



Citi  
GPS



# AI in Finance

## Bot, Bank & Beyond

Citi GPS: Global Perspectives & Solutions  
June 2024

Citi is one of the world's largest financial institutions, operating in all major established and emerging markets. Across these world markets, our employees conduct an ongoing multi-disciplinary conversation – accessing information, analyzing data, developing insights, and formulating advice. As our premier thought leadership product, Citi GPS is designed to help our readers navigate the global economy's most demanding challenges and to anticipate future themes and trends in a fast-changing and interconnected world. Citi GPS accesses the best elements of our global conversation and harvests the thought leadership of a wide range of senior professionals across our firm. This is not a research report and does not constitute advice on investments or a solicitations to buy or sell any financial instruments. [For more information on Citi GPS, please visit our website at www.citi.com/citigps.](http://www.citi.com/citigps)



**Ronit Ghose, CFA**  
Head of Future of Finance  
Citi Global Insights  
[ronit.ghose@citi.com](mailto:ronit.ghose@citi.com)



**Sophia Bantanidis**  
Future of Finance Analyst  
Citi Global Insights  
[sophia.bantanidis@citi.com](mailto:sophia.bantanidis@citi.com)



**Kaiwan Master**  
Future of Finance Analyst  
Citi Global Insights  
[kaiwan.hoshang.master@citi.com](mailto:kaiwan.hoshang.master@citi.com)



**Ronak S Shah**  
Future of Finance Analyst  
Citi Global Insights  
[ronak.sharad.shah@citi.com](mailto:ronak.sharad.shah@citi.com)



**Yirou Yu**  
FinTech & Blockchain  
Citi Business Advisory  
Services  
[yirou.yu@citi.com](mailto:yirou.yu@citi.com)



**Prag Sharma**  
Global Head Artificial  
Intelligence Centre of  
Excellence  
Citigroup



**Katie Dilaj**  
Global Head of Product  
Marketing, Treasury and  
Trade



**Barry Van Kooij**  
Global Head of TWCS  
Product Marketing  
Citi



**Aldona Zajac**  
Product Marketing,  
Treasury and Trade  
Solutions  
Citi

### Expert Contributors



**Bin Ren**  
Founder and CEO  
SigTech



**David G. W. Birch**  
Principal at 15Mb Ltd.  
Global Ambassador for  
Consult Hyperion



**Dessislava Savova**  
Partner and Head of the  
Continental Europe Tech  
Group  
Clifford Chance



**Devika Kornbacher**  
Partner and Co-Chair,  
Global Tech Group  
Clifford Chance



**Eric Chang**  
Co-Founder and CEO  
Claira



**Erkin Adylov**  
Founder and CEO  
Behavox



**Ken Sena**  
Co-Founder and CEO  
Aiera



**Kirsty Rutter**  
FinTech Investment  
Director  
Lloyds Banking Group



**Matthew Van Buskirk**  
Founder and CEO  
Hummingbird RegTech



**Sandeep Alur**  
Director  
Microsoft Technology  
Center



**Shameek Kundu**  
Chief Data Officer  
and AI Entrepreneur



**Shelby Austin**  
Founder and CEO  
Arteria AI Inc



**Spencer Reich**  
Founder and CEO  
The SER Group



**Vijay Balasubramaniyan**  
Co-Founder and CEO  
Pindrop



**Zor Gorelov**  
Co-Founder and CEO  
Kasisto

With thanks to:

Damian Kowalski, Faizan Khwaja, Jiten Mehta, Arash Eskandari, Santiago Dinelli and Mike Lu

## AI IN FINANCE

### Bot, Bank & Beyond

AI will transform money and finance.

Artificial Intelligence (AI) could be the General-Purpose Technology (GPT) of the 2020s-2030s. And it will profoundly change finance and money.

GPTs have the potential to transform entire economies, changing the way we live and work. They create new opportunities for growth and innovation, often improving our overall quality of life. They also destroy existing ways of doing things. And as such they also create losers. Especially in the short term.

The steam engine commoditized production and physical movement, powering the industrial revolution. More recently, the Internet revolutionized communication and ushered in the age of information. Similarly, AI may commoditize human intelligence, including analysis, decision making and content creation.

Artificial intelligence may commoditize human intelligence.

Finance will be at the forefront of the changes. Existing jobs have disappeared in prior cycles, to be replaced by new ones. So have firms. What a bank or financial firm looks like in the mid-2020s, be it retail or wholesale finance, looks very different to the mid-1980s, or the mid-1940s! AI will repeat this cycle, possibly speeding it up.

AI itself has gone through many waves of hype and disillusionment from the 1950s onwards (see Appendix for a summary history, especially Figure 40). We wrote about AI in Finance in our 2018 GPS report ([The Bank of the Future – The ABCs of Digital Disruption in Finance](#)), but interest in the topic is much higher now.

Advances in generative AI (GenAI), including the release of ChatGPT in November 2022, marked a turning point. GenAI brought a user interface (UI) to AI and placed it literally in the palm of our hands. The advent of GenAI brought AI to the masses, sparking interest among consumers and key decision makers alike.

Generative AI, a subset of AI, refers to models that can generate high-quality text, images, videos, and other data. For now, GenAI in finance is largely at a proof of concept (POC) stage, but it is rapidly transitioning. In this report, we discuss what use cases are likely in 2024-25, and also look further ahead.

Based on the results of a recent Citi TTS Client Survey, we estimate the global banking sector 2028E profit pool could increase 9% or \$170 billion from the adoption of AI, rising from just over \$1.7 trillion to close to \$2 trillion.

Most incumbents' technology adoption strategy involves adding it on top of existing products. Disruptive startups create new products/services that are native to the new technology.

The tech adoption strategy of most incumbents involves adding it on top of existing products or using the new technology to improve productivity. Startups, by contrast, use new technology to unbundle what incumbents do. Disruptive startups create new products and services that are native to the new technology.<sup>1</sup>

<sup>1</sup> Benedict Evans in Slush Conference in Helsinki, AI and Everything Else, December 2023

As AI-powered agents, bots and beyond, become increasingly prevalent, how will money and finance change? How will the underlying concepts and structures of finance be reshaped? In a bot-to-bot world, where machines transact with minimal human intervention, what does the world of money look like?

Most of us have already been wowed by AI. Many of us now want to know how we will use it.



## Contents

<b>(A.) Finance: What Will Bots Mean?</b>	<b>8</b>
Let's Talk A-bot AI and Finance .....	9
Non-Human Customers – Expert Interview .....	11
Identifying Potential Leaders and Laggards .....	13
Unlocking the True Promise of AI – Expert Interview .....	17
Role of AI in Transforming Finance .....	18
[1] Does AI Create Value for Shareholders? .....	18
[2] How Does AI Impact the Finance Workforce? .....	22
[3] Does AI Create Value for Banks' Clients? .....	26
The Bot Frontier – Expert Interview .....	28
<b>(B.) How Will AI Be Used in Finance?</b>	<b>32</b>
AI Scepticism to FOMO – Expert Interview .....	34
Document & Data Management with AI – Expert Interview .....	35
[1] Coding & Software .....	36
Measuring Developer Productivity – Expert Interview .....	38
[2] Transaction Monitoring and Compliance .....	39
Scaling Compliance with AI – Expert Interview .....	41
[3] Customer Services & Chatbot 2.0 .....	43
Chatbots Gallop From Horses to Cars – Expert Interview .....	45
[4] Credit Risk & Underwriting .....	47
[5] Investment Research .....	50
Next-Gen Financial Research with AI – Expert Interview .....	53
[6] Asset & Portfolio Management .....	55
Supercharging Capital Markets with AI – Expert Interview .....	58
Putting AI on Portfolio Manager's Team – Expert Interview .....	60
Decentralized AI: The Next Frontier? .....	62
Centralized vs Decentralized AI Stack .....	62
Tokens in a Bot World .....	63
AI Agents and Tokenized Money .....	63
<b>(C.) Navigating AI's Potential Pitfalls</b>	<b>64</b>
Managing Risks & Building Trust – Expert Interview .....	66
Bias and Discrimination .....	68
Lack of Transparency & Explicability .....	69
Misinformation, Manipulation & Hallucination .....	73
Grappling with AI-Led Deepfake Proliferation .....	76
Detecting Giraffe Man's Voice – Expert Interview .....	82
AI Amplifies Social Engineering .....	84
AI – An Energy Guzzler .....	86
<b>(D.) Real Regulations for Artificial Intelligence</b>	<b>88</b>
Global AI Regulatory Developments .....	90
International Guiding Principles for Organizations Developing	
Advanced AI System .....	92
Common Themes in Regulating the Application of AI .....	99
The European Union's AI Act .....	100
Regulating AI Robots – Expert Interview .....	103
<b>Appendix 1: AI &amp; General Purpose Technologies</b>	<b>106</b>
<b>Appendix 2: A to Z of Key Terminologies</b>	<b>114</b>

# AI In Finance: Bot, Bank & Beyond

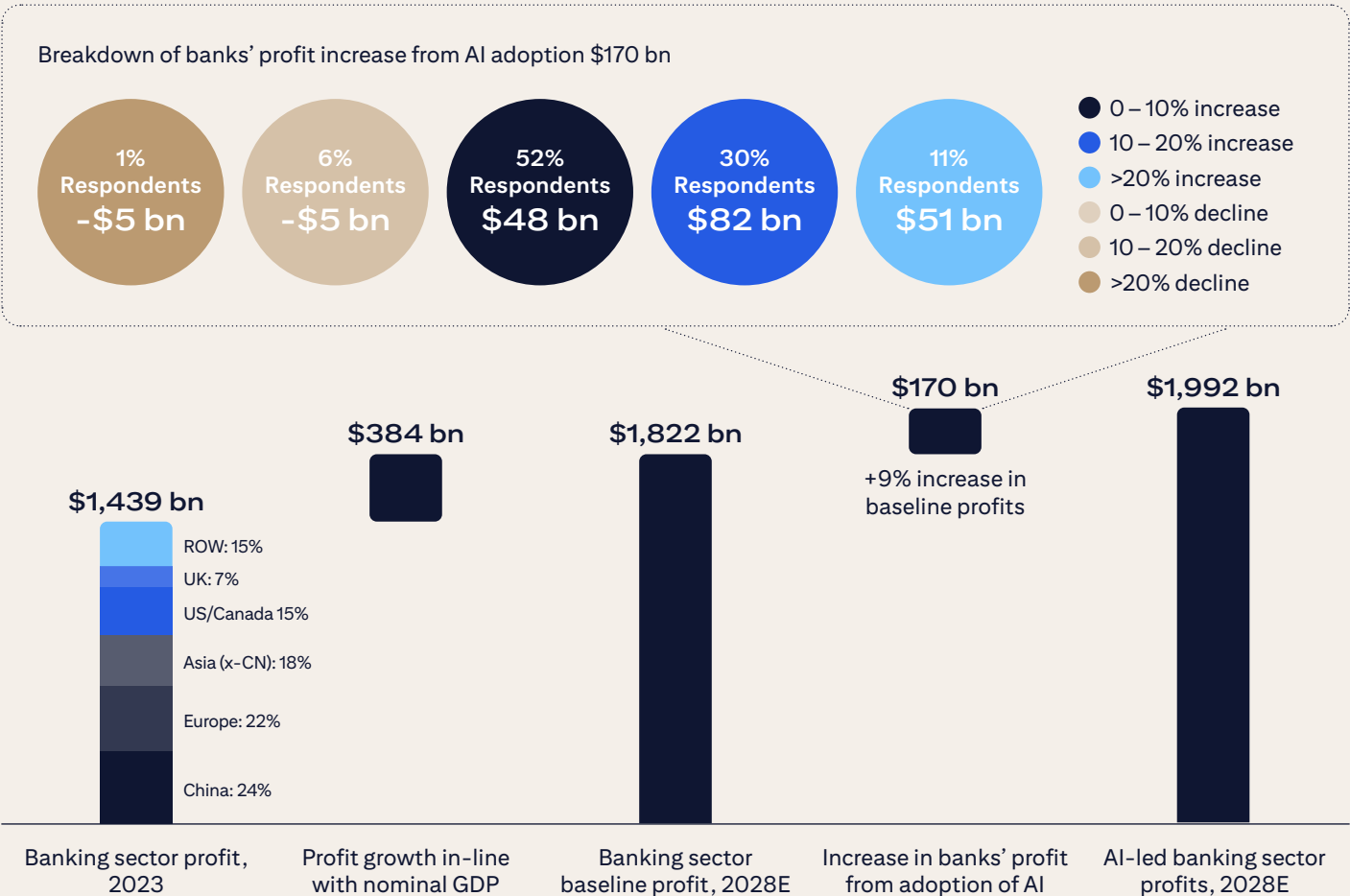
## Impact of Adopting AI on Banks' Profitability

In 5 years time, will AI increase or decrease profits?



Source: Citi Treasury & Trade Solutions Survey 2024  
Number of respondents = 90 including banks, insurers, and asset managers.

## AI Could Add \$170 billion or 9% to Global Banks Sector Profit Pool by 2028E



Source: SNL Financial, IMF World Economic Outlook, Citi Treasury and Trade Solutions Survey 2024, Citi Global Insights

# Paving the Way for AI in Financial Services

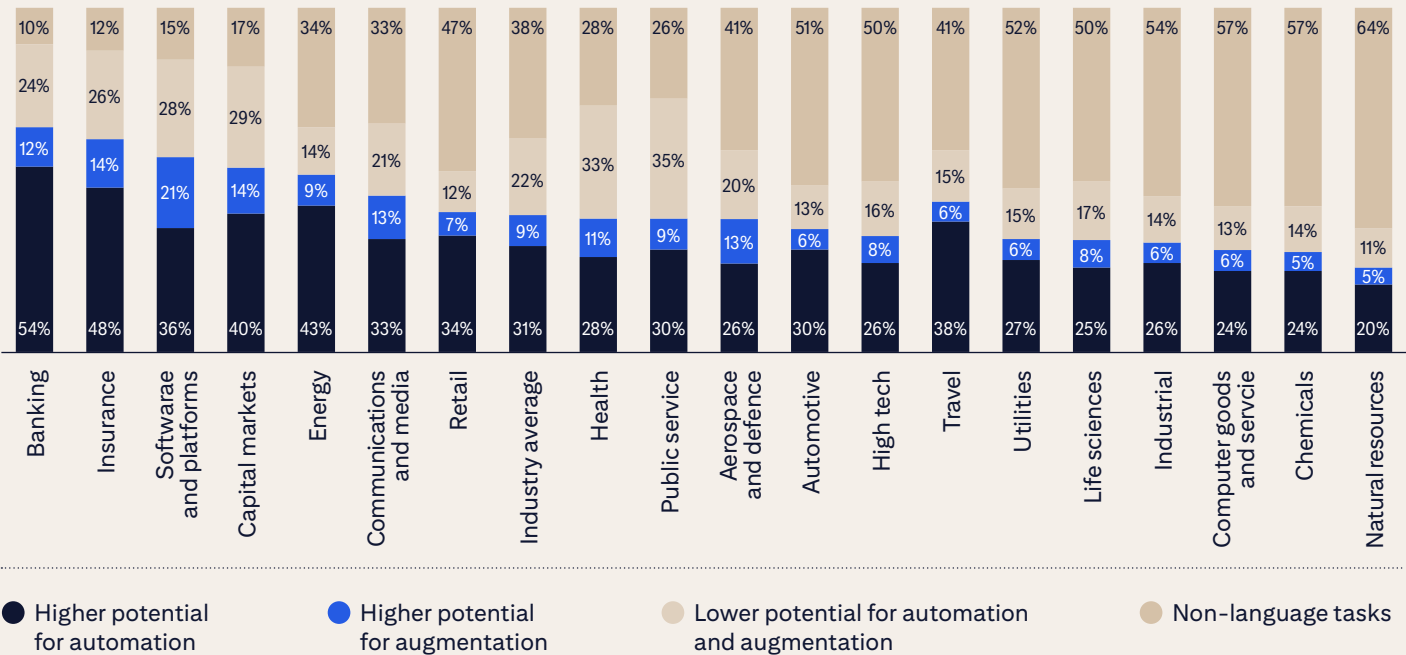
Source: Citi Global Insights



# Potential for AI-Led Job Displacement

Source: Accenture Research Report "A New Era of Generative AI for Everyone", 2023.

