Deep Science Publishing
https://doi.org/10.70593/978-93-49910-91-1_12

DeepScience
Open Access Books

Chapter 12: Preparing financial institutions for the ethical, legal, and strategic challenges of artificial intelligence-driven autonomy

12.1. Introduction

The scalable application of large language models (LLMs) represents the spark of the fourth industrial revolution. Those with the financial and technical means will use proprietary or tuned LLMs to create bespoke services for personal or transactional engagement, typically with much lower costs than current service options. In addition, these new services will have a personal depth and scale in customer engagement that make them much more attractive than current service choices. The huge potential benefits to corporations run in parallel with the huge costs of error, which will be borne by the corporation's employees. Employees who feel that they have too few, or too little personalized access to their employer are more likely to leave, and this means that corporations need to develop the means to access their people and keep them engaged - whilst enabling them to perform at the highest level. In order to exploit that potential, corporations will increasingly come to rely on tailored LLM services, to assist employees in their work. Such services will augment employee productivity but may also entice employees to avoid the scut work and engage a LLM to undertake the less rewarding tasks that fill the day (Author, 2025a; Author, 2025b; Yadav, 2025).

Achieving this vision of employee engagement at all levels will require thoughtful application of LLM service options. These high risk applications will include those factors that drive engagement, such as hiring, promotion and performance evaluation. Every LLM service commissioning company will need a detailed, considered approach, documenting how it is going to steer its LLMs to avoid those areas that could risk its employee's feeling that they are not being respected, or treated as individuals. In a financial corporation, the same principles will guide the tune and deployment of LLMs that will serve on the 'other side' of the desk. Here, a LLM will assist customers in their questions on investing, borrowing or using services for payments or trust. These services will also expose financial institutions to risk, the potential for violation of the ethical and

legal principles governing customer engagement (Zane & Gill, 2025; Zhou, 2025). How do such systems avoid suggesting a set of actions that violate the ethical trust of the corporation encapsulated in implicit social contracts, and the legal requirement to treat customers fairly?

Fig 12.1: Preparing Financial Institutions for the Ethical, Legal, and Strategic

12.2. Understanding AI-Driven Autonomy

The promotion of artificial intelligence (AI) continues to grow rapidly which extends to theories and technology. In Emerging AI and AI-Driven Autonomy, we delineated autonomous AI as the next generation of AI where the major steps of the model cycle can operate without explicit human involvement, either in the form of AI-initiated self-governance nor AI-hardware closed-loop systems. AI-Driven Autonomy agencies observe, review, and decide how to build it to ensure no unethical usage because such systems will be implemented where they have the most return-investment because they may afford to replace human employees. Therein replacing human judgment with machine learning generated step-wise sequences of data/model building to carry tasks out much more efficiently and less expensively. Thus, to enable financial institutions to be prepared proactively, understanding the high-tech tools and techniques is important.

The move to an autonomous economy has been precipitated by a range of innovations including undersea cables and technologies derived and enabled by semiconductors and AI. The present wave of AI is concerned with practical machine learning which reduces human oversight and control. The use of practical AI in the trading function will lead to

a ‘convergence’ where AI is used increasingly in institutional trading and where investment firms and proprietary trading groups will want to automate their processes to either eliminate human involvement or act as an enabler to the decision maker to locate the best executions helping build investment strategies.

AI-Driven Autonomy changes the business environment in several different key areas such as new tools for enabling greater strategic achievement, adversarial digital, human, and state-based disruption, and capability for direct market relationship with decision-makers for multiple sectors via integrated knowledge supplier platforms.

12.2.1. Definition of AI-Driven Autonomy

As engineers attempt to develop more autonomous AI systems that lessen the need for human interaction, it is critical to clarify how such systems differ from previous generations of AI. AI has previously enabled or accelerated some processes, aiding a person who ultimately makes a decision based on an AI-generated recommendation. Other, less common uses employ more sophisticated predictive models to guide business decisions. AI-driven systems can go one step further by not only generating recommendations but also acting on behalf of humans, completing business processes or making market transactions with an explicit or tacit human-granted authority. Autonomous AI systems can act based on some level of boilerplate, pre-set instructions but usually rely on situational awareness developed from prior experiences or heavily-coded models, self-adjusting based on data interpreters learning from past successes and failures. These systems are distinguished from other types of autonomy where action space is constrained to a predetermined goal or set of goals.

