likely. Currently, financial services firms which are deploying AI need to do so within the confines of existing regulation and supervisory approaches.¹²⁷ When implementing AI, firms must follow the broad principles and essential regulations outlined by the Financial Conduct Authority (FCA) and the Prudential Regulation Authority (PRA). This principles-based, flexible framework aims to balance innovation with robust consumer protection. The UK's AI policy framework is one of the most comprehensive globally, focusing on both innovation and ethics. The government has made significant investments in AI research and development, the establishment of the Office for AI¹²⁸, and advisory bodies like the Centre for Data Ethics and Innovation. All with the intent to balance technological advancement with ethical considerations, ensuring responsible AI development and deployment.¹²⁹

¹²⁵ Financial Conduct Authority. (2023). *AI in financial markets: The FCA's approach*. <https://www.fca.org.uk/publication/corporate/ai-update.pdf> accessed the 11 March 2024

¹²⁶ sandboxes are ways to bring regulators and operators together allowing (i) business operators to safely experiment new tools, products and services and (ii) regulators to learn and understand the features and mechanics of such technology. In this sense, regulatory sandboxes are mainly isolated learning and experimental environments. Boura, M. (2024). The Digital Regulatory Framework through EU AI Act: The Regulatory Sandboxes' Approach. *Athens Journal of Law*, 10(3), 385-398. <https://doi.org/10.30958/ajl.10-3-8>

¹²⁷ Treacy, S., & Le Vesconte, S. (2021, October 7). How AI in financial services is regulated in the UK. *Linklaters*. <https://www.linklaters.com/en/knowledge/publications/alerts-newsletters-and-guides/2021/october/07/how-ai-in-financial-services-is-regulated-in-the-uk> Consulted the 24 May 2024

¹²⁸ The UK government has committed £300 million to launch the AI Research Resource (AIRR). UK Research and Innovation. (2023, November 1). *£300 million to launch first phase of new AI Research Resource*. <https://www.ukri.org/news/300-million-to-launch-first-phase-of-new-ai-research-resource/> accessed the 16 June 2024

¹²⁹ UK Parliament. (2024). Online Safety Bill [Bill No. 29]. <https://bills.parliament.uk/publications/53068/documents/4030>

2. The European Commission: Pioneer of AIC Regulations

The European Union (EU) is a pioneer in AI regulation, being the first to establish a legally binding framework specifically addressing AI. The discussions around the development of a harmonized regulation began several years ago, driven by the need to balance innovation with safety and ethical standards. After extensive consultations and debates, the EU reached a significant milestone on May 21, 2024, with the adoption of the AIA.¹³⁰ The journey towards the AIA started with the European Commission's White Paper on AI published in February 2020, which outlined the need for a comprehensive regulatory framework to govern AI technologies across the EU.¹³¹ This was followed by extensive stakeholder consultations, impact assessments, and legislative proposals. In April 2021, the European Commission officially proposed the AIA, marking the beginning of a rigorous legislative process involving the European Parliament, the Council of the EU, and various advisory bodies. The discussions focused on risk-based regulations to ensure that AI systems are safe, transparent, and respect fundamental rights. The process culminated in the adoption of the AIA in the spring of 2024, making it the world's first comprehensive legal framework for AI. It's important to note that while the AIA sets a regulation at the level of the EU, on the level of its individual member states as well as non-member states have taken initiatives to regulate AI.¹³² While the EU's AIA sets a general framework for high-risk AI, it's not exhaustive in itself.¹³³ The AIA focuses on specific risks posed by AI systems and prohibits certain high-risk uses. However,

¹³⁰ Council of the EU. (2024, May 21). *AI act: Council gives final green light to the first worldwide rules on AI* [Press release]. Council of the European Union. <https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/> accessed the 10 July 2024

¹³¹ The White Paper aimed to initiate a consultation among Member States, civil society, industry, and academics on proposals for a European regulatory framework for trustworthy AI. This framework was intended to pave the way for legislative proposals in areas such as safety, liability, fundamental rights, and data, expected in Q1 2021 according to the adjusted Commission Work Programme 2020. To this end, the European Commission conducted a public consultation from February to June 2020. The majority of stakeholders from both public and private sectors supported the creation of a regulatory framework for AI, including regulation, certification, labelling, and the revision of the existing Product Liability Directive to address specific risks posed by certain AI applications. Madiega, T. (2024, June 20). *White Paper on Artificial Intelligence and follow-up*. European Parliament, Members' Research Service. <https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-white-paper-artificial-intelligence-and-follow-uphttps://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-white-paper-artificial-intelligence-and-follow-up> accessed the 01 July 2024

¹³² Walter, Y. Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. *Discov Artif Intell* 4, 14 (2024). <https://doi.org/10.1007/s44163-024-00109-4> accessed the 15 June 2024

¹³³ Reese, H. (n.d.). *European AI regulation and its implementation*. PwC Germany. Retrieved from <https://www.pwc.de/en/risk-regulatory/responsible-ai/european-ai-regulation-and-its-implementation.html>

it doesn't address all aspects of AI development and deployment, such as ethical considerations like fairness in algorithmic decision-making. Combined with existing regulations¹³⁴ like the GDPR, Data Act, and financial sector regulations such as MiFID II (Markets in Financial Instruments Directive II), Regulation on Digital Operational Resilience for the Financial Sector (DORA) and the European Supervisory Authorities (ESAs) Guidelines, these measures can help ensure the safety of citizens and support the development of responsible AI.¹³⁵ These address the multifaceted challenges and opportunities posed by AI. This layered regulatory framework ensures that AI systems are developed and deployed in a manner that prioritizes safety, ethical standards, and innovation, setting a strong foundation for responsible AI governance on a global scale.