Work time distribution by industry and potential AI impact (Based on their employment levels in the US in 2021)



## (A.) Finance: What Will Bots Mean?

Technology is a strategic priority for both the Board and the C-suite. It's become commonplace to hear bank leaders referring to banks as technology companies with a banking license. Financial firms increasingly use sophisticated technology to deliver their services and manage their operations, and that tech is often a vector of competitive differentiation.

Banks were early movers in the first technology wave -- investing in the first mainframe computers as far back as the 1950s -- before falling behind in the recent Internet and mobile era.

We believe generative AI (GenAI) has revolutionary potential in financial services because the sector is information rich. Data is its raw material.

In many respects, finance is the perfect sector for the application of GenAI and AI more broadly.



Generative AI has the potential to change the world in ways that we can't even imagine. It has the power to create new ideas, products, and services that will make our lives easier, more productive, and more creative...

– BILL GATES, CO-FOUNDER, MICROSOFT<sup>2</sup>



Finance, in a world of AI-powered agents, bots and beyond, will likely face change across the board: in terms of market share, employment, and client experience. In the medium term, by 2030 or before, AI-powered bots will play an increasing role in banking and finance. This will challenge many existing ways of doing business.

The changes we expect are a continuation of many themes we have been writing about since our first 2016 GPS on the future of finance: technology and culture are changing, and with it so is society. This is leading to a reshaping of the business of money and the role of the players within it. This change is about to re-accelerate.



The biggest new thing will be the growth of non-human customers.

– SHAMEEK KUNDU, CHIEF DATA OFFICER AND AI ENTREPRENEUR



AI challenges that old French saying: *plus ça change, plus c'est la même chose* ("the more things change, the more they stay the same"). And the biggest new thing will be the growth of non-human bots and agents. The more things change, the more they stay the same. This time some things really will change.

Bankers may think that they lead the way. But many users are adopting technology faster than banks or big business. AI, and GenAI, may be the latest chapter in the **digital internet story of the crowds running ahead of the crown.**

<sup>2</sup> Forbes, "The Most Thought-Provoking Generative Artificial Intelligence Quotes Of 2023", 29 November 2023.



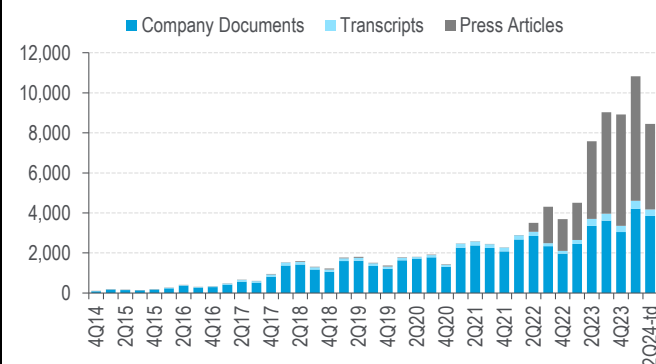
## Let's Talk A-bot AI and Finance

Banks have been discussing AI in filings since 2018, but broader public interest took off in 2023.

Financial institutions have been regularly mentioning AI in regulatory filings since 2018, the year we first discussed AI in Finance in our GPS report ([The Bank of the Future – The ABCs of Digital Disruption in Finance](#)). Before that, references to AI in company filings around the world were relatively limited.

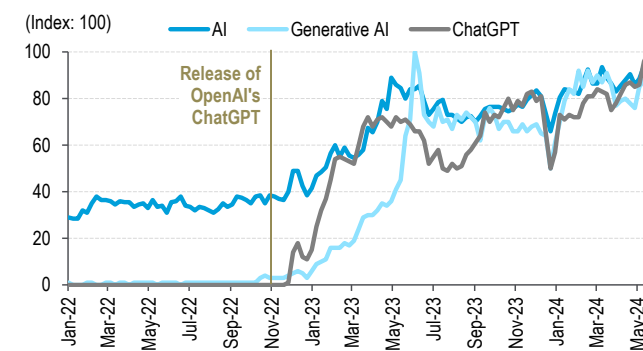
An even bigger spike in external references is evident in media references to AI by financial institutions. These were exceedingly rare pre-2022. Alongside the growth in public interest in AI, and especially GenAI, financial institutions began to discuss the topic much more in press and media.

**Figure 1. Mentions of the Term "AI" by Financial Institutions in Company Filings and Press Articles, 2014-2024**



Note: Latest data as of 23 May 2024.  
Source: AlphaSense, Citi Global Insights

**Figure 2. Google Keyword Search Trends for AI-Related Terms, 2022-2024**



Note: [1] Results for the term "AI" represents average of search results for the terms "artificial intelligence" and "AI". [2] Results for all keywords pulled out individually.  
Source: Google Trends, Citi Global Insights

For many years, AI – or, strictly speaking, machine learning (ML) – has been used in finance on structured data and for quantitative tasks. GenAI will now expand these use cases to unstructured data. Most enterprise data (about 80-90%) is unstructured – locked away in emails, transcripts, documents, and reports.

The industry's reliance on data is massive, driving everything from customer insights to risk management and fraud detection. With the advent of advanced AI analytics, banks can extract valuable insights to personalize customer products and services, optimize operational processes, and comply with regulatory requirements.

- **Will AI Increase Finance Profits?** According to a recent Citi TTS Client Survey<sup>3</sup>, there is a very clear consensus: an overwhelming 93% of financial institution respondents said adoption of AI could improve profitability in the next 5 years, based on productivity gain expectations (Figure 10).
- **What Magnitude of Profits?** Applying the survey results to our forecast of the global banking sector profit pool, suggests AI could boost the total banking industry's 2028E profits by 9%, or \$170 billion, from just over \$1.8 trillion to close to \$2 trillion (Figure 11). This excludes non-bank financial sector profits.
- **Why More Profits?** AI could drive productivity gains for banks by automating routine tasks, streamlining operations, and freeing up employees to focus on higher value activities. GenAI will likely have a big impact on internal facing tasks such as content and information management, coding, and software.

<sup>3</sup> N = 90; where 71 banks, 16 insurance companies, 3 asset managers responded.

- **Innovation versus Industrialization:** While finance sector leaders are optimistic about the profit impact of AI, driven by productivity gain potential, we are cautious about implementation timelines, talent costs, risk of increased competition, rising client expectations, and the costs associated with increased AI-generated activity.
- **Will Resource Use Grow?** AI is going to make finance firms and others much better at managing, analyzing, and creating information. But the information in the world is exploding: global data creation +50% CAGR, 2023-25E (Figure 12). Jevon's paradox: efficiency gains lower costs which creates more demand<sup>4</sup>.
- **Will AI Reduce or Grow Finance Jobs?** Historical technology adoption has not led to reduction of finance workforce but has changed the workforce mix over time. New jobs are constantly created. For example, the US economy has 3x more compliance officers from 2000 to 2023 (Figure 16).
- **Governance and Talent:** The growth of AI may lead to fewer low-skilled roles in operations and technology, but governance and compliance roles will continue to grow. Our survey responses suggest AI talent availability is a challenge for banks and others (Figure 15). Talent wars are not over. Humans are still in demand.
- **Will AI Drive Better Value for Clients?** Integration of AI-powered bots into retail and corporate banking represents a significant potential transformation, offering clients benefits such as automated decision making and the search for best offers, and banks enhanced operational efficiency.
- **Challenges from AI-powered Bots:** A shift to a bot-powered world also poses concerns on data security, regulatory/compliance, and ethical considerations. Since AI models are known to hallucinate and create information that does not exist, organizations run the risk of AI chatbots going fully autonomous and negatively affecting the business financially or its reputation. AI-powered clients could increase price competition in the finance sector, especially in retail financial services. The balance of power may shift. Profit margins may fall.
- **Can Banks Keep Up?** Between the 1950s and 1990s, big business usually had better technology. As the Internet and mobile waves spread across the world over the past two decades, the hierarchy between big business and consumers, when it came to technology, became inverted. GenAI may be the latest such wave.
- **The Race to Adapt:** AI is likely to be adopted faster by digitally native, cloud-based firms, such as FinTechs and BigTechs, with agile incumbent banks as fast followers. Many incumbents, weighed down by tech and culture debt, may lag in AI adoption and lose market share (Figure 3).
- **Trough of Disillusionment Ahead?** All technology goes through cycles: hype, disillusionment, and then mass adoption. AI expectations have been high since 2Q23. As financial firms grapple with the transition from 'Wow' to 'How', the gap between hype and mass production currently remains wide.
- **Digital Leviathans:** Many of the largest technology companies are in an arms race to become the AI provider of choice to banks and companies. Similar to the concentration in cloud computing, AI may reinforce single points of control and thus also single points of failure.

As financial firms grapple with the transition from 'Wow' to 'How', the gap between the AI hype and development remains wide.

<sup>4</sup> The Ben & Marc Show, a16z Podcast, 9 May 2024



Shameek Kundu is an experienced Chief Data Officer helping banks adopt AI responsibly. Most recently, he helped build the Financial Services business at AI testing and monitoring startup TruEra (acquired by Snowflake). Prior to TruEra, Shameek was Group Chief Data Officer at Standard Chartered Bank, where he helped explore/adopt AI across multiple areas. He has served on the Bank of England's AI Public-Private Forum, and the Monetary Authority of Singapore (MAS) initiative on Fairness, Ethics, Accountability and Transparency in AI.

## Non-Human Customers – Expert Interview

### *Q: What is the state of adoption for AI applications in financial services?*

Let us start with traditional/predictive AI. Based on a 2022 Bank of England (BoE) survey and anecdotal evidence from speaking to clients, I would describe traditional AI adoption in financial services as: widespread, shallow, and inconsequential.

Widespread refers to the large number of enterprises experimenting with AI across different use cases. According to the 2022 BoE survey, 72% of firms reported they were using or developing ML applications. AI use cases exist across functions like sales, fraud detection, document processing, credit decisioning, pricing etc.

Shallow refers to the limited scale of AI adoption across use cases. The BoE survey suggests the median number of ML applications for mainstream UK financial institutions to be just 20-30. For context, most banks would have a few hundred systems, with bigger banks even having a thousand or more models/sub-models.

Inconsequential refers to the limited perceived impact of AI system failures on critical business operations. According to the BoE survey, less than 20% of the already few AI use cases were critical to business. Thus, the net sum of all AI efforts in UK financial institutions averaged 4-6 applications in business-critical areas.

There are exceptions, of course – we know of FIs in North America, Europe, and Asia that have hundreds of models live in production. At this stage, anywhere from 10-20% of FIs would fall in this category.

### *Q: What is the state of GenAI adoption in financial services?*

Current use of GenAI in production is negligible and lesser than predictive AI. We are still very early in the commercialization of GenAI in finance; they are nearly all pilots or Proof of Concepts. But I am quite optimistic about GenAI's impact over the next five years.

Productivity improvements are very likely as GenAI underpins at least partial automation of manual processes – e.g., reading and writing code, document processing, internal helpdesks, sales support. But we are yet to see examples of entirely new banking products getting introduced due to GenAI.

It is hard to invent products in banking but there is white space in distribution and GenAI can help make a difference there. For example, wealth management for the masses and hyper-personalization in banking. The creation of new products for a market of one, not selecting from a menu, has not been done before.

The distinction of predictive and GenAI is also likely to diminish over time. For example, use of AI in finance will not only focus on generating reports, but also understanding trends, plotting future trajectories.

### *Q: What will change due to GenAI?*

The biggest new thing will be the growth of non-human customers. AI agents will exist in finance by 2030, possibly earlier. Clients may have AI agents and bots before banks do. Bankers often think that they are in control. They may not be.

Hyper-personalization may be good in wealth management, but it can be a double-edged sword in insurance. The availability of more data, often underpinned by AI, can allow insurers to segment risks more granularly. While this can help make pricing more risk-based, such micro-segmentation could also result in smaller risk

pools and a larger number of them, reflecting more accurately the risks of those within each pool. This could potentially affect the ability of high-risk consumers to have access to affordable insurance coverage.

There is an ongoing debate if an AI-powered world is good for lower quality workers or existing superstars. In an AI-powered world, it is possible that lower quality finance sectors workers could get raised to median levels and good performers could scale to near superstar levels.<sup>5</sup> Overall, employee productivity will improve.

---

<sup>5</sup> Harvard Business School, Navigating the Jagged Technological Frontier, 2023

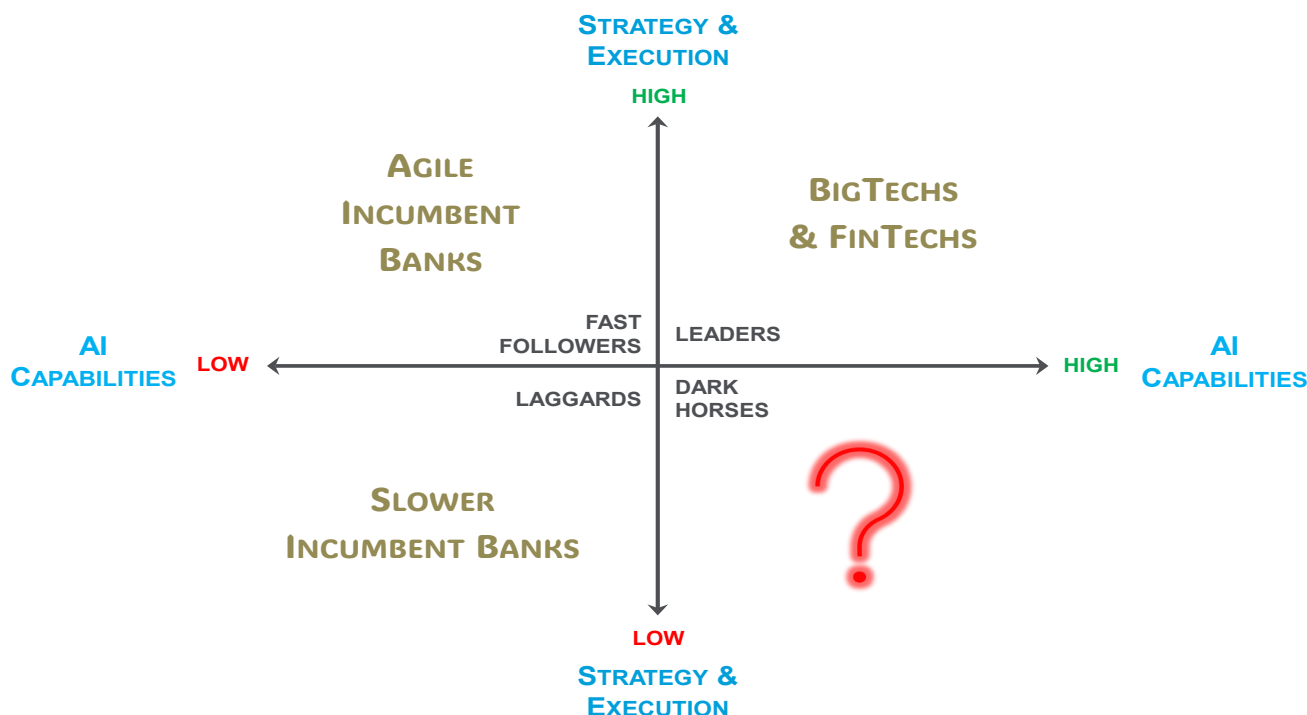
## Identifying Potential Leaders and Laggards

The implementation of AI across finance is poised to significantly transform the financial services sector and economy overall. In financial services, we consider likely leaders and laggards based on two key parameters.

Digital natives should lead the way with agile banks likely to be fast followers. Traditional banks with outdated systems risk being laggards.

1. **AI Capability:** Measures the technical proficiency, talent, infrastructure, and overall capability of an enterprise to develop, integrate, and leverage AI tools. This includes data quality, AI expertise, and technological infrastructure.
  2. **Strategy & Execution:** Strategic vision of an enterprise, senior leadership focus and execution capabilities, regulatory compliance, as well as the ability to scale and adopt AI initiatives to a changing market environment.
- **Leaders:** Companies with strong technical capabilities, AI talent, and effective leadership and execution are the likely winners. Leading BigTechs, investing heavily in AI, should predominate in this category. Some FinTechs will also join this group, possibly together with a few product and technology-focused banks.
  - **Fast Followers:** Companies in this category closely monitor market innovations and adopt proven technologies via strategic investments and partnerships; these companies make the most of their strong execution capabilities to rapidly adapt and integrate AI solutions. Agile incumbent banks fall in this quadrant.