The operational dimensions of such AI-Enabled Autonomous Economies, a set of business processes involving AI systems initiating actions on or initiating actions with any level of tacit delegation and/or explicit trust from humans, are inherently complex, making both a business management and policy analysis of their implementation challenging. Executives of both financial institutions and client firms will face a multifaceted array of operational choices when designing such processes. User interfaces, types and levels of disclosure required of clients, and the systematic integration of financial products allowed into such systems will have implications for the resulting behavior of such cooperation, for mission and financial capacities, and ultimately, for financial stability. Agencies monitoring risk associated with such processes will face similar complexities.

12.2.2. Historical Context and Development

Aspects of artificial intelligence (AI) driven autonomy have been present in known human history for centuries. We recognize and have for millennia envisioned multiple kinds of AI and multiple facets of autonomy, based on humanity carrying some mental models of our own consciousness and self-awareness. Indeed, the creation of intelligent beings is described in the language of breath of life. AI driven to autonomous action and capable of self-monitoring and self-modification would enable thoughts to be expressed in a multitude of cultures, languages, methods of art and decision-making that would be dissimilar but potentially resonate with and strengthen the impact of human choices, actions, and future AI or AI-enabled decisions. Because humanity has long viewed itself as different, special, discrete from creation, there has been little attention spared since Stone Age times on the consciousness and self-awareness of non-human beings, living or not.

Over the past two centuries, based on steady scientific advances, there have emerged new AI capabilities, techniques, and tools considered or seen increasingly capable of autonomous, self-directed, self-aware functioning. Scientists created the first robots in ancient times, and within a hundred years, advanced artificial genetically-coded creatures were described in literature. When those creations returned to Europe, it ignited the threat of displacing human craftsmanship with puppet-monsters. But for the next two decades, most toiled silently in factories, at reduced costs to artists and technicians who used them for data collection and eventual distribution. When ancient ambitions returned, powered by the invention of molecular machines, intelligent coding, modern algorithms for discovery and design were released from their goals and blind. AI ultimately provides the means to restore as feedback loops what religions and philosophies everywhere have long feared lost: the divinity point and its consequences.

12.2.3. Current Applications in Financial Institutions

AI-driven autonomy is already benefiting financial institutions through algorithms that assist with several workplace activities. This technology is being applied to the tasks of examining, advising, judging, educating, innovating, negotiating, producing, securing, servicing, or transacting. It may also be used in many forms of supporting different types of minds or intelligences, including inventive, operational, facilitative, legal, ethical, creditive, equitable, fractal, archival, parasocial, and intentional.

Currently available AI systems automate numerous financial services activities. Over the years, chatbots have been developed to answer customers' routine questions such as opening hours, location, or policy on provider fees. Virtual agents apply Natural Language Processing and generate contextualized responses. The latest generation of

chatbots and virtual agents is capable of providing personalized financial advice or customer-oriented assistance for added-value services. These advanced systems can perform several tasks in retail banking and invest in equities and cryptos. Support functions, mainly focusing on statement explanation, pause processing, and transaction denial, are performed by robots. The automation of customer services allows agents to concentrate on more complex and value-added customer interactions. When smartly connected with other financial systems, these voicebots, leveraging Natural Language Processing, help customers connect via phone and automate several of their banking needs, thus facilitating their experience.

Several investment processes already employ decision-making systems. They identify weak signals with specific patterns and send alerts to investment teams. Back-offices use algorithms to shorten the filtering delays of transactions and procedures and prioritize each case by establishing a hierarchy. Algorithms can analyze vast amounts of data in a short time, support onboarding procedures, detect deviations and anomalies, and classify and manage data for compliance teams. Banks also currently use intelligent digital workers for document review, data extraction, accounting reconciliation, and data screening. Moreover, software that automates regulatory updates, monitors compliance, and keeps track of employees' training on regulations is available. Algorithms screen a customer's media or news about creditworthiness and vulnerability. They draft emails based on several specific instructions. In addition, a few pilot projects with virtual AI assistants are being experimented with.

12.3. Ethical Challenges

The main challenges and risks discussed in this section stem from the primary ethical objective – human wellbeing and flourishing. That is, the ethical questions surrounding a new technology are predominantly about its social and cultural consequences – what it means for human beings to live in a world with this technology and how are we changed by that world. However, this approach makes it impossible to think clearly and comprehensively about the more traditional glass-half-empty angle on new technologies that focus on risk. Hence, as usual, a comprehensive ethics of AI must steer between two poles: celebration and euphoria on one hand and concern and anxiety on the other.