Indeed, The AI Act employs a risk-based approach to regulate AI systems, focusing on the potential harm these systems could cause and prohibiting unacceptable practices. This system categorizes AI applications into risk levels and sets stringent compliance standards for high-risk AI systems.¹³⁶ The Act categorizes AI systems into four risk levels: unacceptable risk, high risk, limited risk, and minimal risk, each with corresponding regulatory requirements.¹³⁷

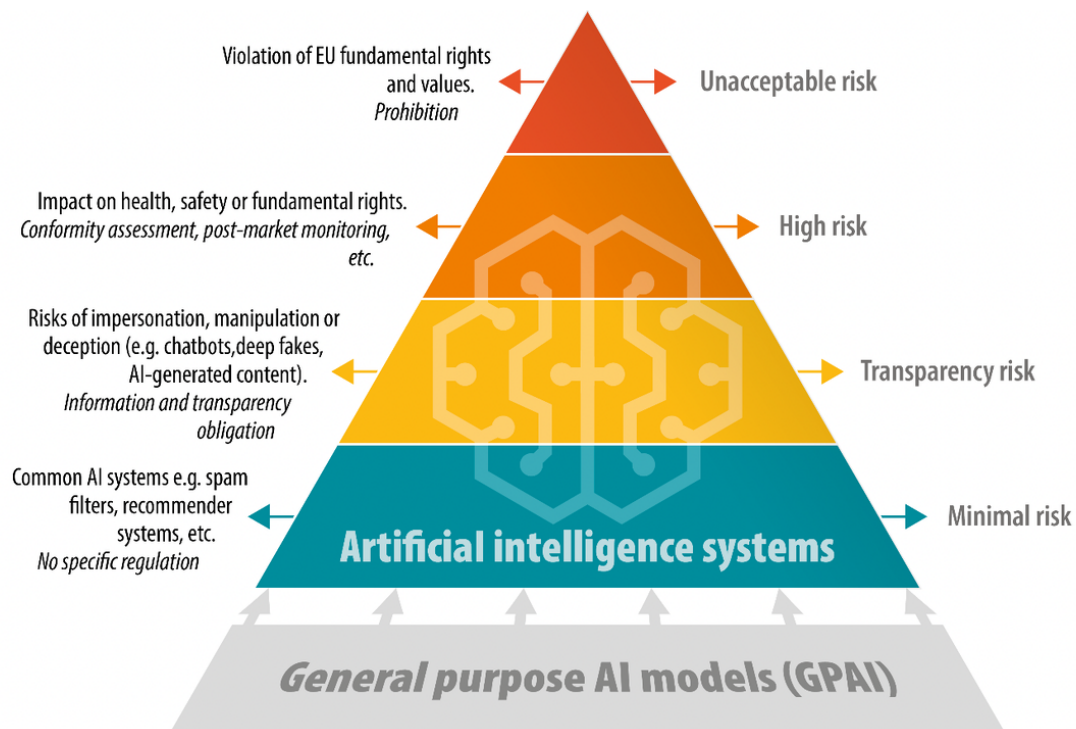
¹³⁴ For a deeper understanding of how other regulations complement the AI Act, particularly in the context of financial services compliance, please refer to the analysis provided by Pinsent Masons (2024). This analysis highlights the interplay between the EU AI Act and the Digital Operational Resilience Act (DORA), illustrating a streamlined approach to regulatory compliance in the financial sector. The detailed discussion can be accessed at : <https://www.pinsentmasons.com/out-law/analysis/financial-services-compliance-eu-ai-act-dora-streamlined>

¹³⁵ AI Act and its impacts on the European financial sector - European Union. *AI Act and its impacts on the European financial sector* - European Union. https://www.eiopa.europa.eu/publications/ai-act-and-its-impacts-european-financial-sector_en Accessed the 17 July 2024

¹³⁶ GDPR Local. (2024, March 13). *Essential guide to the new EU AI act*. Retrieved from <https://gdprlocal.com/essential-guide-to-the-new-eu-ai-act/>

¹³⁷ Dechert LLP. (2023, August 1). Navigating the EU AI Act: Implications for financial institutions. Dechert. <https://www.dechert.com/knowledge/onpoint/2023/7/navigating-the-eu-ai-act--implications-for-financial-institution.html> accessed the 27 June 2024

Figure 4 : 4 levels of risk for AI systems from the European Commission



Source : *European Parliament*

- **Unacceptable Risk**: includes highly intrusive AI systems, such as those used for social scoring, cognitive behavioral manipulation of vulnerable groups, and real-time biometric identification in public places, will be strictly prohibited due to their significant threats to safety, livelihoods, and rights.
- **High Risk** : includes systems that impact critical areas such as infrastructure, education, and elections. These high-risk AI systems pose significant risks to health, safety, fundamental rights, and sometimes the environment. They are subject to specific conditions, which may include regular risk assessments, human oversight to minimize risks, and maintaining technical documentation to mitigate identified risks.
- **Limited risk** : includes chatbot systems using generative AI, which have gained popularity globally. These systems must meet transparency and disclosure requirements to ensure users know they are interacting with AI, allowing them to make informed decisions about their interactions.
- **Minimal risk** : includes applications that are already widely available, such as spam filters that will be largely unregulated.