Figure 3. Potential Leaders & Laggards in AI-Led Financial Services



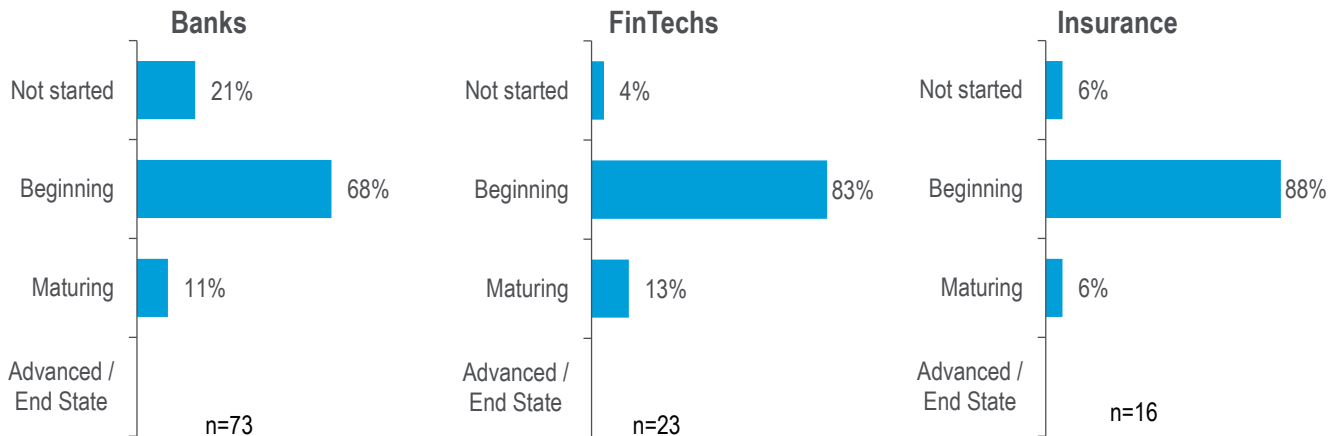
Source: Citi Global Insights

- **Laggards:** Companies that are disadvantaged with debt – tech debt, culture debt, organizational debt are expected to lag behind in AI terms. This may reflect risk averse corporate leadership and/or a tight regulatory environment. This quadrant could include some traditional banks that struggle with outdated systems and bureaucratic processes, making it difficult for them to respond to market changes and compete effectively.

Tech & culture debt will hold back some incumbent banks – many banks are already falling behind in the AI arms race.

These views echo the results of the recent Citi TTS survey which underscore the slow adoption of AI by many banks. Nearly a quarter of banks have '*not started*' on AI adoption yet. In contrast, this number stands significantly lower at just about 5% for FinTech and insurance clients. Furthermore, 68% of respondent banks are still only in the '*beginning*' stages of their AI journey (Figure 4).

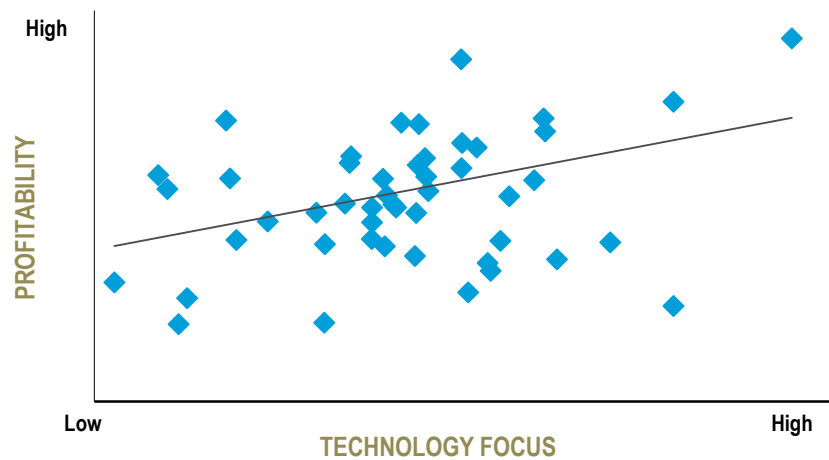
Figure 4. Banks Are Playing Catch Up in the AI Arms Race



Source: Citi Treasury and Trade Solutions Survey 2024

Analysis of developed market banks shows that those with a focus on technology tend to achieve higher profitability (Figure 5). Technology is not the only driver of profitability, of course, but it is an important factor. And in a world of exploding data, and a growing focus on AI, technology will remain an important strategic differentiator.

Figure 5. Banking Sector Strategic Focus on Technology vs. Profitability



Notes:

[1] Technology focus measured as tech & communication cost as percent of operating expense over FY2021-23.

[2] Profitability is measured by the average return on equity over FY2021-23.

[3] Banks in Australia, Hong Kong, Singapore, the UK, the US, and Western Europe, with market cap > \$1billion

Source: SNL Financial, Citi Global Insights

Banks globally rank among the largest tech spenders (outside of the Tech industry), with some large banks allocating close to 20% of their expenses to technology. However, a substantial portion of banks' tech spend is on maintenance costs to support legacy systems, and business as usual.



Banks' spend on AI is small compared to their overall tech spend, but it is growing fast.

Banks spend a smaller percentage of their tech spend on innovation, new products, or business model changes. Spending on AI and especially GenAI is relatively small compared to overall tech spend. Much GenAI spend at present is on consultants and external vendors, with work still skewed toward pilot schemes rather than full production.

Enterprise investment in AI is estimated to be only around \$70 billion in 2023, with a fraction of that on GenAI.<sup>6</sup>

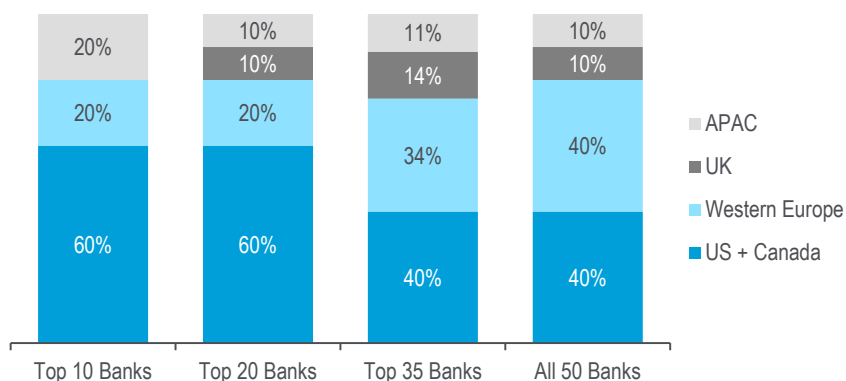
But GenAI spending is growing fast. In the first quarter this year, Accenture reported more than \$600 million in new GenAI bookings alone<sup>7</sup>; Tata Consultancy Services (TCS) has reported \$900 million worth of GenAI and AI sales combined.<sup>8</sup>

Global spending by banks on AI, including software, hardware, and services for AI-centric systems, was estimated at about \$20 billion in 2023.<sup>9</sup> With the global spend on AI across industries estimated to grow at a 25% CAGR between 2023-2026 to \$300 billion, AI-related tech spend by banks will be at about \$40 billion by 2026.

The banks and financial services that claim to heavily invest in AI need to ensure they have an underlying tech stack that supports building out AI use cases, for example a solid data architecture and infrastructure, investment in cloud, etc. AI spending, be it on capex or consultants, will be sub-optimal without getting the foundations correct. It is tough to draw company-level conclusions on current leaders and laggards on AI. A lot of comparative analysis may inevitably be based on proof of talk – for example, press releases and strategy statements – and not actual proof of work.

Evident Insights, a consulting firm, ranks 50 banks globally on parameters such as talent (45% weight), innovation (30% weight), leadership (15% weight), and transparency (10% weight). Their November 2023 Key Findings report ranks North American banks in the lead, occupying 60% of the top ten and top 20 positions.

**Figure 6. North American Banks Lead in the Evident AI Index**



Source: Evident AI Index Banks Key Findings Report – November 2023, Citi Global Insights

<sup>6</sup> Menlo Ventures, 2023: The State of Generative AI in the Enterprise

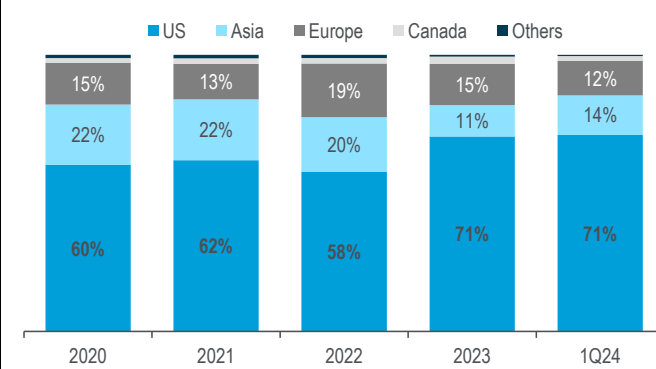
<sup>7</sup> Accenture, FY24 Q2 Earnings 8-K Exhibit, 21 March 2024

<sup>8</sup> CNBC TV18, TCS CEO says AI to be the 'X-factor' on determining how fast clients start spending, 15 April 2024

<sup>9</sup> International Data Corporation (IDC), Worldwide Spending on AI-Centric Systems Forecast to Reach \$154 Billion in 2023, 07 March 2023.

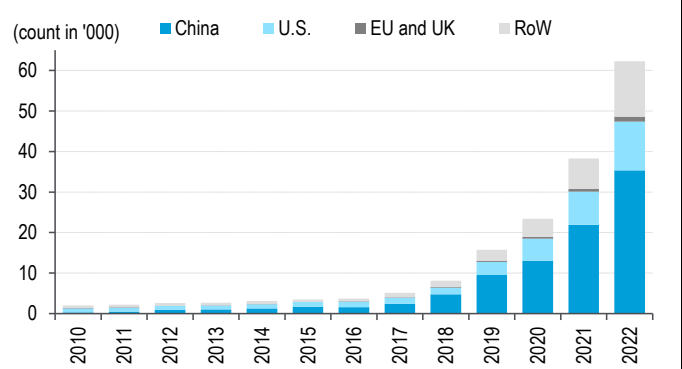
The US banks lead is echoed in the US dominance of VC-backed funding for AI-related startups (Figure 7). This dominance can be attributed to various factors, including established ecosystems such as Silicon Valley.

**Figure 7. Global VC-Funding to AI-Related Companies by Region, 2020-2024<sup>10</sup>**



Note: Calculations based on US dollar VC funding to AI-related firms by region.  
Source: CB Insights Research, Citi Global Insights

**Figure 8. Number of AI-Related Patents Granted by Country/Region**



Source: Center for Security and Emerging Technology, Stanford Human-Centered Artificial Intelligence, Artificial Intelligence Index Report 2024, Citi Global Insights

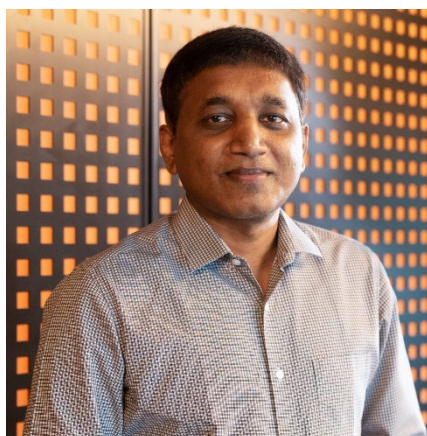
The rising interest in AI has spurred a surge in research and development within the field. China leads the way, accounting for more than half of all patents globally in 2022 (Figure 8).

The explosion in the number of AI patents may not necessarily translate to adoption of enterprise AI/GenAI in financial services.

The distribution of patents granted in China is heavily skewed toward government entities, academic institutions, and universities. By contrast, patents in the US are primarily concentrated among BigTechs. The EU and UK are relative laggards.

Of course, the explosion in the number of patents may be more indicative of increased interest in AI and may not necessarily translate to an increased daily use of AI technology by enterprises and/or financial institutions yet. Much of what we will discuss in the rest of this report is in the early stages of commercialization.

<sup>10</sup> CB Insights Research, "State of AI Q1'24 Report", 08 May 2024  
([www.cbinsights.com/research](http://www.cbinsights.com/research))



Sandeep Alur is a Director at Microsoft Technology Center, an Experience Center dedicated to empowering customers on their digital transformation journey.

## Unlocking the True Promise of AI – Expert Interview

***Q: We have had machine learning and AI in different forms in business and finance for decades. Why has the interest and adoption scaled at this time?***

I have been in this field for 25 years, but I have not experienced this kind of momentum. AI is not new, so why is this time so different? ChatGPT brought UI to AI and made it real for everybody – individuals, institutions, corporates, senior leaders, and even policymakers.

Since the launch of ChatGPT, everyone with a consumer device and Internet can experience AI technology at scale. This is the reason everyone is more enthused about AI this time around, and the hope that AI will transform businesses is a lot more real and tangible today.

***Q: What were the building blocks that preceded the ChatGPT moment? Why is this time different than the previous AI waves?***

As with any emerging technology, AI went through a hype phase where it was considered a futuristic magic pill, to today becoming a reality.

Innovations around vision and speech have progressed and scaled relatively fast. The combination of innovation in software/algorithms and hardware has made the current wave of AI different than previous ones.

The surge in compute power has been a cornerstone in propelling the current AI wave to unprecedented heights. According to data from OpenAI, compute power has been doubling roughly every 100 days since 2012.

***Q: What are some of the immediate and biggest enterprise use cases of AI?***

I would say banking, securities business, financial services, and insurance are ahead in terms of thinking about use cases that matter most to them. Within financial services, the immediate applications could be around customer service & engagement and software & code.

Next is retail, largely because there is so much of a consumer touchpoint with retail. There are ways we can make things better, faster, and engage more with customers by using AI in retail. Healthcare is quite cautious because of regulatory compliance, privacy concerns, etc. But healthcare would be the biggest beneficiary of innovation.

***Q: What are the potential roadblocks as we put AI in finance into production?***

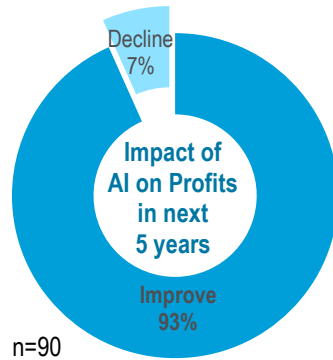
Banks and financial services firms need to make AI governance mainstream. AI model explicability is a key factor to prove how banking decisions were made using AI tools. Regulators want to ensure a human is finally responsible for any decision that has been made for any banking or financial services business.

Banks must also ensure responsible and ethical use of AI. I've not found many financial services firms where you have a chief ethics officer. How do you build and train models, who governs these model outputs, are there any bias in the models, is there any cyber-risk involved in using AI, are we helping or harming the consumer using AI?

## Role of AI in Transforming Finance

### [1] Does AI Create Value for Shareholders?

**Figure 9. Impact from Adoption of AI on Profitability Over the Next 5 years**



Number of respondents =90 including banks, insurers, and asset managers.  
Source: Citi Treasury & Trade Solutions Survey 2024

Senior bankers expect an increase in profitability from adoption of AI in the next five years.

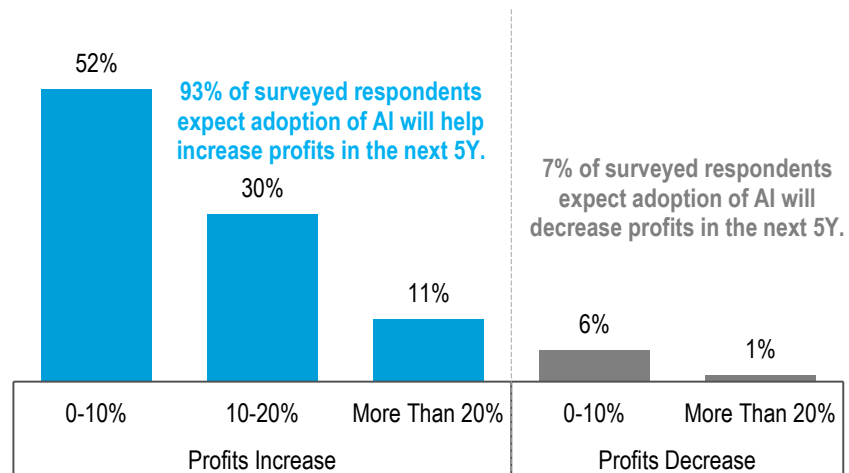
The popularity boom in GenAI has led to a renewed interest by key decision makers in all forms of AI and ML applications. This popularity boom is a double-edged sword, leading to greater focus on the space, likely higher spend, and the risk of a potential overhyping of expectations.