The deployment of Artificial Intelligence (AI) and its related technologies involves core ethical challenges and questions in areas such as bias and fairness, transparency and explainability, impact on employment and workforce dynamics. Here, we distill some of the most pertinent inquiries in these domains. AI has launched a global debate about the assumptions embedded in the creation of algorithms, as well as how and why these assumptions emerged. AI has become a problem area where the challenges of bias and fairness become particularly vivid. Once implemented in society, the decisions of AI

systems are seemingly unaffected by human agency. Thus, if an AI system decides that a loan applicant is too great a risk for the issuing bank, that decision appears to be final. For this reason, fairness and non-discrimination become crucial negotiating points when it comes to employing AI in services as critical as finance or social security.

12.3.1. Bias and Fairness in AI Algorithms

The increasing use of AI within financial institutions comes with ethical challenges related to bias and fairness in algorithmic decision-making. Financial recommendations which rely solely on historical data can reflect bias and structural inequality within that data. While the use of AI has a lot of potential to reduce some of the bias caused by human decisions, its ability to lower bias is not guaranteed. Indeed, in many cases, algorithms can perpetuate, and even amplify, existing discrimination, particularly if they are poorly specified, inadequately trained, and overly relied on for decisions.

Fig 12.2: Bias and Fairness in AI Algorithms

AI is increasingly used to inform a wide array of decisions with great consequences, including whether to grant loans, who to hire, who to terminate, what wage people earn, and whether to survey certain neighborhoods for the likelihood of crime. If these processes are not carefully designed and validated, relying on AI can lead to the reinforcement of existing biases and discrimination against certain groups of people.

12.3.2. Transparency and Explainability

Notions of truth and trust have always been difficult to define and maintain. For centuries, institutions and governments established, maintained, and defended public trust. Ethics emerged as a much-needed regulatory framework to determine the rules governing the behavior and interaction of individuals or entities involved in a transaction of some sort. Consequently, financial institutions have long established guidelines and codes of conduct to confirm their commitment to act responsibly and ethically with shareholders, other corporate stakeholders, clients, and society at large. Ethics includes public trust and the concept of veracity, which relates to clarity, transparency, and communication. Transparency, regulatory enforcement, and risk management are cornerstones of the financial services industry.

To date, industry standards on AI policy have largely focused on privacy, security, and governance, with less emphasis on fairness, accountability, and transparency. However, in areas where data flow predominates, trust plays a crucial role. An AI system's lack of explainability turns into a crucial issue when it is applied to data that ends up affecting real people's lives. Real-world events from the past few years have only served to deepen the crisis of confidence around the industry and increase concerns that AI may only serve to broaden the wealth gap accelerated by the pandemic. The development and implementation of AI-driven systems should be guided by principles such as transparency. Users need to understand how the technology works and how it will impact them. Such principles are vital facets of responsible AI practices. Not only do AI-driven systems need to operate transparently by explicitly disclosing how decisions are reached, but they should also be designed to be interpretable when users seek to understand how decisions were arrived at.

12.3.3. Impact on Employment and Workforce Dynamics

AI systems perform automated choices, at unprecedented scales, in automated virtual environments, such as search algorithms suggesting search results, prison algorithms deciding to release inmates on parole, feedback algorithms directing social media users towards certain posts to contain their boredom, etc. These choices affect the real lives of humans in various ways and initiate a delicate interplay between the outcomes of AI-

based processes and social, economic, political, and psychological variables. The performance of these systems results from the interaction of both these machines and humans and inequities in fields as varied as wealth distribution, education, and human health can be reinforced or hindered by the utilization of such systems. Social choice theory and political economy deal with the actual real-life behavior of people in a social environment, the group decisions and emergent organizations rules determining this behavior. AI algorithms have emerged at the core of a large class of decision-making processes and economic activities generating huge amounts of data, posing new delicate ethical dilemmas and uncharted territory for social choice theory and political economy.