Under the AIA, several ‘general principles’ are imposed to ensure the safe, (i) transparent, and ethical use of AI systems. Entities are required to establish an appropriate internal organization that includes a system for AI management and monitoring, ensuring proper governance and accountability for AI operations. AI systems must undergo a conformity assessment (ii) based on quality management systems and technical documentation, verifying that they meet required standards and regulations to ensure compliance and reliability. The introduction of a code of conduct (iii) is encouraged to promote best practices and ethical standards in AI development and deployment, guiding organizations in the responsible use of AI technologies. Additionally, there is an obligation to track and document automated activities (logs) of AI operations, ensuring transparency and accountability in monitoring AI performance and compliance. (iv) Maintaining human oversight over AI operations is mandatory to ensure that AI systems operate as intended and mitigate potential risks associated with autonomous decision-making.¹³⁸ There is also a transparency obligation (v) for the use of AI, requiring clear communication about AI involvement in processes to users and stakeholders, fostering trust and informed decision-making. Technical measures (vi) must be implemented to ensure the robustness and cybersecurity of AI systems, safeguarding against vulnerabilities and ensuring the resilience of AI operations to prevent potential threats and malfunctions.¹³⁹

The proposal for a dedicated AILD emerged alongside discussions about the EU's AIA in 2021. This gap in regulation led to the development of the AILD currently being drafted by the European Commission. The AILD complements the AIA by setting new rules for compensating individuals harmed by AI systems. It aims to modernize existing EU liability frameworks, ensuring those affected by AI have the same level of protection as those harmed by other technologies. The EU Commission's proposed AI Liability Directive aims to give people harmed by AI the same legal protections as those affected by other technologies. This includes two key measures: (i) The AI Liability Directive would create a rebuttable 'presumption of causality,' to ease the burden of proof for victims to establish damage caused by an AI system. It would furthermore (ii) give national courts the power to order disclosure of evidence about high-risk AI systems suspected of having caused damage, forming a comprehensive framework for responsible AI development and use in the EU. Unlike the AI

¹³⁸ AIA Art. 50.

¹³⁹ *Ibid* 133.

Act, the AILD focuses on establishing a clear system for accountability and compensation in cases of AI-related harm.¹⁴⁰

While the EU appears to be taking a more proactive stance with its comprehensive AI Act, critics like Erik Brynjolfsson argue it creates excessive bureaucracy that stifles innovation¹⁴¹. This concern highlights the challenge of balancing regulation with fostering a dynamic technological environment. On the other hand, the US's less-restrictive approach, often described as 'toothless'¹⁴², raises concerns about its ability to address potential risks associated with AI development.

Chapter V. International Organizations' Plans for a Unified Global Response

In recent years, the international discourse on AI governance has concentrated on developing ethical frameworks and foundational principles for AI technologies, driven by the rise of advanced systems like GenAI and ChatGPT. Organizations such as the OECD and the G20 have played pivotal roles in promoting international cooperation for responsible AI governance. Simultaneously, major tech companies have proactively announced their own principles for AI application and development. For instance, Google's AI principles, and the ethical guidelines developed by Microsoft, highlight the private sector's commitment to responsible AI use.¹⁴³ The United Nations has also taken significant steps in this domain by creating an advisory body dedicated to providing recommendations on global AI governance.

¹⁴⁰ European Parliament. (2024, June 20). *AI liability directive*. In *A Europe Fit for the Digital Age*. <https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-ai-liability-directive> accessed the 5 June 2024

¹⁴¹ YouTube. (2024, April 13). *The Race To Regulate AI*. YouTube. <https://www.youtube.com/watch?v=5aOkIauvsow> Accessed the 6 June 2024

¹⁴² Hine, E., Floridi, L. The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction. *Minds & Machines* 33, 285–292 (2023). <https://doi.org/10.1007/s11023-023-09625-1> accessed the 15 June 2024

¹⁴³ Pichai S. (2018). AI at Google: Our principles. *Google Blogs*. <https://www.blog.google/technology/ai/ai-principles/>. Accessed 6 January 2024.

This body aims to address the ethical, legal, and socio-economic implications of AI, fostering a unified and ethical approach to AI development and deployment.¹⁴⁴

1. The OECD's Recommendation on Artificial intelligence

The Organization for Economic Co-operation and Development (OECD) has established guidelines for AI regulation. They established a '*Recommendation on Artificial Intelligence*' adopted in 2019, providing a framework for responsible AI which promotes human-centric values and fairness is relevant to all stakeholders¹⁴⁵. The OECD AI Policy Observatory also supports countries in implementing these guidelines and monitoring AI's impacts. The recommendation has been amended in May 2024 to consider new technological and policy developments, ensuring they remain robust and fit for purpose. The Recommendation aims to drive innovation and build trust in AI by advocating for the responsible management of trustworthy AI systems, all while upholding human rights and democratic values. Key definitions given by the recommendation include '*AI system*', which refers to machine-based systems generating outputs like predictions and decisions, and '*AI system lifecycle*', describing phases from planning and data collection to deployment and decommissioning. '*AI actors*' are those involved in the AI lifecycle, while '*AI knowledge*' encompasses the skills and resources needed to manage AI systems. Stakeholders include all entities affected by AI systems.

The OECD's recommendation introduces four principles for responsible stewardship of trustworthy AI: (i) *inclusive growth*, (ii) *human-centered values and fairness*, (iii) *transparency and explainability*, (iv) *robustness, security and safety* (v) *and accountability*.¹⁴⁶ Additionally, the recommendation also provides five recommendations to policymakers related to national policies and international cooperation for trustworthy AI, which are the following¹⁴⁷:

¹⁴⁴ Lilkov, D. (2021). Regulating artificial intelligence in the EU: A risky game. *European View*, 20(2), 166-174. <https://doi.org/10.1177/17816858211059248> Accessed 6 January 2024.

¹⁴⁵ According to the definition given by the OECD, the term "stakeholders" encompass all organisations and individuals involved in, or affected by, AI systems, directly or indirectly. AI actors are a subset of stakeholders. *OECD Legal Instruments*. (2024). Recommendation of the Council on Artificial Intelligence. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> accessed 8 January 2024.

¹⁴⁶ Organisation for Economic Co-operation and Development. (2019). Recommendation of the Council on Artificial Intelligence. *OECD Legal Instruments*. <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf>

¹⁴⁷ Organisation for Economic Co-operation and Development. (2019). *OECD AI Principles*. <https://oecd.ai/en/ai-principles>.