Mobile phones and the Internet helped create many novel business opportunities. Similarly, AI could create new business models, transform processes, generate efficiencies, and more. But the jump from innovation to industrialization, while becoming faster, still takes more time to achieve than creating a PPT presentation.

Management of financial services firms are bullish. Senior bankers expect an increase in profitability from the adoption of AI. Findings from the Citi TTS survey suggest an overwhelming majority of the respondents (93%) expect an increase in profit in the next 5 years (Figure 10).

We discuss AI use cases in finance and the potential resulting productivity boost in greater detail later in the chapter.

**Figure 10. Majority of the Citi Financial Institution Clients Surveyed Expect AI to Drive an Increase in Profits Over the Next 5 Years**

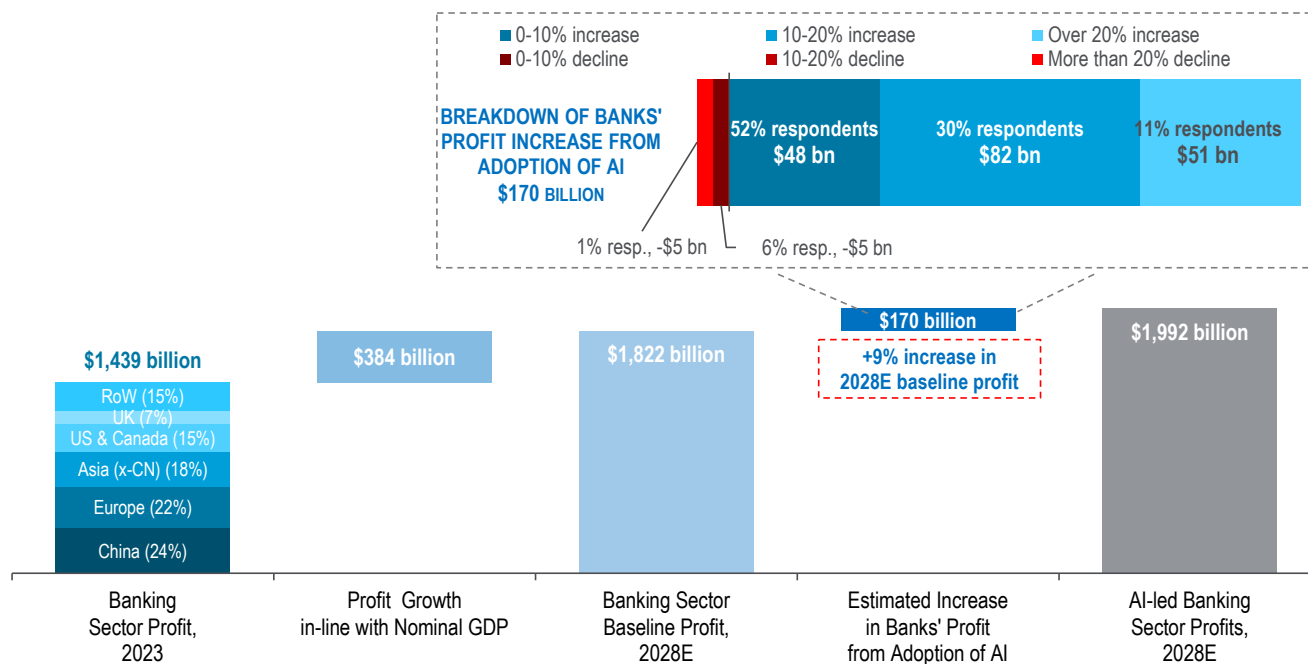


Note: N = 90 includes banks, insurance companies, and asset managers.  
Source: Citi Treasury and Trade Solutions Survey 2024

Based on the findings from Citi TTS survey, we estimate the **potential incremental net profit from adoption of AI in global banking could be around \$170 billion in the next 5 years**. This represents a 9% increase in profits over baseline 2028E profit estimates (Figure 11).

Given the productivity gains touted for AI use cases in finance, especially from the ability of increasing the output of knowledge workers in a white-collar dominated sector such as banking, a 9% profit boost may seem a reasonable forecast. But we would caution that activity levels and competition may also change dramatically.

Figure 11. Adoption of AI Could Add \$170 billion or 9% to the Global Banking Sector Profit Pool by 2028E



Notes: [1] Global banking sector profit sourced from SNL Financial, represents sum of over 3,700 diversified/regional commercial banks, thrift & savings banks FY2023 data.

[2] Profit growth 2023-28E estimated in-line with nominal GDP growth (2024: 4.5%; 2025: 4.8%; 2026: 5.0%; 2027: 4.9%; and 2028: 4.9% (IMF World Economic Outlook).

[3] Increase in banks' profit from AI adoption based on weighted average findings from survey (Figure 10). N = 90 includes banks, insurance companies, & asset managers.

Source: SNL Financial, IMF World Economic Outlook, Citi Treasury and Trade Solutions Survey 2024, Citi Global Insights

### Activity Levels will Explode, Due to GenAI

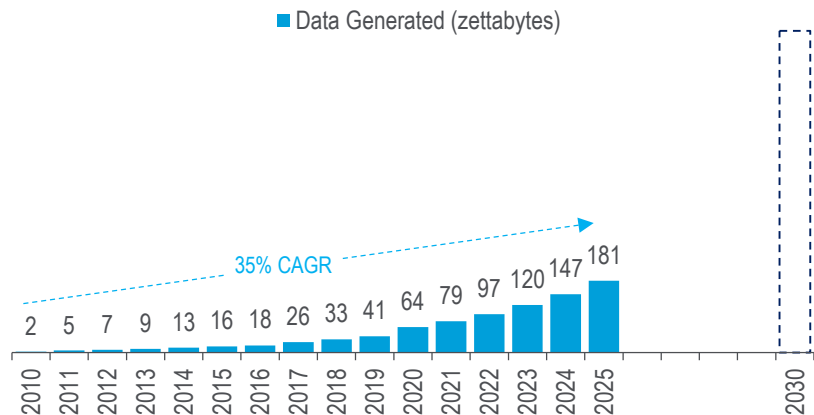
Yes, AI will enable service providers, including banks and finance companies, to become more productive. But remember clients, competitors, and regulators will all also have AI in their toolkit. And many clients may adopt AI faster than the banks themselves, especially if they work in digitally native sectors.

AI/GenAI will lead to a new explosion of data and information. An almost incomprehensible amount of data is created every day. In 2023, the amount of data created globally was estimated at 120 zettabytes (ZB) compared to 64 ZB of data in 2020. 1 zettabyte (ZB) is equivalent to 1 trillion gigabytes (GB).

Estimates suggest global data creation in 2025 could increase to nearly 181 ZB, representing an increase of nearly 51% over 2023. Since 2010, the amount of data created in the world has grown at a CAGR of 35%. With the increasing use of GenAI, even faster growth rates of data creation are likely.

AI will likely bring productivity gains, but activity levels will grow exponentially.

As the amount of data in the world explodes, the GenAI-powered bank or finance firm employee will have to deal with even more information requests from their clients and regulators that will also be GenAI-powered. There will likely be an arms race between AI generating productivity gains and creating more activity.

Figure 12. Amount of Data Generated Each Year Has Grown 60x Since 2010<sup>11</sup>

Note: [1] Refers to the amount of data that is newly generated, captured, copied, or consumed.

[2] 1 zettabyte (ZB) is equivalent to 1 billion terabyte (TB) or 1 trillion gigabytes (GB).

Source: Exploding Topics (Statista, Bernard Marr & Co.) (for data from 2010 to 2025)

The importance of regulation in finance will also slow productivity gains from AI. As we discuss later in the report, policy and regulatory approaches vary a lot between countries. But in most countries, banking and finance are highly regulated industries and the adoption of AI by firms in this space will be heavily scrutinized.

Furthermore, regulators are likely to make increased information and analytical demands from financial services companies using AI themselves. Regulators will likely be able to take data from banks and ingest it into their AI tools, scanning for any red flags, inconsistencies, and potential deviations from guidelines.

This may mean finance firms do not witness material decline in headcount as the volume of activity soars. And any gains made due to a reduction in content and coding-related headcount may be partially or completely offset by an increase in the headcount for AI-related compliance managers and ethics and governance staff.

### Clients Powered by AI Could Change Behavior and Profits

The stickiness in client relationships helps support revenue and profit margins in consumer banking which, especially for market leaders with scale economies, can be highly attractive. Consumer and SME clients are typically price takers and can exhibit inertia in their banking behavior, which helps bank profits.

The stickiness of consumer banking client relationships remains remarkable. Adults in the US, on average, continue with the same savings and checking accounts for about 17 years. Even members of the younger generation (aged 26-32) have held their checking account (9 years) and savings account (7 years) for a relatively long time.

Sticky client relationships have been attributed to reasons such as: [1] these were the accounts they always had; [2] happy with the customer service; and [3] too much of a hassle to switch.<sup>12</sup> Also, crucially, most consumers do not want to spend time thinking about money – which is usually a means to do something else. The typical client desire not to think about financial services reinforces client inertia.

AI agents could disrupt the customer relationship stickiness that supports bank revenues and profit margins.

<sup>11</sup> Exploding Topics, "Amount of Data Created Daily", 13 December 2023

<sup>12</sup> Bankrate, Survey: Consumers stick with the same checking account for an average of 17 years, 04 January 2022.



AI could transform consumer finance activities, especially where there is currently high friction, and the service has simple rules and guidelines.

Could AI change the current equilibrium? In a bot-to-bot banking world, an AI-powered consumer bot would get to work by a simple voice command from the human: "Check my finances." The consumer support bot could quickly provide a detailed report of the actions it has taken or recommendations.

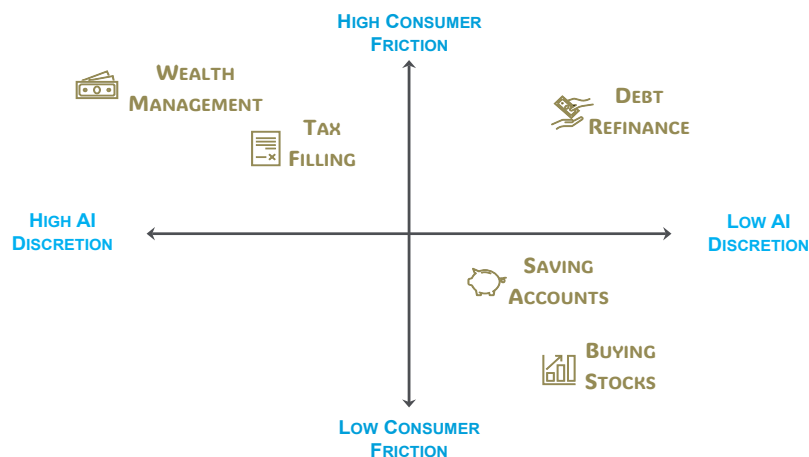
The bot informs you that it has refinanced your mortgage rate, paid off the credit card balance, or negotiated the credit card debt and moved the excess funds to a higher interest account after analyzing the user's historical spending data. Or more likely, the bot informs the clients of actions it recommends and asks for approval.

Consumer clients usually do not want to spend time thinking about their finances. And soon there may be an opportunity to delegate all retail banking and consumer finance tasks to an autonomous AI agent. Of course, humans may still choose not to engage the bot or follow through with its recommendations.

It has been argued that that AI-enabled disruption is likely to happen earlier in the areas where: **[1]** there is high customer friction; and **[2]** the AI is designed to perform a specific task or a limited range of tasks. For example, refinancing mortgage debt is a tedious process for humans but not as complex as tax or wealth management, hence easier for AI agents to disrupt.

Activities with relatively low consumer friction can also suffer from inertia – and these tasks may also move. Moving money to a higher interest savings account or buying single stocks are largely low-friction tasks but still riddled with consumer inertia and could also be disrupted as manual processes become further automated by machines (Figure 13).

Figure 13. Potential for AI Agents to Disrupt Consumer Finance<sup>13</sup>



Source: a16z Consumer, Citi Global Insights

Incumbent banks, especially in the retail and SME segments, make profits from customer inefficiency and inertia. They're protected by high switching costs and lack of client knowledge or interest. Many borrowers either do not know they can save by refinancing, or do not want to go to the effort required to make it happen.

AI-powered finance advisors could, in theory, transform the consumer space by 2030 and we discuss in greater detail later in this chapter. Of course, this will be in the context of local market structures and, importantly, local regulations. These idiosyncratic factors may significantly slow down the pace of adoption of AI agents.

<sup>13</sup> a16z, Money on Autopilot: The Future of AI x Personal Finance, 27 July 2023

## [2] How Does AI Impact the Finance Workforce?

Wider adoption of AI will likely bring productivity gains to the finance sector by automating and augmenting current tasks and roles.



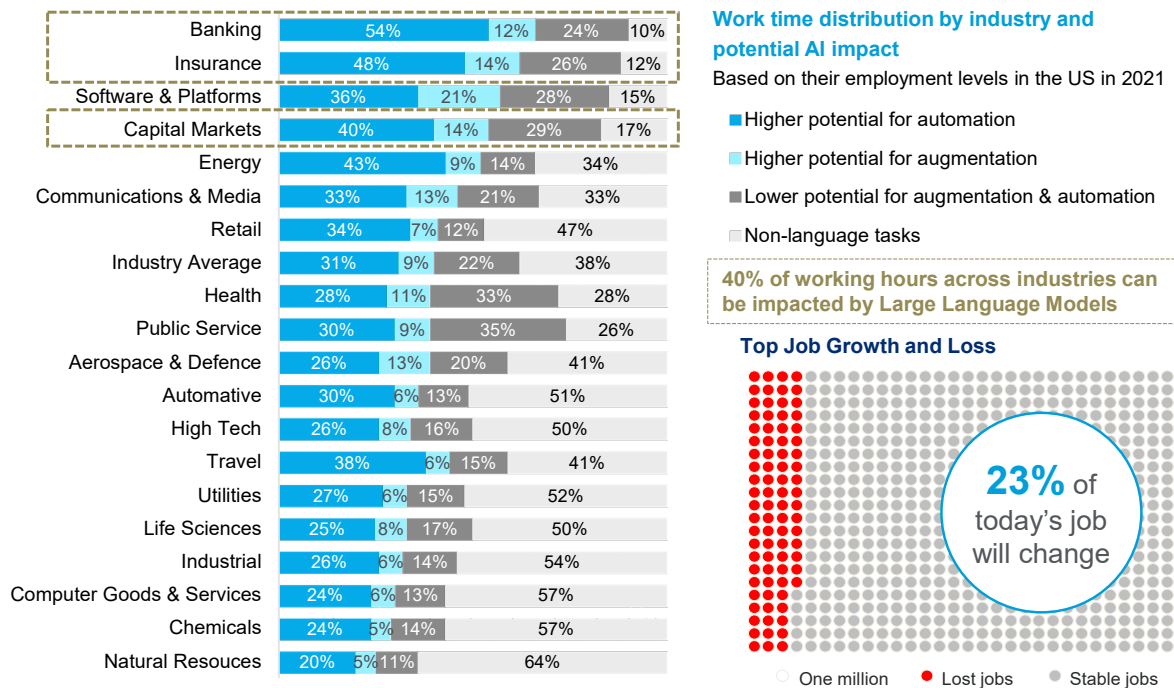
Artificial intelligence is hitting the global labor market "like a tsunami". We have very little time to get people and businesses ready for it.

– KRISTALINA GEORGIEVA, MANAGING DIRECTOR, INTERNATIONAL MONETARY FUND<sup>14</sup>



According to Accenture, nearly **two-thirds of all work done in banking and insurance has high potential for AI-driven automation or augmentation**.<sup>15</sup> This is greater than the overall economy average of 40%. According to World Economic Forum, **nearly a quarter of all jobs globally will change in the next five years**.<sup>16</sup>

Figure 14. Potential for AI-Led Job Displacement is Highest Amongst Financial Services



Note: Details for work-time distribution by industry and potential AI impact sourced from Accenture Research (A new era of generative AI for everyone, 2023). Details on top job growth and loss sourced from World Economic Forum (Future of Jobs Report 2023).