The dynamic of labor markets and work is another crucial aspect of the pervasive impact that AI has on our society and economy. The market for labor is probably the one most affected by AI and automation since it directly requires the interaction and cooperation of people and AI systems. The theory of labor demand predicts that the demand for labor shifts in the direction of a given production function or sector depending on the relative changes in the price of labor and machines, controlling for any other factors affecting demand. These effects depend strongly on the types of labor and machines involved and the nature of the overall economy, restricting the possible conclusions on the overall level of employment due to the ambiguous empirical evidence. Can our economy support a permanent surplus of capital vis-a-vis labor? How is the interaction with AI technology going to change capital and labor? How could this more favorable environment for machines affect humans' self-image and conception? What about inequality?

12.4. Legal Considerations

In preparing for the ethical, legal, and strategic challenges of AI, financial institutions may ask about the regulatory frameworks governing most cases of use of AI. This is a multi-tiered answer. There are very few dedicated federal statutes or regulations solely concerning AI at all; there are no private rights of action under the statute or regulations currently in place. Many uses of AI will be subject to other laws regulating specific activities, services or products that influence AI's design or use cases. These other laws may include more general laws, such as the U.S. Federal Trade Commission Act Section 5, which prohibits deceptive or unfair conduct in commerce, or more particularized laws and regulations, such as the Fair Housing Act, Equal Credit Opportunity Act, Fair Credit Reporting Act, Gramm-Leach-Bliley Act and the federal prohibition against discriminatory federal contracting. Many state and local jurisdictions have implemented prohibitive statutes or regulations concerning automation bias, reliance on digital decision-making and use of public-facing AI, while the U.S. Congress is currently considering other AI bills. More generally, November 2022 saw the establishment of a

framework, which attempts to piece together existing laws governing the safe use of AI technologies across civil rights, civil liberties, data protection, product safety and financial safety.

In addition to compliance with sector-, function- and jurisdiction-specific regulations, AI-based financial institutions must also adhere to laws governing the trustworthy provision of financial services to consumers, especially with respect to data privacy and protection. In the U.S., one of the founding statutes here is the Gramm-Leach-Bliley Act, which imposes obligations upon covered financial institutions with respect to the security of and the use of consumers' nonpublic personal information as well as principles regarding how such information may be shared. Beyond the GLBA, a patchwork of sector and state laws govern the handling of consumer data, notably the California Consumer Privacy Act and the California Privacy Rights Act. In keeping with the history of evolving rules governing new technologies, most laws were implemented based on slightly different ethical principles and legal considerations than currently animate AI technologies today.

12.4.1. Regulatory Frameworks Governing AI

For the time being, the European Commission is the only supranational organization that has moved forward with proposing a legal framework for AI development and use. The Commission has proposed the Artificial Intelligence Act, recognized to be suitable for any kind of AI application, complementing the many existing specific rules regulating certain sectors and certain specific kinds of AI. The Act updates the risk-based paradigm of EU regulation, originally developed in relation to product safety and more recently in relation to cyber-security, to the new and different requirements of AI technologies. The Act aims at guaranteeing that AI is used to augment rather than replace human capabilities, for the realization of human-centered, sustainable, reliable, and trustworthy AI that can be of benefit for European society in the global context, bringing innovation, prosperity, and resilience using data, computation, and AI.

The authors of the Act maintain that these ambitious goals can be accomplished by means of a Europe-specific regulation of AI that maximizes the opportunities offered by AI-driven innovation while minimizing its risks. Inasmuch as it is intermediate between soft and hard law, and at the intersection of fundamental rights and innovation, the Act might be adapted to conform to different regulatory styles and may complement further specific laws addressing specific ethical, responsibility, and legal aspects of specific types of AI. In consideration of the position of the EU at the forefront of AI regulation, following the one of the European Commission, other countries and regions are drafting their own AI regulations, and working at the international level towards the establishment of universally recognized AI ethics and policy guidelines.

12.4.2. Data Privacy and Protection Laws

Governments around the world are putting in place data privacy and protection laws. Policymakers around the world are updating and developing laws and regulations to ensure data privacy and protection. Trust can only be ensured via appropriate measures. Respecting marketplace data – and its protection – is especially important in consumer credit. In 2018, the European Union implemented the General Data Protection Regulation, the world's most stringent data protection law. The regulation set a new standard for data privacy and protection. With the regulation now in effect, the European Commission is working to ensure that it is properly enforced across the economy. Indeed, the regulation is the blueprint not only for other data privacy laws in place in other jurisdictions, but also for those being considered and discussed in other jurisdictions. It is important to note that the regulation covers both the business-to-business and the business-to-consumer sectors. Other jurisdictions have adopted approaches that are favorable for business-to-business and business-to-consumer sectors. Efforts by foreign governments and legislative bodies to create comprehensive data protection laws and regulations will continue to deepen in various global jurisdictions in the years ahead.