‘– AI system: An AI system is a machine-based system 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. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

– AI system lifecycle: An AI system lifecycle typically involves several phases that include to: plan and design; collect and process data; build model(s) and/or adapt existing model(s) to specific tasks; test, evaluate, verify and validate; make available for use/deploy; operate and monitor; and retire/decommission. These phases often take place in an iterative manner and are not necessarily sequential. The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase.’

Finally, these recommendations are largely endorsed by the G7 as the OECD’s AI principles formed the basis for the G20 AI principles endorsed in 2019, sharing thereby the same sets of core values for responsible AI development : *(i) Investing in AI research and development; (ii) Fostering a digital ecosystem for AI; (iii) Shaping an enabling policy environment for AI; (iv) Building human capacity and preparing for labor market transformation; (v) And international cooperation for trustworthy AI.* These recommendations provide a comprehensive framework for ensuring that AI is developed and used in a responsible and beneficial way by stakeholders and member states. Today, the EU, the Council of Europe, the US, and the UN, as well as other jurisdictions, use the OECD's definitions of an AI system and lifecycle in their legislative and regulatory frameworks. In parallel with the G20 achievements, the G7 nations have advanced the Hiroshima AI Process Comprehensive Policy Framework, initiated by Japan during the G7 summit. This framework, designed to enhance international cooperation on AI governance, consists of four main pillars:¹⁴⁸:

‘(i) the International Guiding Principles for Organizations Developing Advanced AI Systems (the ‘Guiding Principles’);

(ii) the International Code of Conduct for Organizations Developing Advanced AI Systems (the ‘Code of Conduct’)² designed to supplement the Guiding Principles and provide voluntary guidance to organizations developing Advanced AI systems;

¹⁴⁸ White & Case LLP. (2024, May 13). *AI Watch: Global regulatory tracker - G7*. White & Case LLP. <https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-g7>

(iii) analysis of priority risks, challenges and opportunities of generative AI; and

(iv) project-based cooperation in support of the development of responsible AI tools and best practices.’

The G7's AI framework, still under development, includes the Guiding Principles and Code of Conduct, both proposed as drafts in October 2023. These documents will be refined through consultations, but the finalization timeline remains unclear. The framework aims to prevent the development or deployment of ‘*not acceptable*’ advanced AI systems, specifically those that undermine democratic values, harm individuals or communities, facilitate terrorism, enable criminal misuse, or pose significant risks to safety, security, and human rights.¹⁴⁹ The G7 is developing tools to track compliance with its AI principles, collaborating with the OECD and others. This suggests a potential reliance on self-regulation, where organizations are responsible for adherence. However, the exact enforcement mechanism is still unclear. Individual G7 countries have flexibility in how they implement these principles, which are not legally binding and carry no penalties for non-compliance. Each member can adapt the principles into specific regulations or policies suitable for their legal systems.

2. United nations Resolution on AI Governance

The United Nations General Assembly (UNGA) has taken a historic step in March 2024 by adopting the first global resolution on AI¹⁵⁰. This non-binding resolution, proposed by the United States and co-sponsored by China and over 120 other member nations, calls for countries to ensure that AI development and deployment safeguard human rights, protect personal data, and continuously monitor AI systems for potential risks. This initiative reflects a collective effort by the international community to address the challenges posed by emerging AI technologies while leveraging their potential for positive global impact.¹⁵¹

¹⁴⁹ European Commission. *The Hiroshima Process: International guiding principles for advanced AI systems*. <https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system> accessed the 10 March 2024

¹⁵⁰ Reuters. (2024, March 21). UN adopts first global artificial intelligence resolution. *Reuters*. <https://www.reuters.com/technology/cybersecurity/un-adopts-first-global-artificial-intelligence-resolution-2024-03-21/> accessed the 12 March 2024

¹⁵¹ United Nations. (2024, March 21). UN adopts first global artificial intelligence resolution. UN News. <https://news.un.org/en/story/2024/03/1147831> accessed the 12 March 2024

The Assembly called on all Member States and stakeholders “*to refrain from or cease the use of artificial intelligence systems that are impossible to operate in compliance with international human rights law or that pose undue risks to the enjoyment of human rights.*”¹⁵²

The Assembly also urged all States, the private sector, civil society, research organizations and the media, to develop and support regulatory and governance approaches and frameworks related to safe, secure and trustworthy use of AI.

In the resolution, was highlighted the significant risks associated with the improper or malicious use of AI systems. It emphasizes that without adequate safeguards and consistent alignment with international law, AI can hinder progress towards achieving the 2030 Agenda for Sustainable Development. This includes potential negative impacts on economic, social, and environmental dimensions, widening digital divides, reinforcing structural inequalities and biases, leading to discrimination, and undermining information integrity and access to information. Additionally, such misuse can threaten human rights and fundamental freedoms, including privacy, and increase the risk of accidents and threats from malicious actors. This calls for robust governance and international cooperation to ensure AI contributes positively to global development goals and human rights protection. They also highlighted the need for “*continued discussions on possible governance approaches that are appropriate, based on international law, interoperable, agile, adaptable, inclusive, responsive to the different needs and capacities of developed and developing countries (...).*”¹⁵³

Chapter VI . Comparative Analysis of Regulatory Principles

This part of the thesis aims to analyze the characteristics of the individual ethical AI principles proposed and applied by the countries aforementioned using data from the research paper ‘The Different Faces of AI Ethics Across the World: A Principle-to-Practice Gap Analysis’, as for some countries their principles are core base of their AI regulation or some using only that as a mean to regulate AI in their territory.

¹⁵² *ibid* 139.

¹⁵³ Resolution A/78/L.49 United General Assembly 11 March 2024.