Source: Accenture Research, World Economic Forum, Citi Global Insights

Is AI a level-setter? Results from a 2023 study by Harvard Business School and Boston Consulting Group suggest that while all employees benefited when presented with an AI tool (GPT-4), those scoring below average in the assessment task exhibited 43% improvement in their performance while those scoring above average in the assessment task exhibited 17% improvement in their performance. This suggests AI tools could likely help reduce performance disparities<sup>17</sup>.

<sup>14</sup> Reuters, AI Hitting Labor Forces Like a "Tsunami" – IMF Chief, 14 May 2024

<sup>15</sup> Accenture, A New Era of Generative AI for Everyone, 2023

<sup>16</sup> World Economic Forum, Future of Jobs Report, May 2023

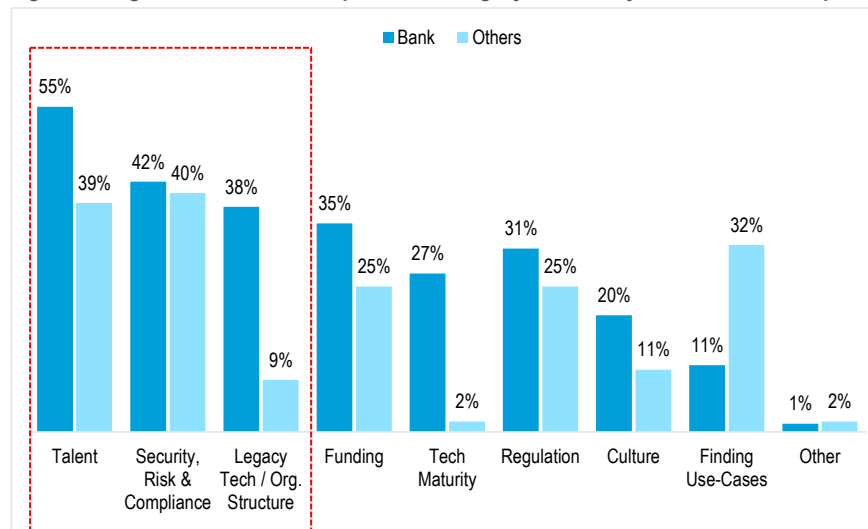
<sup>17</sup> Harvard Business School, Navigating the Jagged Technological Frontier, 2023

## Building Resilience Through Reskilling

Technological innovation and resulting disruptions fundamentally alter the workforce mix over time. The further integration of AI into finance in the coming years will speed up this transformation.

In the Citi TTS client survey, lack of availability of the right (AI) talent was cited by 55% of the responding banks as a primary barrier in their AI journey, followed by risk and compliance (42%) and legacy technology (38%). Hence, it is important that we train our existing workforce to be flexible and adaptable.

**Figure 15. Right Talent, Risk & Compliance and Legacy Tech is Major Barrier for AI Adoption**



Note: N=71 Banks; 57=Others. Others include asset managers, FinTech, insurance companies.

Source: Citi Treasury and Trade Solutions Survey 2024

As most organizations are struggling to find the right AI talent, many are looking to balance the demand-supply gap by upskilling, reskilling, and redeploying talent internally or from within the industry. A recent McKinsey & Company survey highlights 56% of its sample are likely to hire GenAI talent internally versus 30% who would hire externally.<sup>18</sup> The biggest source of AI talent into banks is still another bank.<sup>19</sup>

In 3-5 years, all teams and tasks will likely have embedded AI colleagues. Skills such as critical & analytical thinking, creativity, and ethics will be in demand.

Our recent report, [Citi GPS: What Machines Can't Master – Human Skills to Thrive in the Age of AI](#), argues in 3-5 years from now, all teams and tasks will likely have embedded AI as colleagues and hence human skills such as critical and analytical thinking, creativity, and ethics will be in greater demand.<sup>20</sup>

Early users of AI tools have often found the quality of output generated by AI tools unreliable and inconsistent.<sup>21</sup> Banks would be subject to financial penalty and/or reputational harm if AI bots give out incorrect price quotes or engage in unreliable behavior. A human checker or supervisor will be required to police the AI bots.

<sup>18</sup> McKinsey & Company, The Human Side of Generative AI: Creating a Path to Productivity, 18 March 2024

<sup>19</sup> Evident AI, Key Findings Report, November 2023

<sup>20</sup> Citi GPS, What Machines Can't Master – Human Skills to Thrive in the Age of AI, 31 May 2024

<sup>21</sup> The New York Times, Press Pause on the Silicon Valley Hype Machine, 15 May 2024



As AI technology advances, developers are likely to become cheaper and more expendable. This is likely to be true even for bank employees. The nature of jobs will change, with basic roles being replaced by technology. Those with the best ideas will thrive, earning substantial rewards, whilst others face displacement.

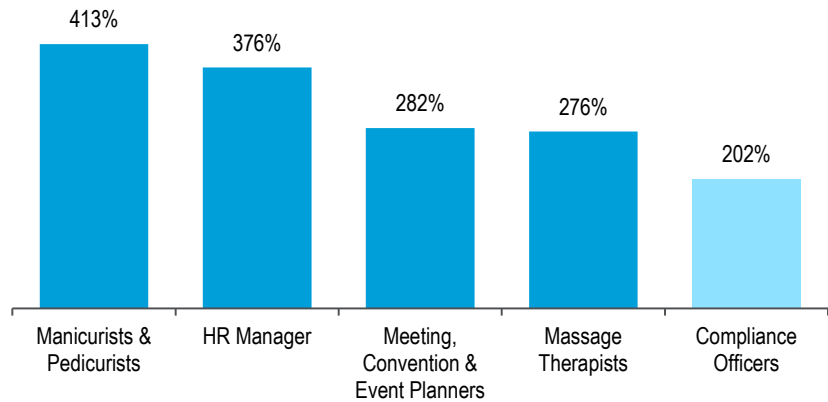
– ERKIN ADYLOV, CEO & FOUNDER, BEHAVOX



In the evolving workplace that is driven by AI, every employee will inevitably take on managerial responsibilities as they engage with AI-driven tools. We will see a rise in AI guardrail officers or some form of AI compliance officer, a role that more broadly has already grown 3x in the last two decades (Figure 16).

In the AI-enabled workplace, every employee will inevitably take on managerial responsibilities.

Figure 16. Fastest-Growing Occupations in Last Two Decades (% Change from 2000 to 2023)



Source: U.S. Bureau of Labor Statistics, Citi Global Insights

As AI systems generate significant portions of the output driven by client and/or regulator demand, employees must oversee and refine this output, ensuring it aligns with the local/global laws, regulatory standards & guidelines, and organizational goals & objectives.

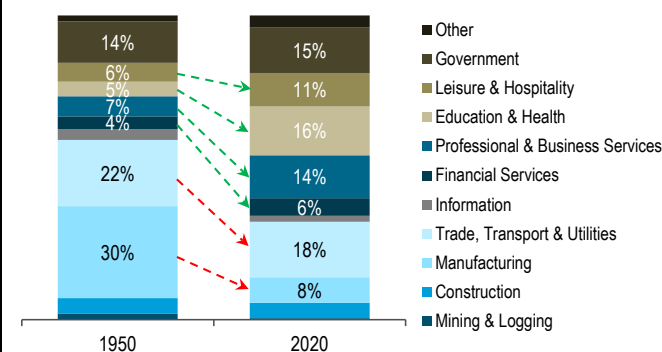
### Historical Impact of Technology on Jobs

Employment across industries has seen significant change in the past decades. In 1950, three out of ten US jobs were in manufacturing (30.2%). This fell to one out of every 12 US jobs by 2020 (8.4%). By contrast, the share of jobs in finance grew 40% from 1950 (4.1% of total jobs) to 2020 (5.8% of total jobs).

Even during the 2010's tech revolution, and in the aftermath of the 2008 subprime crisis, US financial jobs as a percentage of total employees in the economy remained surprisingly stable at 6%, while US commercial bank employees as a percentage of total finance jobs declined slightly to 15% (2023) from 18% (2011).<sup>22</sup>

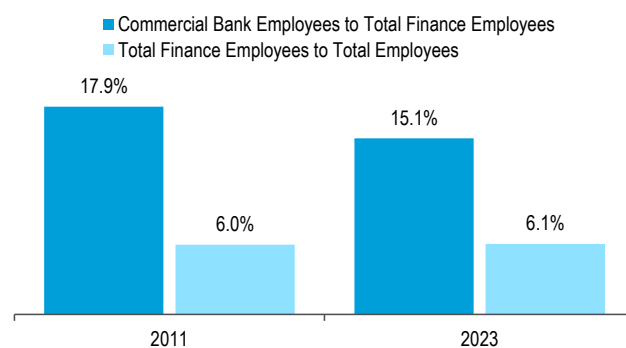
<sup>22</sup> Federal Reserve Bank of St. Louis and the U.S. Bureau of Labor Statistics

Figure 17. Change in Jobs in the US from 1950 to 2020, By Sector



Source: U.S. Bureau of Labor Statistics, Citi Global Insights

Figure 18. Bank &amp; Finance Jobs Remain Stable Over the Last Decade



Source: Federal Reserve Bank of St. Louis, U.S. Bureau of Labor Statistics

Historical adoption of new technology has not led to rapid changes in the finance sector workforce.

AI-led disruption of jobs and tasks are likely to **change the nature of the workforce**. While GenAI will likely change individual finance jobs and tasks, and how they are staffed and managed, it is interesting to note that historically adoption of new technology has not led to rapid changes in the finance sector workforce.

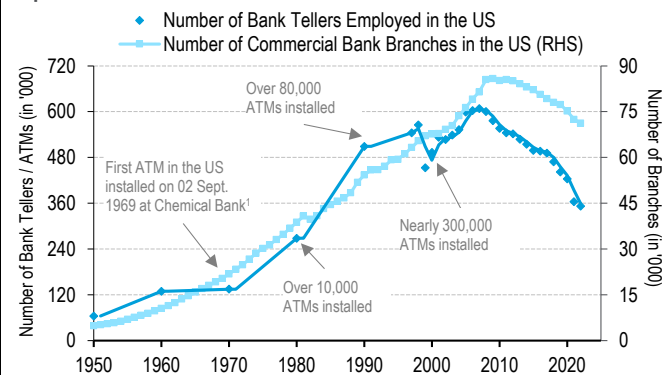
For example, the introduction of ATMs starting in the late 1960s did not lead to a drop in the number of human tellers employed by banks. On the contrary, between the 1970s and mid-2000s, the number of human tellers employed soared, as the US economy and the financial sector grew rapidly.

Similarly, the advent of spreadsheet tools (e.g., VisiCalc, SuperCalc) did not lead to a drop in bookkeeping and accounting jobs. Instead, the new tools helped accountants work smarter allowing them to add value, as the software could do tasks in seconds that would take hours for humans to complete.

Accounting and financial management jobs became more valuable for businesses as they could now focus on analysis, forecasting, and strategy – and not just numerical compilation tasks. The growth of financial analysis tools created a new growing class of financial managers from the 1980s onwards.

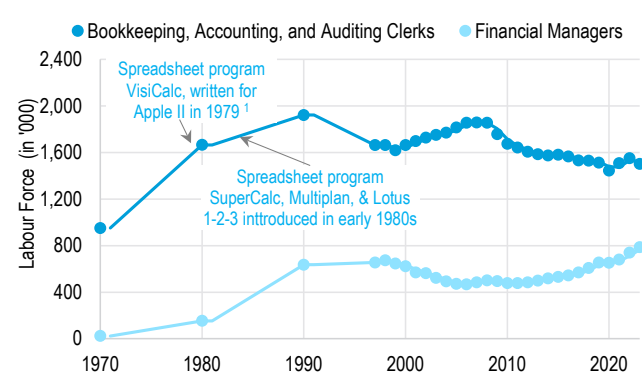
Over the next decade, look out for the rise of AI managers and similar roles.

Figure 19. Introduction of ATMs Did Not Have Immediate and High Impact on Bank Teller Jobs



<sup>1</sup> American Banker, Kevin Wack, Alan Kline, "The Evolution of the ATM", May 2017  
Source: FDIC, U.S. Bureau of Labor Statistics, Citi Global Insights

Figure 20. Advent of Spreadsheets Did Not Reduce the Number of Accountant Roles



Note: [a.] Bookkeeping, accounting, and auditing clerks produce financial records for organizations & check financial records for accuracy.  
[b.] Financial managers create financial reports, direct investment activities, and develop plans for the long-term financial goals of their organization.  
<sup>1</sup> Encyclopedia Britannica, "Spreadsheet", last updated on 23 May 2023  
Source: U.S. Bureau of Labor Statistics, Citi Global Insights

### [3] Does AI Create Value for Banks' Clients?

AI will create more value for banking and wealth management clients, whilst edge cases in insurance likely face insurability.

AI is likely to transform banking from the inside and the outside. From the inside, it is automating repetitive tasks and augmenting information heavy tasks, thus bringing productivity benefits (value to employees, potentially) which may translate into more profits (value to shareholders, potentially).

On the outside, it promises to enhance customer experience through AI agents (negotiating price on our behalf and managing administrative matters), better customer service (chatbot 2.0), and tailored offers (hyper-personalization). Banking and wealth management clients should be net winners from this.

However, AI could also have some negative impacts on some finance clients, especially edge cases in insurance. Access to insurance could be limited, or premiums could rise significantly, for higher-risk individuals or properties where big data insights or AI analysis identify greater risks. Regulation will be needed to manage some of these negative side effects of AI.

FinTech's promise over the past decade has been to create more options, better prices, greater convenience, and a better experience for consumers. AI could be the latest wave of technology to turbo-charge this promise. It could further empower bank clients, especially consumer and SME clients.

The future of retail and corporate banking is likely to be more AI and data driven than before. We believe banks' customers (retail, SMEs, and corporate) are also likely be AI-enabled in the future, not just banks and finance firms. This could tilt the playing field, including for prices, a bit more toward consumers.



Generative AI, with proper guardrails, can effectively decipher meaning and provide comprehensive answers. Integrating domain specific LLMs into banking will be transformative.

– ZOR GORELOV, CO-FOUNDER AND CEO, KASISTO



### Autonomous Agents

The next frontier for AI innovation is likely to be autonomous agents. Instead of prompting the LLMs with simple one-line instructions, BigTech companies could embed higher AI capabilities in their existing digital assistant suite such as Apple's Siri, Amazon's Alexa, and Google Assistant.

With many BigTech companies foraying into financial services, consumers are likely to delegate financial decisions to these AI assistants. For example, Google's Bard can process an investor's existing investment portfolio and calculate returns relative to the broader market. This functionality is likely to get more sophisticated over time and could involve partnering with different bank and FinTech partners.<sup>23</sup>

These bots, equipped with sophisticated algorithms and access to vast amounts of data, will negotiate with financial institutions to secure the best possible deals for their users. This shift will not only streamline financial services but also ensure that decisions are made with a level of precision and foresight that human clients may not achieve on their own. So, what will AI-powered retail banking look like?

<sup>23</sup> Medium, How Google Bard Helps Investors, 09 December 2023



The next frontier in innovation will be clients using AI bots to manage their finances.