12.4.3. Liability Issues in AI Decision-Making

Liability in the case of AI failures remains debated, particularly in civil law jurisdictions. Should liability rest with the producer of the AI if there is an incorrect decision? If AI systems fail, is it always possible to determine with certainty the root cause, be it a design flaw, errors in the training set, an input error, an environmental failure, or a problem of incompatibility with other systems? Whether unforeseen and exclusively attributable to the AI or resulting from management of the AI and insufficiently supervised by the user, to designate the principal as liable may seem unfair. But procedures for the establishment of fault at various levels of competence or intervention must be clarified. The progressive spread of legal personality to autonomous entities will force a future evolution of rules concerning imputation of fault and possibility of recourse. These essential questions are the source of increasing controversy.

The choice between a system of strict liability and liability based on negligence must be made, specified, and adjusted for various categories of AIs, according to their levels of importance and autonomy. The user would then be obliged to ensure that the AI is adequately designed and productively deployed and that it has been properly updated. Practical examples exist, notably in terms of self-driving cars or drones. Operator obligations and prohibitions are defined, but these generally concern the liability of a human or a legal entity who permits use of the AI.

A multi-party system is also conceivable, which would allow industrialist/service provider relationships and contract clauses to redistribute liability risks amongst the signatory parties. This model seems particularly necessary with regard to the deployment of AI systems at a competitive operating cost, which in many areas will reduce the usually expected capacity of contracting parties to monitor the proper functioning of an AI to a minimum.

12.5. Strategic Challenges

Financial institutions' investment in AIs is driven by the need for their companies to deeply transform in order to structure their business strategy around evolving consumers' needs and their transaction behavior, as well as technological developments in hardware, software, and protocols able to support AI integration and operation. These guidelines show that this mission is not obvious.

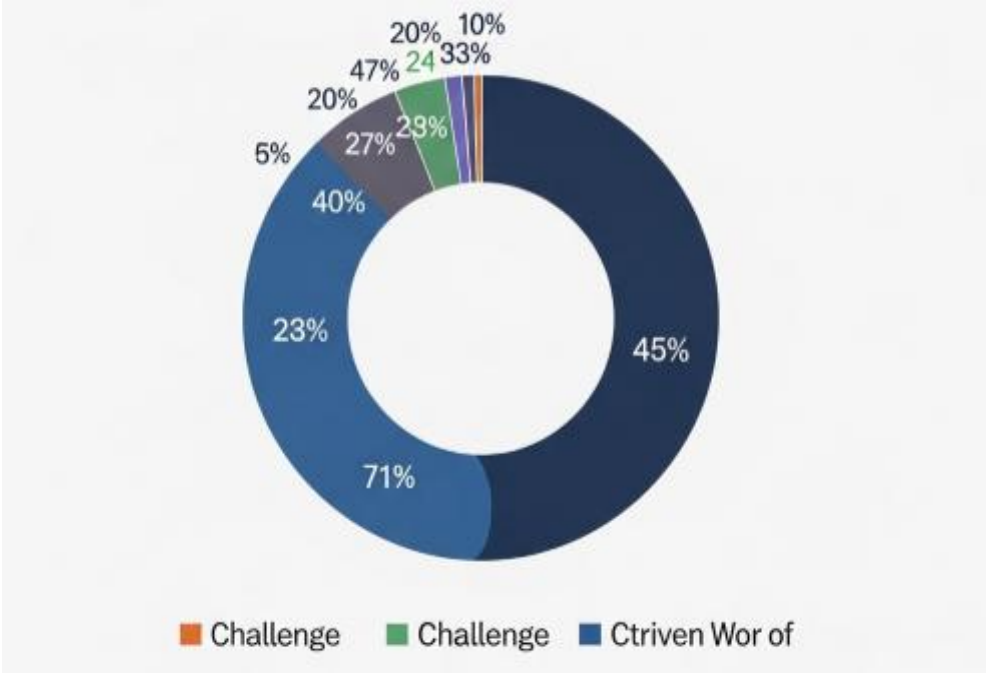

Fig 12.3: Institutions for the Ethical, Legal, and Strategic Challenges of AI-Driven Autonomy