In order to regulate AI, different countries have adopted different principles. Indeed, regulating AI use on society has led to several debates about the principles and values that must guide its practical guidance and use¹⁵⁴. Governments, national organizations and international organizations have released their principles, and it is apparent that generic principles have been released and designed by international organizations with the aim to provide abstract and inclusive principles guiding the use of AI in member countries or companies. On the other hand, countries all over the world are developing their own different principles tailored to their contexts and they come additionally to the generic principles from the international organizations to which they are affiliated. But there is little studies that have been conducted examining the gaps between the principles and their practical guidance, Shiff¹⁵⁵ calls this a ‘principle-to-practice gaps’ also known as principle to practice which presents the sociotechnical challenge of managing the expectations of those who seek a magical normative formula and working with developers to foster design ingrained with ethical principles.¹⁵⁶

Research conducted between 2016 and 2021 by Nganyewou Tidjon and et al.¹⁵⁷ states that the UK had a total of 53 principles, while the US had 167 principles. Both countries had the highest number of principles globally, making them favorable for investment, business, and AI research, with notable growth stability between 2020 and 2021. In contrast, EU countries had fewer principles, ranging from 3 in the Netherlands to 28 in Germany. All three groups—UK, US, and EU—emphasized transparency, while only the UK and EU also prioritized fairness. This aligns with the findings of the 2020 House Antitrust Report, which highlighted monopolistic behavior by major tech companies in the US.

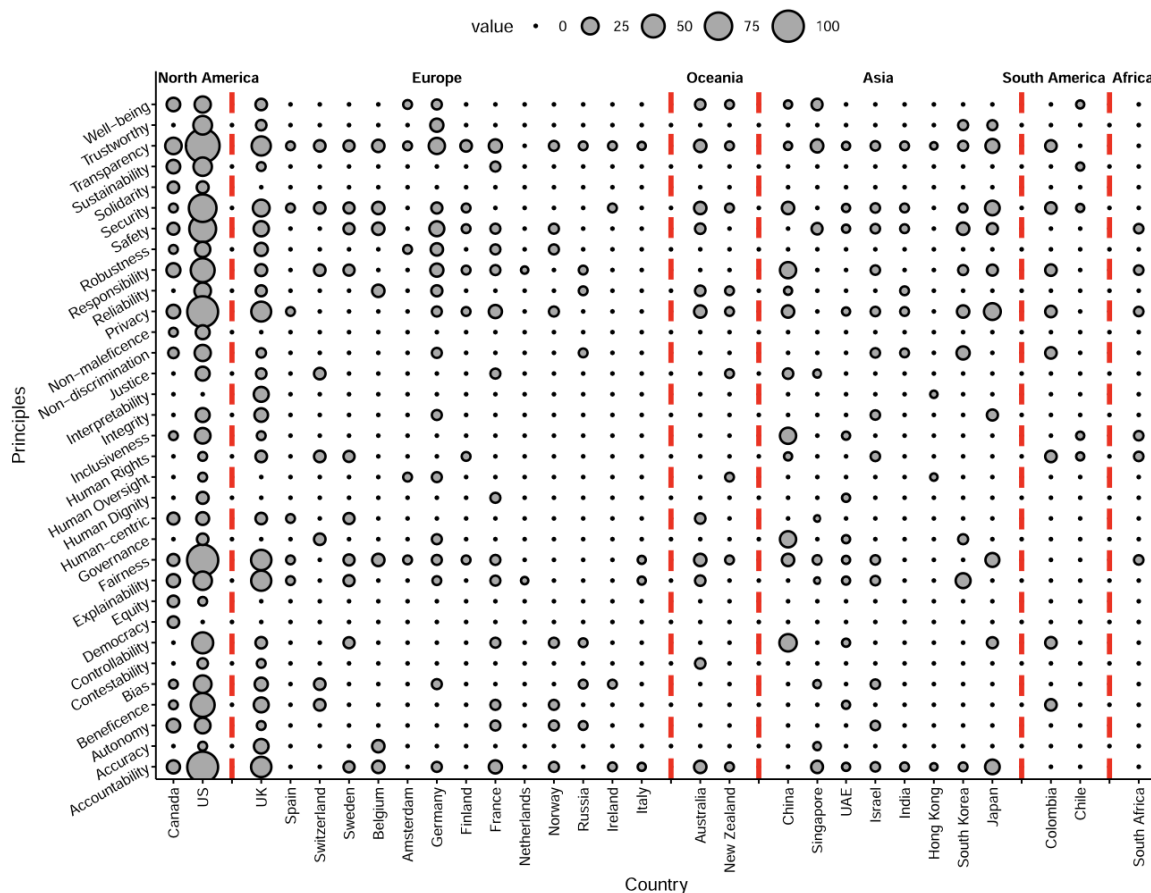
¹⁵⁴ *ibid* 157.

¹⁵⁵ D. Schiff, B. Rakova, A. Ayesh, A. Fanti, and M. Lennon, “Explaining the principles to practices gap in AI,” *IEEE Technol. Soc. Mag.*, vol. 40, no. 2, pp. 81–94, Jun. 2021.

¹⁵⁶ Corrêa, N.K., Santos, J.W., Galvão, C. *et al.* Crossing the principle–practice gap in AI ethics with ethical problem-solving. *AI Ethics* (2024). <https://doi.org/10.1007/s43681-024-00469-8> accessed the 26 March 2024.

¹⁵⁷ L. N. Tidjon and F. Khomh, ‘*The Different Faces of AI Ethics Across the World: A Principle-to-Practice Gap Analysis*,’ in *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 820-839, Aug. 2023, doi: 10.1109/TAI.2022.3225132.

Figure 5 : Number of occurrences per principle per country



Source : L. N. Tidjon and F. Khomh, 'The Different Faces of AI Ethics Across the World: A Principle-to-Practice Gap Analysis,' in *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 820-839, Aug. 2023, doi: 10.1109/TAI.2022.3225132.