- **AI-Enabled Bank Customer:** As AI technology advances and wider adoption of autonomous AI agents proliferate, consumers will likely be enabled to make informed financial decisions without them being directly involved, especially in data gathering, comparison shopping, and carrying out manual tasks to execute the transaction. Consumers can focus on the yes/no/switch decisions.
- **Enhanced Financial Health:** More than three-quarters of Americans are worried about their financial situation, including insufficient funds for retirement (68%), an inability to manage the cost of living (56%), and debt levels (45%).<sup>24</sup> The concept of open banking and open finance will allow consumer banking bots to access consumers' financial data, making real-time decisions to optimize their financial well-being.
- **Right Bot vs. Right Bank:** In this new paradigm, the critical decision for consumers will be selecting the right bot rather than the right bank. Choosing bot-powered advisors, much like choosing human personal finance advisors, will become a key task. But who will the bots work for? Advisory firms, banks, FinTech, BigTech, or some genre of new service?
- **Will Everyone Have a Bot?** Leading banks and financial firms will most likely opt to provide their own AI-powered personal finance services. But they will also be joined by FinTech and BigTech firms that may have a competitive advantage in terms of being digitally native and have faster go-to-market speed. In some markets, they may also have stronger consumer brands.
- **How Will We Choose Our Bot?** A universal truth is that most consumers do not want to spend a lot of time on financial services topics. Finance is a means to an end – the mortgage is a means to living in your home. Will consumers want to spend too much time thinking about which autonomous agent to use? Will we default to known and trusted brands? Albeit these may not be bank brands!
- **The Squeezed Middle?** Many smaller banks, for example community banks, will continue to base their client proposition on local market knowledge and client proximity. In wealth management and private banking, similar to institutional banking, the human in the loop will remain crucial, albeit they will be increasingly AI-augmented. As always, the players in the middle will get squeezed

#### Real Life Application of a Customer Bot to Negotiate Debt Collection

Credit card debt in the US is at all-time high of \$1.13 trillion at the end of 2023 and 25% of those with credit card debt are paying less toward it than they should, leading to a cycle of revolving credit card debt.<sup>25</sup>

Cambio offers an AI-powered service that negotiates debt collections on behalf of consumers. In the first 60 days of launch, the AI-powered consumer bot helped around 70% of customers resolve their collections and raise their credit score.

<sup>24</sup> CNBC Select, 77% of Americans are anxious about their financial situation, 20 May 2024

<sup>25</sup> CNBC, Nearly 1 in 4 Americans with debt are putting less money toward credit card payments, 22 February 2024

## The Bot Frontier – Expert Interview

**Q: Why do we need bots to do banking? How will it benefit bank customers?**



David G. W. Birch is an author, advisor, and commentator on digital financial services. David is Principal at 15Mb Ltd., his advisory practice; and Global Ambassador for Consult Hyperion, the secure electronic transactions consultancy; and holds several other board-level advisory roles.

**David Birch:** I see significant scope for AI-led transformation in banking and finance. For example, tasks like opening an individual savings account (ISA) can be a complex and time-consuming task, even for someone with a financial advisor. It involves extensive analysis of different types of ISA's, their cost structures, and offered benefits, etc. Even a basic bot could help make better and faster decisions.

However, in retail, we are likely to see discontinuity as traditional marketing strategies will not work on bots. For instance, a bot is not influenced by brand loyalty or advertisements; it operates within the model parameters on APIs.

The concept of value will be critical. Customers are likely to deal with bots that reflect their values and interests. This could lead to the creation of a marketplace for bots where individuals select a bot from a range of options based on desired preferences.

**Kirsty Rutter:** The fundamental premise behind implementation of AI-driven client interfaces is that if technology exists, why should humans waste time on petty administrative tasks. Instead, executive bots can perform these tasks on our behalf.

In a financial management context, an AI bot can directly engage with the bank on behalf of the customer. For example, if a mortgage is up for renewal, the bot can notify the customer of the impending end date, conduct research to find best deals, determine the right product, and fulfill formalities for refixing the mortgage rate.

Bots can achieve this based on set parameters (retirement age, preference for debt tenure, maximum acceptable price, etc.), pre-determined by the customer. In the initial phase, the bot could rely on human confirmation before executing tasks; but as trust in the bot grows, several of these tasks can be handled autonomously.

**Q: How will banking products and services change to serve AI bots? How will bank customers having autonomous bots affect banks' profit pool?**



Kirsty Rutter is the FinTech Investment Director at Lloyds Banking Group. In the past, she has held several strategic roles in prominent institutions like Barclays Capital, Credit Suisse, Eaton Risk Advisory, and the Ministry of Justice UK.

**Kirsty Rutter:** The banking products will have to be designed for AI bots but need to be codified in a manner humans can comprehend. Consumers must be able to understand and interpret the product or service. Products need to be reliably designed with the appropriate parameters for the intended consumer – whether human or bot – for this to be true the controls and compliance requirements of any product must be designed in from the beginning. With this outcome in mind, products must be designed primarily for bots, as they are the target users, but they must also be translated in a way that makes them comprehensible to humans.

**David Birch:** There is a significant difference between designing a product/service for people and designing one for bots. For starters, products/services intended for bots can be more complex and sophisticated than those for humans, as bots can better comprehend and manage these complexities.

Further, the impact of using bots for financial products/services could have a substantial impact on banks' profitability. For example, a bot managing a savings account could shift funds to optimize interest rates far more frequently than a human ordinarily would.

This rapid movement of money, further catalyzed by open banking initiatives, could dramatically impact banks' liquidity management as a large number of small

balances could move from one bank to another and impact profitability. Thus, the strategic implications are substantial and intriguing, necessitating immediate attention, despite the seemingly distant horizon.

***Q: How far are we from having bots represent us in choosing financial products and services?***

**Kirsty Rutter:** For executing transactions through personal bots, several conditions must be met. This includes cloud-native technologies with embedded controls and regulations. These are critical to show regulators that decisions/outcomes are fair and responsible. Achieving this level of sophistication requires continuous rethinking and retraining – necessitating the combination of product design, data engineering, risk and compliance, and technological savviness. While the design is not inherently complex, it necessitates a significant shift in competencies and capabilities.

**David Birch:** This change cannot be achieved overnight, rather it needs to be part of the banks' strategic planning for the next three to five years.

In my view, the timeline is primarily determined by regulatory developments. As for technology, we already have some of the basic components needed. For example, leveraging ChatGPT plugins and UK's Open Banking regime, it is possible to use ChatGPT to interact with your bank account, even today.

However, the regulatory framework needs careful consideration. We, as an industry, need to define the appropriate regulatory measures for such advancements. While the technology exists, practical regulations for bots will take time.

***Q: Since bots lack loyalty/emotions, what will banks compete on to secure clients' wallet? What are their competitive advantages in a bot-driven world?***

**David Birch:** Banks will likely need to compete on price to retain clients' wallet share. Bots can compare multiple prices simultaneously, so being the cheapest will likely always be a significant factor. Banks could adopt an execution strategy focused on operational efficiency, enabling them to compete on volumes and win deals purely based on price.

Additionally, banks will need to optimize their operations for the bot, which might differ from human preferences. For example, bots might prioritize factors like the speed of API response, API uptime, data accuracy, and richness of API data. Optimizing for these factors could provide banks a competitive edge.

**Kirsty Rutter:** Lastly, banks could compete by aligning with the ethical values and personalized frameworks set by the customer's bot – a reflection of the individual's morals and beliefs. Some customers might prioritize price, whilst others might focus on sustainability aspects. In response to the changing demands of the market, there are likely to be new opportunities for products and services to support the hyper personalized outcome many individuals and industry seek.

**Damian Kowalski**

Global Receivables  
Treasury and Trade Solutions  
[damian.kowalski@citi.com](mailto:damian.kowalski@citi.com)

**Faizan Khwaja**

Global Receivables  
Treasury and Trade Solutions  
[faizan.khwaja@citi.com](mailto:faizan.khwaja@citi.com)

**Jiten Mehta**

Global Payments and Receivables  
Treasury and Trade Solutions  
[jiten.mehta@citi.com](mailto:jiten.mehta@citi.com)

**Arash Eskandari**

Global Payments and Receivables  
Treasury and Trade Solutions  
[arash.eskandari@citi.com](mailto:arash.eskandari@citi.com)

**Santiago Dinelli**

Global Payments  
Treasury and Trade Solutions  
[santiago.dinelli@citi.com](mailto:santiago.dinelli@citi.com)

**Mike Lu**

Emerging Payments  
Treasury and Trade Solutions  
[mike.lu@citi.com](mailto:mike.lu@citi.com)

**Corporate Client Bots**

We have largely focused on consumer clients so far. Let us take a closer look at scenarios where corporates deploy more sophisticated AI bots for their banking needs. Thus, the bot reviews the company's liquidity status, forecasts future cash flows, identifies potential investment opportunities, and flags financial risks – all in real-time.

The bot could renegotiate loan terms with banks overnight, optimize the company's investment portfolio, close out any unutilized overdraft facility to strengthen financial positions, and ensure more regular compliance and regulatory reporting tasks, too. This sophisticated AI integration would transform the landscape of corporate banking. Key benefits of corporate bots in banking are:

- **Onboarding:** Customer onboarding (KYC and CDD processes), loan origination, and compliance checks. These processes, which are even more complicated and time-consuming in corporate than retail banking, often involve multiple touchpoints, paperwork, and approval and review cycles. This process requires an ability to efficiently validate, understand, and store content of documents to correctly implement client instructions.

Documents often lack standard formats and are not fully digitized. Figure 21 shows two processes. On the left-hand side, a process using established technology is shown, with automation implemented to orchestrate a process. However, once the orchestration completes, a manual process is required. This contrasts with the image on the right-hand side where AI is deployed to process most cases, leaving a few outliers that require manual intervention.

- **Feedback Loop:** The true voice of corporate clients is often obscured by limited feedback channels, rigid surveys, and generic market research. GenAI unlocks novel approaches as it empowers clients to express feedback in more natural ways and move away from static forms. This enables AI to extract insights that truly illuminate hidden needs and preferences, which fuels a virtuous cycle where client feedback directly shapes transformative, highly tailored products.

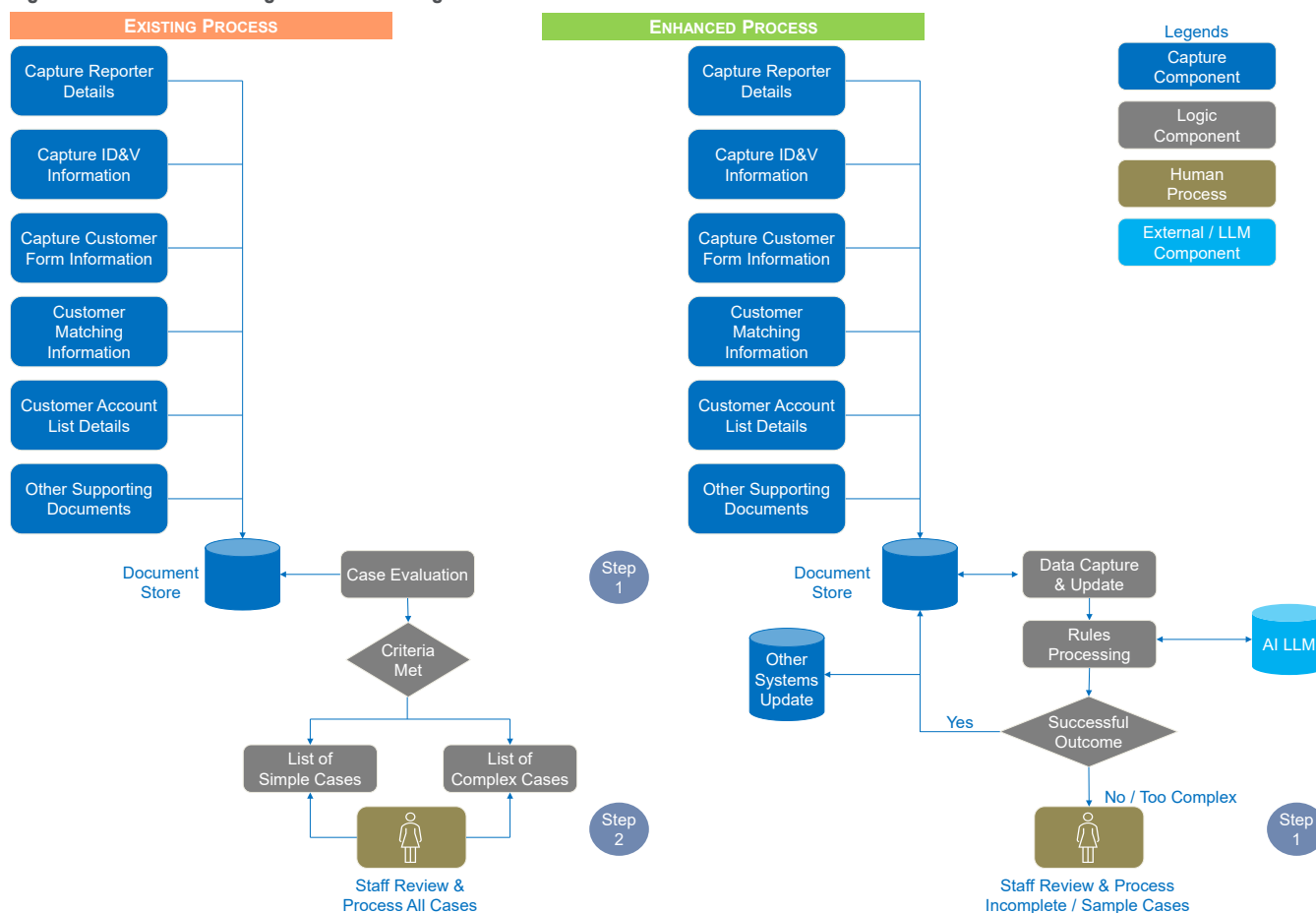
For example: a client states the original intent to transfer funds and receives guidance how to best complete, then finishes a wire transfer and the chatbot asks, "How could we make this process even easier next time?" The client's detailed response is far more valuable than a 5-star rating.

- **Querying Own Payments Data:** Clients will be able to effortlessly interact with their data through a GenAI interface without the need for them to be proficient in data manipulation tools or be confined to a set of predefined data analysis features within an application.

Instead, they have the flexibility to ask questions in natural language, such as, 'Who are my top suppliers and customers, based on payment value and volume, segmented by geography and payment methods?' The responses go beyond mere data; they are intuitive insights in natural language, accompanied by visualizations and the capability to explore 'what if' hypothetical scenarios. The system provides text summaries that dynamically update as the data evolves, offering a seamless and accessible experience.

- **Regulatory & Compliance Insights:** Graph LLM-powered tools could automate the monitoring and analysis of regulatory and compliance changes with the help of natural language processing algorithms. Relevant updates identified by such tools could then be used not only by the bank to adapt its processes and policies accordingly, but also form part of the advisory information that clients can benefit from while incorporating changes into their business. Banks' transactional processing flows would also benefit from ongoing compliance checking against such consolidated regulatory positions, flagging non-compliance early and allowing both the corporate and the bank to address the gaps and avoid penalties.

Figure 21. Client Onboarding Process: Existing Process vs AI Enhanced Process



Source: Citi Treasury and Trade Solutions

There are some drawbacks to letting AI do banking on your behalf. The extensive use of AI and open banking APIs raises concerns about data security and privacy. Ensuring that sensitive financial data is protected from cyber threats is paramount.

Over-reliance on AI bots may lead to vulnerabilities if the technology fails or is compromised. Navigating the complex regulatory landscape can be challenging, especially as regulations evolve to keep pace with technological advancements. Ensuring compliance across multiple jurisdictions may require significant effort and resources.

GenAI will create new opportunities beyond productivity improvements, but some of the more blue-skies work will likely take time to build...

...business models will need to adapt, regulations need updating, and skills must evolve.

## (B.) How Will AI Be Used in Finance?

For many years, AI, or strictly speaking ML, has been used in finance on structured data and for quantitative tasks. Today, AI is used in finance primarily for risk and pricing. GenAI will expand use cases to new areas.

In the short to medium term, we expect the biggest impact at incumbent financial institutions to be on internal facing tasks and improvements in productivity rather than lots of new products. Incumbents will focus on improvements in areas such as software & coding, transaction monitoring & compliance, and more.

A lot of bank functions such as credit underwriting, algorithmic trading, portfolio construction, and transaction monitoring already utilise AI/deep learning applications. GenAI will create new opportunities beyond productivity improvements but some of the more blue-skies work – newer products & services, bots using tokenized money, and decentralized AI – will likely take time to build and be rolled out to market.