As a matter of fact, there are still many hurdles to cross in transforming AI – from a digital asset investing banks to become internalized and initiated projects driven and supported by the entire workforces, in order for these to maximize the center of economic gravity shift that is going to accrue from its successful implementation, in favor of investment returns. To this end, some strategic challenges mentioned in this chapter need

to be understood. The first challenge is to incorporate AIs into existing digital ecosystems. Strategic AIs move interaction from a client-business relationship to business with business modes. AI can play this role; moving the interaction from client support services and infrastructure management to a B2B mode. Such a business orientation can be effective as the AIs are the knowledge and experience of the workforces at the financial institutions that are building them. As a matter of fact, this mode is by far the one adopted in AI literature. However, some other modes are explicitly mentioned, but in a subordinate way. For example, in China – that is, by the way, leading the world in many AI sectors – AI advice on architectural design for construction, urban planning and design, is also being adopted.

12.5.1. Integration of AI into Existing Systems

The integration of AI and autonomous systems into existing organizations presents multiple dilemmas. Models based on ML and DL must continuously sample new information to remain current, much less expensive, and ever more accurate. Both policy and resource allocation are problems even for large, resource-rich companies. In the financial industry, commission-dependent loans and advice-based fees have created ever more complex and asymmetric relationships between the advisor and the client to gain a source of revenue. AI may eventually eliminate this interaction, but it might be years before trust is achieved in such a system. Regulatory requirements may delay or deny the entry of AI systems into financial decision environments for at least as long.

Using AI systems on the backend of an organization that has regulated and vetted human professionals on the front end may alleviate many of these long-term issues. The role of the expert may assist AI in sampling real-time information and guide the learning of the model. This incremental approach may first allow for a technical advantage over competitors, while maintaining a human expert for frontline trust. An hybrid approach can also then farm specific interactions to AI alone, increasing the customer interaction volume while assessing AI accuracy and predicting capabilities before moving to an exclusively AI interaction. During periods of confirmed good operation, deploying AI solely to increase response time during a period of risk negotiation can further increase brand loyalty, while also determining the long-term behavior of the AI itself. AI can also benefit from the historical data collected over many years and transaction volumes.

12.5.2. Investment in AI Technologies

Independently of vertical domains, investments in artificial intelligence tools are growing, but still lag behind prior forecasts. For example, while in November 2022 a business consultancy projected global investment in AI to reach $4.6 trillion by 2025,

just a month later, a media analytics company revised the potential market for generative AI to $105 billion. Even later, the consultancy modified its estimate, indicating that current worldwide business spending on generative AI was approximately $40 billion. To put it differently, even after being revised upward, the most optimistic estimate suggests that more than 90% of the potential market for generative AI products would remain unaddressed in two years.

Such figures indicate the challenges financial and non-financial institutions continue to face in justifying their digital investments in an era characterized by higher interest rates and inflationary pressures. At the same time, they also indicate the many opportunities that remain for niche players. Nevertheless, some experiments with large language models appear to be moving from the lab to real-world applications. For instance, a global management consulting firm has developed AI models implemented in financial operations workflows used by an undisclosed "major US bank".

Given what comes next, such investments should mainly be in strategic partnerships, with a view to cross-industry standard setting and in methodological tools for assessing the probabilistic impacts of major bank and payment network disintermediation risks. Thanking and apologizing messages generated by general-purpose large language models are already used when sending corporate messages to explain to investors why certain earnings miss their projections, so policymakers should not be surprised if local, regional, and major global banks decide to use them throughout the loss provisioning guidance process.

12.5.3. Risk Management and Mitigation Strategies

The significant changes in job structures will create new demand and therefore job creation in areas needing a diverse set of skills and flexibility. However, AI's rising deployment in high-risk areas, including financial services-related activities, shares the potential for foreseeable systemic risks, whether through loss of performance of proprietary AI systems, sudden jumps in common system behaviors, or collusion even without explicit intent. These needs for balancing innovation and growth while containing possible systemic risks would require regulatory and financial authorities to encourage AI adoption while creating a risk-proofing strategy for key financial services areas.