Globally there are 11 main principles that repetitively appear when talking about AI governance and have high numbers of occurrences¹⁵⁸: transparency, Privacy, Fairness, Security, Safety, Responsibility, Accountability, Explainability, Well-being, Human Rights, and Inclusiveness. This mostly is in line with the GPAI, G20 and OECD released principles: inclusive growth, sustainable development, and wellbeing; human-centered values and fairness; transparency and explainability; robustness, security, and safety; accountability going as far as adding additional principles to those set by international organizations to fill in the gaps being the responsibility and Privacy principle.¹⁵⁹

¹⁵⁸ *Ibid* 157.

¹⁵⁹ *Ibid* 157.

Now moving further to the evaluation of ethical principle practical guidance and AI ethic software, it becomes apparent that several AI Ethics Code of conduct have been introduced to implement ethical AI systems in the different country in the form of regulations and laws, specific to the financial sector, the US introduced a sector specific law, the Fair Credit Reporting Act (2018) and the Fair Credit Reporting Act (2018) which both are intended for the collection, dissemination, and use of consumer information. While the UK with the EU developed more general Code of conduct, the Data Protection Act (2018)¹⁶⁰ which implement the EU's, GDPR and the EU additionally to the GDPR has implemented the EU AI Act. AI ethic software, on the other hand, includes Aequitas, Deon, SHAP, LIME, AI FactSheets 360, AI Fairness 360, AI Explainability 360, PAIR Facets, What-If Tool, InterpretML, PySyft, Adversarial Robustness Toolbox, Data Version Control, and TensorFlow Privacy.¹⁶¹ These tools were developed to help explain and interpret ML models and improve transparency by collecting relevant information, thereby helping to achieve ethical principles in AI.¹⁶²

The analysis reveals that the identified principal-to-gap issues are due to several factors. Key among these are the lack of practical guidance tools for certain ethical AI principles, as most developed tools only cover a few of the main principles, such as fairness, nondiscrimination, bias, explainability, interpretability, security, integrity, and privacy¹⁶³. Additionally, it is challenging to interconnect AI decisions and processes learned from data, as machine learning model results typically involve multiple optimizations.

¹⁶⁰ Data Protection Act," 2018. [Online]. <https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted>

¹⁶¹ *Ibid* 157.

¹⁶² *Ibid* 157.

¹⁶³ *Ibid* 157.

Chapter VII. Conclusion

AI plays a dual role in the financial sector. It poses significant concerns due to its potential for criminal misuse and the inherent risks of its autonomous nature. Conversely, AI is a crucial tool in fighting financial crimes. Through the research done, we found out that AI is increasingly integrated into the financial sector to enhance efficiency, decision-making, and customer experience. Notable applications include fraud detection and prevention, where AI systems analyze transaction patterns to identify and prevent fraudulent activities; algorithmic trading, where AI-driven algorithms execute trades at speeds and efficiencies beyond human capabilities, optimizing investment strategies; and customer service, where AI-powered chatbots and virtual assistants improve customer service by handling routine inquiries and transactions. However, the utilization of AI in finance also brings risks such as cybersecurity threats, where AI systems can be targeted by cybercriminals, leading to potential breaches and data theft ; bias and discrimination, where AI algorithms may inadvertently perpetuate biases present in training data, leading to unfair treatment of customers ; and operational risks, where reliance on AI can introduce systemic risks if the technology fails or behaves unpredictably. All of this being attached to the ‘black box’ aspect of the machines.

Through the second part of the thesis, it is concluded that the regulation of AI varies significantly across regions, each reflecting different priorities and regulatory philosophies. In the US, a more flexible approach is adopted, allowing significant industry input to shape AI regulations. This strategy aims to foster innovation and maintain the country's technological leadership.¹⁶⁴ However, this flexibility might come at the cost of comprehensive safeguards against AI-related risks, potentially leaving gaps in protections against AIC. The EU leads in the field with a stringent regulatory framework emphasizing human rights protection. Measures such as mandatory human rights impact assessments for high-risk AI applications ensure thorough oversight. While this approach provides robust safeguards, it is concluded it could also hinder innovation and market entry, especially in fast-paced sectors like finance. Finally, the UK appears to adopt a more cautious and balanced approach, utilizing sector-specific regulators to oversee AI implementations rather than hastily advancing the regulation process. This method aims to harmonize comprehensive oversight with a degree of regulatory flexibility, promoting both innovation and safety. As previously mentioned, current

¹⁶⁴ <https://thediplomat.com/2023/11/global-competition-for-ai-regulation-or-a-framework-for-ai-diplomacy/>

discussions on AIC regulations focus on balancing innovation with risk mitigation. Policymakers and industry leaders are working together and debate the extent of regulatory intervention required to prevent AIC without stifling technological progress. In the US, a 2023 consultation between senators and tech CEOs resulted in a unanimous agreement that the government should play a role in regulating AI due to the extensive list of associated risks.¹⁶⁵ The EU's stringent regulations are praised for their protective measures but criticized for potentially hindering innovation. The UK's middle-ground approach is often cited as a potential model for achieving equilibrium.¹⁶⁶ In essence, both the EU, UK and U.S. approaches have their merits and drawbacks. The challenge remains in finding the right balance between promoting innovation and mitigating risks to effectively govern AI. A unified legal framework and a comprehensive definition of AI crimes should be created in order to identify them; -States that have not yet implemented substantive and procedural penal laws for AI crimes shall expedite the amendment and rationalization of their laws to appropriately address such crimes, in order to avoid legislative deficiencies and overcome legal gaps in this regard.