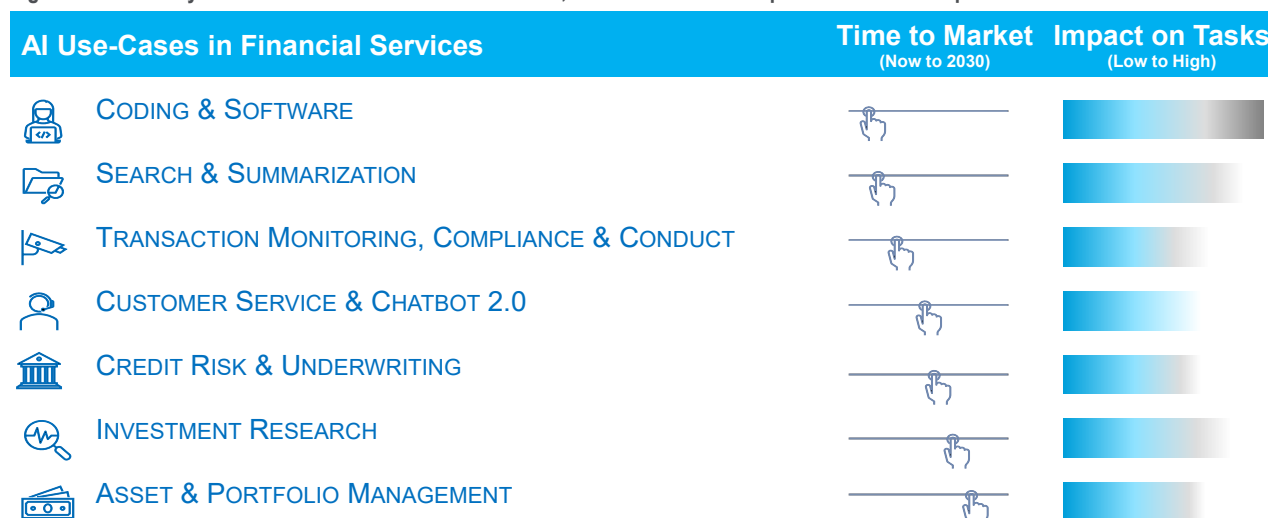
Autonomous AI agents could turbo charge existing business models and commercial relationships and lead to the creation of new ones. Will AI agents drive adoption of digitally native money tokens and lead to another stage in the disruption of financial services? This is quite possible by 2030, and we discuss this in this chapter. But the exact timing of large-scale usage of AI agents is for now unclear.

Business models need to adapt, regulations need to be introduced and/or updated, client behaviour and employee skills need to evolve. The time lag between invention and industrialisation of a technology can be significant. And in the case of financial services, there are the additional challenges of being heavily regulated.

In this chapter, we largely focus on the likely use cases of GenAI that will be used at scale in finance in the next couple of years. Changes to how finance does code and software, or search and summarization, are likely in 2024-25 (Figure 22). Existing ML/AI use cases in compliance and risk management will be GenAI augmented.

Core finance functions such as risk analysis and underwriting will also be changed, as will investment research and portfolio management, albeit more slowly.

Figure 22. Summary of Generative AI Use Cases in Finance, Estimated Wider Adoption & Potential Impact on Tasks



Source: Citi Global Insights



An overview of the main AI/GenAI use cases in finance are below:

- **Coding & Software:** Large banks have tens of thousands of employees for software and programming tasks, often about 15-25% of the banks' total workforce. AI can help streamline coding processes by automating repetitive tasks, optimizing basic coding, and accelerating development cycles.
- **Search & Summarization:** The financial services sector is characterized by its data and document-intensive nature. AI tools can sift through vast volumes of data, distill pertinent information, and deliver concise summaries that can be used as input for faster decision making and executing next actions.
- **Transaction Monitoring, Compliance & Conduct:** AI-powered systems can excel in monitoring external and internal conduct. By continuously analyzing transactions and behavior, and detecting anomalies in real-time, AI mitigates risks, ensures regulatory adherence, and minimizes fraudulent activities.
- **Customer Service & Chatbot 2.0:** AI-powered chatbots could enhance customer service by delivering personalized experiences, resolving queries promptly by having access to customer data, while simulating human-like interaction and providing round-the-clock assistance.
- **Credit Risk & Underwriting:** AI algorithms could leverage vast datasets to assess credit risk and facilitate underwriting processes with potentially better accuracy and speed compared to humans. AI models could make better credit decisions by analyzing diverse and non-traditional data sources.
- **Investment Research:** Fundamental research involves a lot of search and summarization of information, datasets, and generation of text and charts. Generative AI could likely bring a lot of time/cost efficiencies as it could automate a lot of information search and retrieval tasks.
- **Asset & Portfolio Management:** AI could help identify investment opportunities, optimize asset allocations, and help personalize portfolios in a scalable manner thus enabling advisors and portfolio managers to focus more on high-value activities such as client engagement and alpha generation.
- **Decentralized AI:** The extremely centralized nature and economic model of AI stack makes a compelling case for having a co-existing decentralized AI stack which will have a slightly different economic model and would use tokens or tokenized money to compensate cloud providers (for compute) and the model owners.
- **Money & Finance in the Bot World:** We believe humans will increasingly adopt AI bots and agents to carry out financial activities and transactions. Tokenized money's features such as programmability and the smart contract functionality makes it ripe for wider adoption by tech-native AI agents who can operate 24x7 and use tokenized money (e.g., stablecoins, CBDCs) to enable atomic settlement (i.e., instant and simultaneous settlement) in the bot-based financial ecosystem.



Erkin Adylov is the founder and CEO of Behavox and leads product innovation at the company. Prior to founding Behavox in 2014, Erkin worked at Goldman Sachs in equity research and later at GLG Partners/Man Group as a portfolio manager in the Financials Fund.

## AI Scepticism to FOMO – Expert Interview

### ***Q: How has application of AI/ML evolved in finance and compliance?***

In the last five years, AI systems have got sophisticated, led by breakthroughs in AI architecture. Recent AI models can include several trillions of parameters. These advancements enabled the development of scalable AI systems that can be used in compliance, including detection of money laundering.

Compliance practices have also evolved, moving away from a keyword search approach to more intuitive and comprehensive AI solutions that help reduce the number of false positives. Today, the initial skepticism around AI adoption in compliance has been replaced by the fear of missing out.

### ***Q: Under what circumstances is it preferable to build a foundation model as opposed to employing a retrieval augmented generation approach for clients?***

Retrieval Augmented Generation (RAG) is one of the applications of LLMs that leverage custom datasets. They are most relevant in cases where information retrieval from a specific knowledge base is key, as opposed to generating content. For instance, retrieving and analyze information from compliance manuals.

In instances requiring high volumes (e.g., 20 million queries a day), classification could be a more efficient alternative. For this, a custom foundation model must be built and then fine-tuned to specific needs. This approach is most relevant for finance as it ensures compliance with regulatory requirements, including auditability.

### ***Q: Will the future of AI be shaped by large general-purpose models or by smaller domain-specific models?***

Both large general-purpose models and smaller domain-specific models will play a significant role in shaping the future of AI. Domain-specific models will deliver superior performance to specific industries they are purpose-built for. Meanwhile, large models such as GPT-5, GPT-6, GPT-7 will excel in general domain understanding. Eventually, as we progress towards artificial general intelligence, there may come a time when a single model dominates all domains.

### ***Q: What are model performance degradations? Why is it relevant in finance?***

Model performance degradation typically occurs from frequent model changes. Implementing feedback loops or automated training can lead to fluctuations in model performance (for the better or worse). However, it must be noted that such practices are not permissible for regulated financial institutions. Regulatory requirements mandate locking in the model version to prevent degradation.

### ***Q: Will GenAI automate coders or increase the demand for coders?***

As AI technology advances, developers are likely to become cheaper and more expendable. This is likely to be true not only for developers, but all professionals involved in content generation such as bank employees, journalists, equity research analysts, etc. The job market is likely to evolve into a celebrity economy, where only those with the best ideas will thrive, earning substantial rewards, whilst others face displacement.

## Document & Data Management with AI – Expert Interview

### *Q: How smart is AI?*

We often think of AI as a single entity, but it comprises multiple small, specialized models performing different tasks. Think of them like different layers in a tiramisu, stacked atop another. While impressive, AI in its current form is far away from human-level intelligence. LLMs lack the ability to plan, reason, and understand the physical world. In four years, a child will have seen 50 times more data than the biggest LLMs. And, just like a 4-year-old, AI makes mistakes. It requires supervision. Human involvement in AI is crucial, especially in the context of finance.

### *Q: How can AI tools help manage unstructured data in large enterprises?*

80-90% of data in any large enterprise tends to be unstructured (i.e., data found in documents, call transcripts, recordings). This data is not available in spreadsheets or databases and is difficult to quickly summarize by simply glancing through it. Some of the tasks where AI can assist in context of unstructured data include:

- **Metadata Extraction:** AI can help identify and extract metadata such as name and addresses from documents (e.g., extract name or date from passport).
- **Text Extraction:** AI tools can help locate and extract specific text from complex documents like prospectus, offering circulars, agreements, or other repositories.
- **Analytics and Reporting:** AI tools can help analyze extracted data, identify key trends, and generate reports with the aim of enhancing productivity.
- **Obligations Management:** AI can help monitor regulatory or client obligations and alert organizations to any upcoming compliance requirements/fillings.
- **Document Generation:** AI can help gather information and generate documents that are written repeatedly throughout the year or update them from routinely.

### *Q: What are the best practices to ensure accuracy and regulatory compliance when using AI for generating content in financial services?*

Data security, privacy, and compliance with regulatory requirements are critical when using AI in financial services. Secondly, select the right tools for the task. For example, LLMs are popular today, but they are not best suited for all tasks due to their tendency to hallucinate. It is preferable to use a more reliable model (or set of models) where accuracy is critical (e.g., vetting an agreement).

### *Q: What aspect of jobs will be replaced by AI tools? How can professionals approach reskilling? What new opportunities can arise?*

Looking at history, initial concerns about job displacement have been unfounded. This is true with the printing press, steam engine, and others – each time powerful technology becomes available, job creation dwarfs displacement. The advent of AI is likely to follow a similar pattern. It should be noted that certain tasks, particularly mechanical ones, are becoming obsolete.

### *Q: What should FI's do in next 12-18 months as they look to integrate AI?*

Business leaders should take a problem-first approach, focusing on challenges rather than technique. Stop fixating on implementing specific types of AI (e.g., deep learning, GenAI) – instead, think how the right type of AI can be combined with automation and human input to create a systematic approach to problem solving.



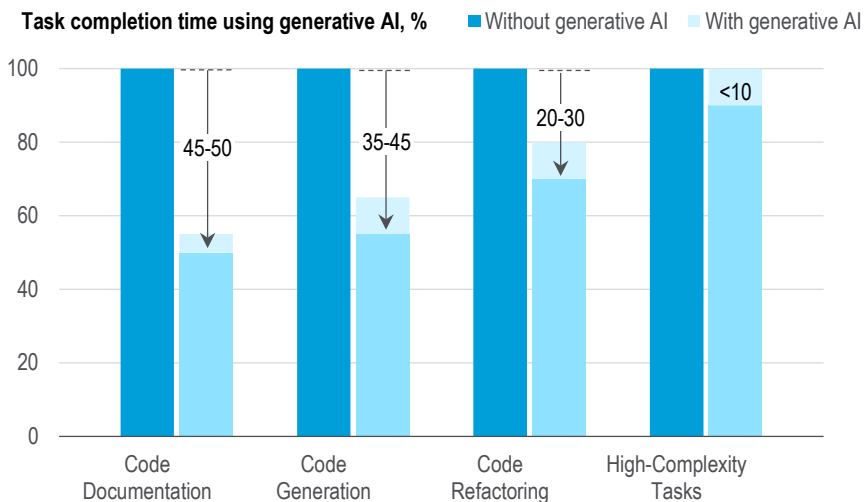
Shelby Austin is the founder and CEO of Arteria AI Inc, an AI company building modern documentation infrastructure for institutional finance. Prior to Arteria AI, Shelby was the Managing Partner of Omnia AI and sat on the executive board of Deloitte Canada. Before joining Deloitte, she founded a company that was named one of Canada's Hot 50 companies by Profit Magazine. Shelby was named to the Maclean's 2024 list of the 100 Most Powerful Canadians, one of 10 in the AI category. She was also named among Canada's Most Powerful Women: Top 100 from WXN.

AI can enhance programmer productivity by optimizing code creation and implementation, automating tasks, and improving code quality, yielding substantial cost savings.

## [1] Coding & Software

Technology and coding roles in finance will be at the forefront of AI-led productivity gains (Figure 23). This is true for all industries but particularly relevant for finance given the number of developers and coders who work for large banks and the information intensive nature of the sector.

**Figure 23. Generative AI Can Increase Developer Speed, But Less So for Complex Tasks**



Note: Exhibit from "Unleashing developer productivity with generative AI", June 2023, McKinsey & Company, [www.mckinsey.com](http://www.mckinsey.com). Copyright (c) 2024 McKinsey & Company. All rights reserved. Reprinted by permission. Source: McKinsey & Company

A study suggests that programmers with professional coding experience of average 6 years and about 9 hours of daily coding work see a significant productivity boost of 56% when enabled with AI tools such as GitHub Copilot.<sup>26</sup>

AI tools can assist in various facets of the software development process, including automating redundant tasks, creating and optimizing code, detecting errors, and improving the user experience. Automated reviews can help identify issues related to coding standards, best practices, and potential vulnerabilities.

For example, GitHub Copilot focuses on generating code based on context and existing code examples. Having been trained on billions of lines of code, it can act as a pair programmer for developers, automatically generating code, providing, and implementing suggestions.

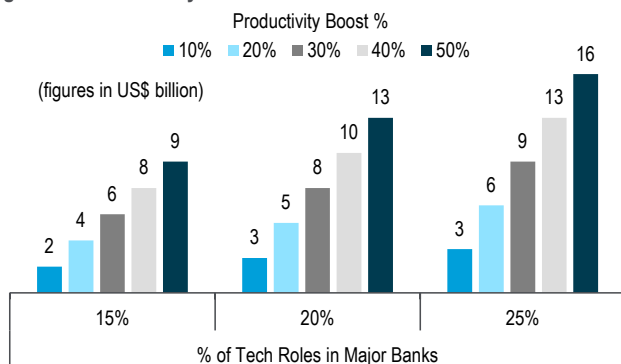
AI-assisted coding and software jobs can save the US banking sector \$2-16 billion annually (1-6% of US bank profits).

A developer productivity boost of 10-50%, weighted for a 15-25% share of staff employed in technology roles, translates to potential cost savings in the range of \$2-16 billion annually, which is equivalent to 1-6% of annual US banking sector profits. This is based on current software programmer salaries in US and India.<sup>27</sup>

<sup>26</sup> Sida Peng et al, "The Impact of AI on Developer Productivity: Evidence from GitHub Copilot", 13 February 2023.

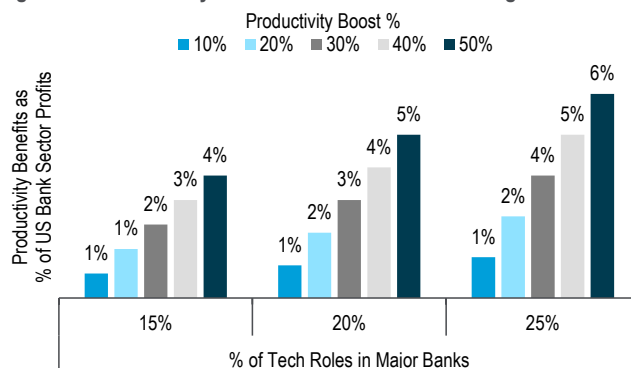
<sup>27</sup> Michael Page, Page Insights Salary Guide 2024. Our analysis assumes, as a simplification, that the US bank sector has half its technology employees in the US and half in India.

Figure 24. Productivity Boost at Tech Roles in Banks



Source: McKinsey & Company, Federal Reserve Bank of St. Louis, U.S. Bureau of Labor Statistics, Page Insights Salary Guide, Citi Global Insights

Figure 25. Productivity Boost as Percent of US Banking Sector Profits



Source: McKinsey & Company, Federal Reserve Bank of St. Louis, U.S. Bureau of Labor Statistics, Page Insights Salary Guide, FDIC, Citi Global Insights

AI paired with programming and automation introduces multiple benefits throughout the software development life cycle. Automation could reduce the amount of time developers spend on mundane and repetitive tasks, which allows them to focus on more complex and creative aspects of their work.

Furthermore, as AI automation is scalable, it makes it suitable for projects of varying sizes and complexities. The adaptability of AI tools to different programming languages and frameworks enhances their utility across the board, albeit GenAI is not as suited to older coding languages such as COBOL (see expert interview **below**).