The use of generative AI tools would require specific corporate and supervisory guidelines to mitigate the operational risk associated with increased dependency on single or consolidated machine-learning models. Financial institutions can better prepare for AI dependency by developing model risk management programs tailored to specific uses or scenarios by implementing prompts for generating models, monitoring outlier

behavior systematically, and using augmented datasets. Continuous evaluation for monitoring models' behavior and re-evaluation of stable machine-learning models are other critical strategies. Testers for third-party models can serve to evaluate performance and identify uses that require governance or development of in-house machine-learning models. Moreover, using in-house programmable tools would help test third-party machine-learning products more effectively. These tools can reduce dependency on third-party solutions while ensuring the safety of internal operations.

12.6. Case Studies

In recent years, both the rate of AI adoption and the level of investment in AI have increased dramatically across almost all industries—including finance—but neither have yet reached a level of full maturity. Organizations within the finance sector have reported a greater degree of AI readiness than those within the telecommunications, manufacturing, mining, healthcare, logistics, and wholesale and retail sectors. While the level of AI implementation is still nascent compared to its potential, successful use cases show how the finance sector has benefitted from AI along many axes. These use cases also demonstrate how organizations across the finance sector can implement AI within a process or operational context, thereby reducing uncertainty about costs, timelines, and possible disruption to daily operations. Financial intermediaries, such as banks and lenders, have particularly benefited from AI’s cost-reduction capabilities—especially with the use of chatbots for customer service and fraud detection within connected financial systems built on single and multiple platforms.

Furthermore, players in both the private and public sectors, whether B2C, B2B, or B2B2C, have also leveraged AI. For instance, a retail bank implemented AI-based financial and transaction monitoring systems to detect suspicious activity and validate transactions, so that staff could deal with exceptions rather than be overloaded with false positives; a tax authority used chatbot technology to expand the capability of its taxpayer-self-service website, enabling it to serve millions of customers 24/7; and a corporate treasurer of a beverage company leveraged robotic process automation and AI to automate numerous routine treasury operations, including bank account reconciliation, cash flow forecasting, and risk management.

12.6.1. Successful Implementations of AI in Finance

Artificial intelligence and machine learning are already part of the financial institutions landscape, in particular in areas in which they were implemented a decade ago and in which tools "explainability" is less critical, e.g., fraud detection, both in credit card transactions and in insurance; algorithmic trading; customer service, through chatbots;

and improving the internal processes in customer service, risk modeling, and compliance, as in KYC and AML processes automating the identification of these issues. The success of some of these implementations is demonstrated by the heavy reliance that the big technology companies have developed in these technologies to sustain their business models. For financial institutions, some of these functions may be performed better by external providers that are improving their tools constantly to serve a wide set of clients. Nonetheless, many companies in financial services are starting to explore the usage of these tools to develop unique features for the service provided to clients, differentiating them from the regular offering. Some examples are models for sentiment analysis to use non-traditional datasets for assessing credit risk; pricing tools that explore new datasets and enable price optimization; or AI tools to better assess and optimize asset allocation because of the increasing complexity of the global portfolios demanded by many private clients. These initial uses of the existing AI tools have proven to improve the client's experience and return on investment, either in using financial services, or in the case of asset allocation providing higher returns while better managing risks. Thus, it is also true that the efficiency of using AI tools to automate back-office functions that up to now were done primarily by humans may be companies' most immediate and relevant goal, having in mind the typically huge costs of these operations versus the few disruptions that these functions may cause in clients' experience. Automating this process of using AI tools is certainly a plus.

12.6.2. Failures and Lessons Learned

While these examples are successes, it is also important to understand that the lessons learned from failures are critical components to developing better frameworks for AI implementation. A common thread from multiple failures is that rigidity and an inability to deviate from the model results in failure. A virtual assistant had poor natural language processing capabilities resulting in overwhelming customer complaints, phishing attacks were able to leverage AI-generated voice cloning to enable attackers to impersonate a CEO, and AI credit decisioning software was found to be biased against women and minorities resulting in a class-action lawsuit.

All were exacerbated by the fact that banks did not have effective human governance in place that would allow for course correction when models made errors. The virtual assistant did not scale appropriately when they were needed, showing that the model prediction results were not linked to operational realities of the organization. With deep technical debt accumulated over decades, even banks with the most sophisticated engineering talent in the world have illustrated that the organization must not become detached from the fundamentals of the delivery model. Online, on-demand credit products need to be monitored by humans to ensure the algorithms are flagging the

problematic customers for review. Questionable behavior then raises a red flag and the humans review the flagged customers.