This thesis provided a thorough analysis of AI integration in the financial sector. The first key question addressed was: - *How is AI integrated into the financial sector, notably in financial services and investment companies, and what risks are associated with its utilization?* - It was found that AI is primarily used not only to process data from clients but also to enhance the effectiveness of different business agents. Technology development has introduced a competitive setting for companies, making it essential for financial institutions to adopt automated systems to remain relevant and competitive. The second question addressed by this thesis was: - *What are the ongoing discussions surrounding the current challenges related to AI regulations?* - This was answered through a comparative analysis of ongoing debates in different jurisdictions, providing an up-to-date picture of the legal discussions from regulators and the available solutions¹⁶⁷. Finally, the third question was: - *How are various countries addressing the emerging challenge of AI in the financial industry?* - This was answered by highlighting a cross-cutting approach as some jurisdictions focus more on supporting Tech

¹⁶⁵ FRANCE 24 English. (2023, September 15). AI regulation: US, EU, UK diverge on how to mitigate many risks. YouTube.

¹⁶⁶ Greene, N. (2024, February 28). UK versus EU: Who has a better policy approach to AI? *Tech Policy Press*. <https://www.techpolicy.press/uk-versus-eu-who-has-a-better-policy-approach-to-ai/>

¹⁶⁷ *Ibid* 31.

companies to secure a leading position in technology development. These jurisdictions also invest more in research for responsible AI.

Bibliography

(A) Articles

- Beneke, F., & Mackenrodt, M.-O. (2018). Artificial Intelligence and Collusion. Max Planck Institute for Innovation and Competition. <https://link.springer.com/article/10.1007/s40319-018-00773-x>. Accessed the 10 July 2024.
- Bradford, Anu. (2023). 'The American Market-Driven Regulatory Model', Digital Empires: The Global Battle to Regulate Technology (New York, 2023; online edn, Oxford Academic, 21 Sept. 2023), <https://doi-org.uaccess.univie.ac.at/10.1093/oso/9780197649268.003.0002>.
- Klein, A., & Lee, S. (2024, July 5). AI and financial stability: Mitigating risks, harnessing benefits. Brookings. <https://www.brookings.edu/articles/ai-and-financial-stability-mitigating-risks-harnessing-benefits/>
- RegTechOne. (2023, February 7). The history and evolution of the fintech industry. <https://regtechone.co/en/the-history-and-evolution-of-the-fintech-industry/#:~:text=The%20fintech%20industry%20as%20we,in%20the%201970s%20and%201980s>
- Sundblad, W. (2018, October 18). Data is the foundation for artificial intelligence and machine learning. Forbes. <https://www.forbes.com/sites/willemsundbladeurope/2018/10/18/data-is-the-foundation-for-artificial-intelligence-and-machine-learning/#12324a7751b4>
- Greene, N. (2024, February 28). UK versus EU: Who has a better policy approach to AI? *Tech Policy Press*. <https://www.techpolicy.press/uk-versus-eu-who-has-a-better-policy-approach-to-ai/>

(B) Books

- Baker, D.J., & Robinson, P.H. (Eds.). (2020). Artificial Intelligence and the Law: Cybercrime and Criminal Liability (1st ed.) (p.34). Routledge. <https://doi-org.uaccess.univie.ac.at/10.4324/9780429344015>.
- Boden, Margaret A., 'What is artificial intelligence?', Artificial Intelligence: A Very Short Introduction, Very Short Introductions (Oxford, 2018; online edn, Oxford Academic, 23 Aug. 2018), <https://doi.org/10.1093/actrade/9780199602919.003.0001>
- Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford University Press.
- Carbó, S., Gardener, E. P. M., & Molyneux, P. (2005). Financial exclusion. Palgrave Macmillan. <https://doi.org/10.1057/9780230508743>

- Dignum, Virginia, 'Responsibility and Artificial Intelligence', in Markus D. Dubber, Frank Pasquale, and Sunit Das (eds), The Oxford Handbook of Ethics of AI (2020; online edn, Oxford Academic, 9 July 2020), <https://doi.org/10.1093/oxfordhb/9780190067397.013.12> , accessed 28 July 2024.
- Lynn, T., Mooney, J. G., Rosati, P., & Cummins, M. (Eds.). (2019). Disrupting Finance: FinTech and Strategy in the 21st Century. <https://doi.org/10.1007/978-3-030-02330-0>
- Remolina, N., & Gurrea-Martinez, A. (2023). Artificial Intelligence in Finance. Edward Elgar Publishing.
- Zhu, J. (2022). Big Data and AI in Digital Finance. Springer. <https://doi.org/10.1007/978-3-030-94590-9>

(C) Reports

- Bank of England. (2023, October). Artificial intelligence and machine learning. Bank of England. <https://www.bankofengland.co.uk/prudential-regulation/publication/2023/october/artificial-intelligence-and-machine-learning>
- European Parliament. (2024, June 20). AI liability directive. In A Europe Fit for the Digital Age. <https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-ai-liability-directive>
- Financial Conduct Authority. (2023). AI in financial markets: The FCA's approach. <https://www.fca.org.uk/publication/corporate/ai-update.pdf>
- New York State Department of Financial Services. (2021, March). Report on Apple Card Investigation. https://www.dfs.ny.gov/system/files/documents/2021/03/rpt_202103_apple_card_investigation.pdf

(D) Master Theses

- Leopoldi-Wieshaupt, D. (2020). The need for ethical algorithms: The European approach to artificial intelligence [Der Europäische Ansatz zu Künstlicher Intelligenz] (p. 26). University of Vienna. <https://phaidra.univie.ac.at/detail/o:1391941>

(E) Papers

- Abdumuratovich, K. S. (2024, April 13). Comparison of AI Policy. SO‘NGI ILMIY TADQIQOTLAR NAZARIYASI, 7(4). <https://bestpublication.org/index.php/sit/article/view/10723/10694>
- Ahmadi, Sina. (2023). Open AI and its Impact on Fraud Detection in Financial Industry. Journal of Knowledge Learning and Science Technology ISSN 2959-6386.

https://www.researchgate.net/publication/377111879_Open_AI_and_its_Impact_on_Fraud_Detection_in_Financial_Industry