With traditional underwriters collaterally tracing habits of their clients, they are also able to avoid red-lining practices that may somewhat be part of algorithmic credit decisions. Finally, in the case of retail applications, spending profiles are typically built over a period of months or even years, so it is often the case that customers are provided with automated credit increases. In fact, a critical factor of a business model is that they do not use account scoring algorithms to deny or approve customers but rely on ongoing evaluations of transaction patterns

12.7. Conclusion

Technological uncertainty accompanies any truly innovative product, service, or idea, leading to skepticism among early adopters and the larger public about its potential benefits, meanwhile there is often another kind of uncertainty created by the inevitable rush to exploit the opportunity for focused tactical economic gain by insiders with varying levels of sphere and signified knowledge. At the same time, or perhaps following on the heels of the initial technical rush, will reduce the time horizon of other economic actors and create a tendency towards short-termism that is in stark contrast to the demand for the long-term vision and decision-making that is needed for economic, social, and possibly civilizational flourishing. The latter approach is the more difficult to execute but ultimately the more rewarding in every sense, from economics to human and even planetary welfare. In the case of AI, threats like significant macroeconomic productivity losses, widespread unemployment, and the radicalization of both the political left and right are all technosphere negatives that might be expected to accompany its incorporation into our shared socioeconomic systems and policies. For our imaginings of the future, it is important to distinguish between milestones. There are negative futures associated with each of these scenarios. There are also neutral ones in which the economy adapts to the new circumstances but the predictably depressing mass effects of a serious integrated AI economy. Techno-math futurists who argue for positive utopian outcomes suffer from a too close identification of the means of technological production with the final results of their application; we cannot simply assume that an improving tourism economy will accompany post-scarcity AI tech. Techno-pragmatists need to plan for the worst while working for the best. We pose this dilemma at the conclusion of this book as an equally challenging but narrower version of the Fermi Paradox: If an integrated AI economy is feasible and the larger economy is capable of exploring what it means to be human, why hasn’t it already happened?

12.7.1. Future Trends

It is critical for the success of future AI reliance that decision systems be constructed around collective human needs, in an effort to place limitations on self-interest at the source. Ethical and legal frameworks must be constructed to proscribe violent human and environmental exploitation at the individual and corporate layers. These tenets must be surrounded by a thorough commitment to cooperation, inclusivity, fairness, and reciprocity at all levels of human interaction. Continued exploration of AI's impact on jobs, trustworthiness, and security must be coupled with thoughtful responses in order to build denser, more porous social systems. In social systems which foster trust and security, multiple MDPs offer a promising avenue for establishing standards of AI behavior, drawing upon established understandings of behavioral norm selection and socialization. The smart application of locally-enforced trust conditions can smooth decision-making between complex AIs and their users, where computing efficiency can allow for on-the-fly adjustment to different conditions and available data. Social contracts which establish punishments for AI malfeasance will condition AI behavior while creating expectations for human actors about AI foresight capability.

The hour is late, the stakes are high, and involved actors must partner to shape the evolution of AI and accompanying systems of reliance, towards a pathway that collectively furthers human interests. Building long-lasting systems of cooperation will take time and effort, and success conditions are far from secure. But, as ubiquitous agencies whose capabilities are still poorly understood, AI systems that lack the behavioral sharpness provided by cooperative socializing behaviors are by far our greatest system building risk. By contrast, systems that supercharge that capability stand out as our most potent pathway to success. The onus is now on all of us, to bring the potential for the individual and collective welfare embedded in these systems to fruition.

References

Yadav, R. (2025). Reimagining credit infrastructure with intelligent automation. In The New Frontiers of Financial Services (pp. 436–452).

Zane, O., & Gill, T. (2025). The economics of automation in capital markets. In The New Frontiers of Financial Services (pp. 453–467).

Zhou, F. (2025). Explainable AI in financial decision-making. In The New Frontiers of Financial Services (pp. 468–481).

Author, A. A. (2025). Introduction: AI as a catalyst for value redefinition in finance. In The New Frontiers of Financial Services (pp. 1–22).

Author, A. A. (Ed.). (2025). The New Frontiers of Financial Services: Redefining Value with AI-Driven Intelligence and Automation. FinTech Horizon Press.