- Bamberger, K. A. (2009). Technologies of compliance: Risk and regulation in a digital age. *Texas Law Review*, 88, 669–739.
- Bradford, Anu. (2023). Chapter 1: The American Market-Driven Regulatory Model. In *Digital Empires: The Global Battle to Regulate Technology* (p. 37). Oxford University Press. <https://doi-org.uaccess.univie.ac.at/10.1093/oso/9780197649268.003.0002>
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*.
- Buchanan, B. (2019, April 2). Artificial intelligence in finance. Zenodo. <http://doi.org/10.5281/zenodo.2612537>
- Buckley, R. P., Zetschek, D. A., Arner, D. W., & Tang, B. W. (2021). Regulating Artificial Intelligence in Finance: Putting the Human in the Loop. *Sydney Law Review*, 43(1), 43+. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3831758
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the ‘Good society’: the US, EU, and UK approach. *Science and Engineering Ethics*, 24, <https://link.springer.com/article/10.1007/s11948-017-9901-7>
- Corrêa, N.K., Santos, J.W., Galvão, C. et al. Crossing the principle–practice gap in AI ethics with ethical problem-solving. *AI Ethics* (2024). <https://doi.org/10.1007/s43681-024-00469-8>
- Demirguc-Kunt, Asli & Klapper, Leora & Ansar, Saniya & Jagati, Aditya. (2017). Making It Easier to Apply for a Bank Account: A Study of the Indian Market. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3043539
- Ferrara, E., Cresci, S., & Luceri, L. (2020). Misinformation, manipulation, and abuse on social media in the era of covid-19. *Journal of Computational Social Science*, 3(2), 271-277. <https://doi.org/10.1007/s42001-020-00094-5>
- Gershman, S. J., Horvitz, E., & Tenenbaum, J. B. (2015). Computational rationality: a converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273-278. <https://doi.org/10.1126/science.aac6076>
- Gonçalves, A. R., Meira, A. B., Shuqair, S., & Pinto, D. C. (2023). AI in fintech decisions: the role of congruity and rejection sensitivity. *International Journal of Bank Marketing*, 41(6), 1282-1307. <https://www.emerald.com/insight/content/doi/10.1108/IJBM-07-2022-0295/full/html>
- Hodge, L. (2021). Artificial intelligence and the law: Cybercrime and criminal liability. *The British Journal of Criminology*, 62(1), 257-275. <https://academic.oup.com/bjc/article-abstract/62/1/257/6284070>
- Kappel, R. (2024, July 1). AI Regulation in Finance: Steering the Future with Consumer Protection at the Helm. <https://www.centraleyes.com/ai-regulation-in-finance/>
- King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. *Science and Engineering Ethics*, 26(2).

- L. N. Tidjon and F. Khomh, 'The Different Faces of AI Ethics Across the World: A Principle-to-Practice Gap Analysis,' in *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 820-839, Aug. 2023, doi: 10.1109/TAI.2022.3225132.
- Lalchand, S., Srinivas, V., Maggiore, B., & Henderson, J. (2024). Deepfake banking fraud risk on the rise.
- Lin, T.C.W., J. Fanto, J. Fisch, J. Heminway, D. Hollis, K. Johnson, et al. (2017). The new market manipulation. *Emory Law Journal* 66: 1253.
- Lui, A., & Ryder, N. (Eds.). (2023). *FinTech, Artificial Intelligence and the Law: Regulation and Crime Prevention* (p. 193). Routledge. ISBN 9781032012469.
- Mhlanga, D. (2020). Industry 4.0 in Finance: The Impact of AI on Digital Financial Inclusion. School of Accounting, University of Johannesburg, Johannesburg, South Africa.
- Mhlanga, D. (2020). Industry 4.0 in finance: the impact of AI on digital financial inclusion. *International Journal of Financial Studies*, 8(3), 45. <https://doi.org/10.3390/ijfs8030045>
- Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 427-436). IEEE. doi: 10.1109/CVPR.2015.7298640.
- Paul, Sandeep. (2019). Use of Blockchain and Artificial Intelligence to Promote Financial Inclusion in India Smita Miglani Indian Council for Research on International Economic Relations.
- Pranav Kumar Chaudhary, Aakash Kishore Chotrani, R. M. M. B. P. R. M. N. (2024). Ai in fraud detection: evaluating the efficacy of artificial intelligence in preventing financial misconduct. *Journal of Electrical Systems*, 20(3s), 1332-1338. <https://journal.esrgroups.org/jes/article/view/1508>
- Schwab, K. "The Fourth Industrial Revolution: What It Means and How to Respond" (2017) World Economic Forum. <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>
- Sheikh, H., Prins, C., Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In: *Mission AI. Research for Policy*. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_2
- Seymour, J., & Tully, P. (2016). Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter. <https://www.blackhat.com/docs/us-16/materials/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter-wp.pdf>
- Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57–84. <https://doi.org/10.1080/17579961.2021.1898300>

- Stinson, C. (2022). Algorithms are not neutral. *AI and Ethics*, 2(4), 763-770.
<https://link.springer.com/article/10.1007/s43681-022-00136-w>
- Toms, S., Simon, D. A., & Vermynck, E.-C. (2023, December 12). How regulators worldwide are addressing the adoption of AI in financial services. Skadden Publication / AI Insights.
<https://www.skadden.com/insights/publications/2023/12/how-regulators-worldwide-are-addressing-the-adoption-of-ai-in-financial-services>
- Wen, W. and Han, X. (2024). An introduction of transaction session-induced security scheme using blockchain technology: understanding the features of internet of things–based financial security systems. *Managerial and Decision Economics*, 45(4), 1817-1834.
<https://onlinelibrary.wiley.com/doi/10.1002/mde.4043>
- Yan, Q. (2023). Legal challenges of artificial intelligence in the field of criminal defense. *Lecture Notes in Education Psychology and Public Media*.
<https://doi.org/10.54254/2753-7048/30/20231629>