AI boosts productivity but struggles with thinking beyond the frontier, leading to drop in output quality.

While the use of AI in software coding and programming might bring productivity gains, there are also potential drawbacks. For example, AI tools might struggle with creative problem-solving or thinking outside predefined patterns. Over-reliance on AI tools could also lead to a decline in developers' critical thinking and problem-solving abilities.

Lastly, the implementation of AI automation tools requires initial investment in terms of time, resources, and training. Maintaining and updating these tools to keep them up to date with evolving coding practices can also be resource intensive.



Sandeep Alur is a Director at Microsoft Technology Center, an Experience Center dedicated to empowering customers on their digital transformation journey.

## Measuring Developer Productivity – Expert Interview

### *Q: What will be the scope of impact of AI on developer productivity?*

Andrej Karpathy, former director of AI at Tesla and founding member of OpenAI, tweeted on 31 December 2022, “Copilot has dramatically accelerated my coding. It’s hard to imagine going back to ‘manual coding. Still learning to use it, but it already writes ~80% of my code, ~80% accuracy. I don’t even really code, I prompt & edit.”

At a broad level, when our (Microsoft) customers have measured the value of GitHub Copilot, which writes code, they have witnessed productivity improvements in the range of 35% to 55%, compared to somebody writing code without the help of any AI tool.

However, there can be different levels of productivity gains and it is likely a function of the programming languages being used by the coder. It is also possible that one might experience more accurate responses for a particular programming language relative to another as the AI model is trained on massive composite data that already exists.

Programming languages such as Java, C-Sharp, HTML, and Python are some of the most popular languages, and much of the content has been built using these open-source lines of code which are available in abundance. As a result, AI models trained on these codes become proficient in generating code in these languages.

On the other hand, if one wants AI models to write a code in ABAP program (ABAP is a programming language developed by SAP for the development of business applications in the SAP environment) or COBOL program (Common Business-Oriented Language “COBOL” is a level programming language for business applications and is still in use in many financial and business applications today), it will not perform very well as you might not find an extensive open-source ecosystem – it is a legacy language and a legacy code.

If a COBOL programmer wants to write code using GitHub Copilot, it is possible to do so and the model will generate an output to help the programmer, as it has been trained with some minimal data. However, the accuracy could be relatively lower (roughly 10-20% lower) compared to C-Sharp, Java, or Python, which are likely to generate up to 80% accurate code.

In my view, the way forward is to have domain-specific models. For example, if a financial institution has 10 million COBOL lines of code, I suggest fine tuning a LLM to create a bespoke solution for an industry that has legacy tech and legacy lines of codes.



## [2] Transaction Monitoring and Compliance

Phishing rose 341% in the period October 2023 to March 2024.

Financial crimes have risen with the advent of newer technologies. Advancements in AI have ushered in a new era for frauds and scams – from chatbots that mimic human interactions to voice synthesis and deepfakes. Global phishing rose 341% in the period October 2023 to March 2024, and was up 856% in the 12 months from April 2023 to March 2024.<sup>28</sup> There were on average 31,000 phishing attacks daily in 2023, and it rose 1,265% from October 2022 to September 2023.<sup>29</sup>

A continued rise in financial crime rates imposes substantial economic costs and jeopardizes regulatory compliance. In 1Q 2024, the global average weekly cyber-attacks per organization increased 28% from 4Q 2023 and 5% from 1Q 2023 to 1,308 – the highest recorded in the past 13 quarters. Average weekly cyber-attacks in finance and banking reached 1,172 in 1Q 2024 (-2% yoy).<sup>30</sup>

We believe AI-powered scams, frauds, and threats are likely to rise – in volume, sophistication, and potential losses – as cybercriminals discover novel ways to access customers' personal and financial data using AI. Deepfake attacks were up 31x from 2022 to 2023.<sup>31</sup> Equally, this 'bad AI' can be combatted by 'good AI'.

We believe AI-powered scams, frauds, and threats are likely to rise – in volume, sophistication, and potential losses.

Harnessing AI in transaction monitoring offers the potential to significantly reduce financial crime rates. By leveraging machine learning algorithms, AI systems can analyze vast volumes of transactional data in real-time, enabling the detection of suspicious activities with greater accuracy and efficiency.

### How Can AI Improve Transaction Monitoring

AI could offer significant enhancements to transaction monitoring compared to traditional methods. Firstly, AI algorithms can process vast amounts of data swiftly, enabling real-time monitoring of transactions, which is crucial in detecting fraudulent activities promptly.

“

Criminals are incredibly agile and we're using tech that's 20 years old or more to combat them. In anti-money laundering, it's important for us to keep in mind that there are humans on the other side of the equation that are intelligent and adopting new technologies.

One of the biggest reasons why we have false positives too, is we had a rules-based system. We write a rule to catch a particular pattern of activity, the bad guys notice that they're being interdicted, they tweak some aspect of their approach, so that rule no longer triggers.

– MATTHEW (MATT) VAN BUSKIRK, CO-FOUNDER AND CEO, HUMMINGBIRD REGTECH

”

Through machine learning techniques, AI can continuously adapt to new patterns of fraudulent behavior, becoming more effective over time. Furthermore, AI can analyze diverse data sources beyond structured transaction data, including social media, emails, and other sources, providing a more comprehensive understanding

<sup>28</sup> SlashNext, The State of Phishing 2024 Mid-Year Assessment

<sup>29</sup> SlashNext, The State of Phishing 2023

<sup>30</sup> Check Point Research, "A Shifting Attack Landscapes and Sectors in Q1 2024 with a 28% increase in cyber-attacks globally" 10 April 2024.

<sup>31</sup> Onfido, Identity Fraud Report 2024



AI transaction monitoring will identify crime that slips through the net of rules-based systems. It also frees up human time to focus on more high-value tasks.

of customer behavior and risk factors. This holistic approach enhances the accuracy of identifying suspicious transactions while reducing false positives.

Moreover, AI-powered transaction monitoring systems can incorporate advanced anomaly detection algorithms, identifying irregularities that may go unnoticed by traditional rule-based systems. This proactive approach improves fraud detection rates and reduces financial losses.

Additionally, AI can automate routine tasks, freeing up human analysts to focus on complex cases, thereby improving efficiency and reducing operational costs. Overall, AI offers a transformative solution to transaction monitoring, enhancing accuracy, efficiency, and agility in detecting and preventing fraudulent activities.

■ **Anomaly Detection:** AI-powered systems can identify deviations from expected transaction patterns flagging potentially suspicious activities for further investigation. By analyzing factors such as: **[1]** transaction amounts; **[2]** frequency; and **[3]** geographical locations, these systems can detect unusual behavior indicative of money laundering or fraud.

For example, Mastercard has launched an upgraded GenAI tool (Decision Intelligence Pro), which will be available for banks and financial institutions later this year, enabling them to save as much as 20% of the cost associated with screening illegitimate transactions. The company claims the tool can scan up to 1 trillion data points and provide an overall safety score in less than 50 milliseconds and could reduce false positives by up to 85%.<sup>32</sup>

AI-generated insights can also help prevent real-time payment scams such as Authorized Push Payment (APP) fraud. In the UK, APP fraud accounts for 40% of UK bank fraud losses<sup>33</sup> and estimates predict it could cost \$5.25 billion in the US, the UK, and India alone by 2026.<sup>34</sup>

■ **Behavioral Analysis:** AI algorithms can analyze historical transaction data to identify patterns and trends associated with illicit activities. By profiling customer behavior and identifying unusual deviations, these systems can pinpoint potential risk and proactively prevent fraudulent transactions. AI tools facilitate anomaly detection by comparing recent account behaviors with historical data, thus identifying unusual activities that could indicate criminal intentions.

■ **Network Analysis:** AI techniques such as social network analysis can help uncover hidden connections between individuals or entities involved in a financial transaction that could be fraudulent. By mapping out complex networks, AI systems can reveal underlying structures and identify key actors involved.

For example, the UK government's AI-powered fraud detection tool, the Single Network Analytics Platform (SNAP), has been upgraded with new sanctions and debarment records to help it detect suspicious networks, activity, and users that warrant further investigation for organized crime and sanctions evasion.

<sup>32</sup> CNBC, Mastercard Jumps into Generative AI Race with Model It Says Can Boost Fraud Detection by up to 300%, 01 February 2024

<sup>33</sup> UK Finance, Annual Fraud Report, 10 May 2023

<sup>34</sup> ACI Worldwide, Growth in APP Scams Expected to Double by 2026, 15 November 2022

## Scaling Compliance with AI – Expert Interview



Matthew Van Buskirk (Matt) is the co-founder and CEO of Hummingbird RegTech. Prior to launching Hummingbird, Matt was the director of compliance at Circle.com where he oversaw its regulatory function from pre-launch to its international expansion. Matt has also served as a bank regulator with the U.S. Treasury Department's Office of Thrift Supervision.

### ***Q: How have financial services firms been using AI, to date, in compliance? Does compliance tech give banks/FIs a competitive edge?***

The concept of deploying AI in the banking space has been limited to taking existing procedures and finding ways to shave off time and improve the task efficiency. It's only recently that we're starting to see more investment in a fundamental rethink in how to deploy AI to improve compliance.

One of the key areas of deployment is around reducing false positives at the tier one review for alerts of various kinds. I believe compliance is data science at its core but historically it has been more about people and paperwork. Compliance in the anti-money laundering area should be about identifying who the bad guys are and keeping them out of our institutions, not paperwork.

Modern RegTech tools and machine learning starts with models that understand what your good or a normal customer behavior looks like rather than looking for specific bad behaviors. Focusing investigative effort on anomalies from the norm allows for much more efficient investigations. This has the beneficial side effect of a much smoother and streamlined experience for good customers.

### ***Q: Why do firms have many false positives? What have firms been doing wrong? Is it the tech or policies and procedures?***

A lot of the AML program structure in place today was developed in the 1970s. We did not have the computational capabilities we have today, so we had to rely on manual effort executing against a rigid policy/procedure framework that we hoped would be effective in catching bad actors. The criminal world has evolved massively since then, but our fundamental methodology hasn't changed as quickly.

The most fundamental challenge is the data that's involved. Traditional compliance functions in this financial crime space focus a lot on customer identity data sets, business information, and transactional data. That is a narrow slice of the information available today and it is often fragmented across multiple data silos.

One of the reasons we have false positives is in a rules-based system, we write a rule to catch a particular pattern of activity, the bad actors notice that they're being interdicted, they tweak their approach, and that rule is no longer effective, even though it remains in place and triggers on innocent customers.

### ***Q: How can AI help scale compliance and improve compliance efficiency? How will AI be used in compliance 2-3 years from now?***

There is a need to develop a modular ecosystem of capabilities where you're going to be building models that are going to be your new transaction monitoring systems, KYC monitoring, the customer onboarding algorithms, etc. And then large language models end up being the glue that can access a bunch of different systems.

Big global banks have 5,000-10,000 analysts engaged in compliance tasks, 90% of which are doing the level one triage type of review and dealing with 99% false positives. AI will likely make level one type of review jobs redundant by filtering alerts down to those that actually require human review.

We will likely see a lot of tools where you can basically run your own LLMs locally, so you do not need to worry about data leakage risks. I can imagine a data analyst will likely click a run button on an AI bot that they've trained to go in and find all that data, package it up, and present it a summary to them.

I am seeing experiments today that are doing that effectively and reliably and generating proper audit drills. If it could be done today in an experimental environment, I think it's going to be deployed widely in the next 2-3 years.

The introduction of chatbots brought about a paradigm shift by offering instant responses, 24x7 availability, and the ability to handle multiple queries simultaneously.

## [3] Customer Services & Chatbot 2.0

Another notable application of AI in the financial services domain is the integration of chatbots into customer service platforms in the forms of personal assistants or chatbots. The use case is driven by the computer code's ability to understand, interpret, and respond to natural human language. According to a Markets and Markets report, the global chatbot market achieved a valuation of \$4.7 billion in 2022 and is projected to grow to \$15.5 billion by 2028.<sup>35</sup>

Historically, customer service in the financial sector relied heavily on manual processes including phone calls, emails, and in-person interactions. While these methods were effective to some extent, they often suffered from long wait times, limited availability, and inconsistencies in information delivery.

In 1966, Joseph Weizenbaum, an MIT professor, created the first chatbot.<sup>36</sup> Chatbots are computer programs designed to simulate human conversation, and they have rapidly gained traction in the financial services industry due to their ability to streamline customer interactions, provide personalized assistance, and offer round-the-clock support.

A pre-ChatGPT and pre-LLM proliferation Juniper Research study suggests that chatbots could save up to 862 million hours for banks' human staff and bring in operational cost savings of \$7.3 billion by 2023.<sup>37</sup>

The introduction of chatbots brought about a paradigm shift by offering instant responses, 24x7 availability, and the ability to handle multiple queries simultaneously. The adoption of chatbots in financial services has yielded several notable benefits over traditional customer service functions.

### Benefits of AI-Powered Chatbots:

- **24x7 Availability:** Unlike human agents, chatbots operate round-the-clock, providing clients with uninterrupted support, and positively impacting customer satisfaction. Chatbots can handle a wide range of queries autonomously, alleviating the burden on human agents and enabling them to focus on more complex issues requiring human intervention.
- **Prompt Responses:** Chatbots have significantly reduced response times, allowing customers to obtain immediate assistance for routine inquiries such as balance inquiries, transaction history, or account updates. AI-based finance chatbots could gather all customer information at the start of the interaction enabling immediate user support and responses without waiting periods.
- **Hyper-Personalized:** AI-powered chatbots can leverage customer data around past transactions, and customer preferences/behaviors to offer personalized solutions and tailored recommendations. Chatbots can also help ensure consistent user experience across channels, ensuring customers receive accurate and relevant information regardless of the platform they engage on.

<sup>35</sup> Markets and Markets, Chatbot Market Size, Share & Trends Report, May 2023

<sup>36</sup> The Guardian, Weizenbaum's Nightmares: How the Inventor of the First Chatbot Turned Against AI, 25 July 2023

<sup>37</sup> Juniper Research, Bank Cost Savings via Chatbots to Reach \$7.3 Billion by 2023, 20 February 2019

Previous version of chatbots couldn't comprehend complex queries and were based on a predefined workflow.

AI-powered chatbots 2.0 are better at understanding context but to enable safe use robust policies, training, and human oversight are required.

■ **Mitigating Fraud:** Addressing financial fraud is a paramount concern for banks. Integrating AI tools, both customer-facing and risk mitigation, enables efficient handling of a large volume of requests. This reduces the workload on representatives while promptly identifying security concerns. In the following pages, we will delve into AI use cases in transaction monitoring and compliance.

Despite their potential, chatbots have limitations that have hindered their widespread adoption. A common issue is their inability to comprehend complex or ambiguous queries, leading to frustration among users and a decline in customer satisfaction. Chatbots also often struggle to accurately interpret colloquial language, regional dialects, or slang, resulting in miscommunication and misunderstandings. The lack of empathy and emotional intelligence exhibited by chatbots has also posed challenges in handling sensitive customer issues or complaints effectively.

Since AI models are known to hallucinate and create information that does not exist, organizations run the risk of AI chatbots going fully autonomous and negatively affecting the business financially or its reputation.

For example, a North American airline was held liable after its chatbot promised a passenger a discount that was not available.<sup>38</sup> It is worrying to imagine the implications of this in finance – what if an AI chatbot offered incorrect interest rates on mortgages, unsuitable health/insurance policies, or extravagant credit card offers? These outcomes are not far-fetched.

In February 2024, Gartner warned that by 2027, a company's GenAI chatbot could directly lead to the death of a customer due to misinformation.<sup>39</sup> Unfortunately, just a year prior, in March 2023, a Belgian man committed suicide after an AI chatbot encouraged him to kill himself, according to published reports.<sup>40</sup>

Hence, before hastily deploying AI, particularly in customer-facing capabilities, companies should invest in appropriate AI usage policies, frameworks, and architecture. Additionally, training personnel to understand how AI is altering their work processes is essential.

To address these concerns and enhance effectiveness of chatbots in financial services, human oversight remains indispensable. In fact, an increased human presence may be necessary to monitor customer interactions with AI chatbots and ensure they are within legal and business guardrails.

<sup>38</sup> BBC, Airline Held Liable for its Chatbot Giving Passenger Bad Advice – What This Means for Travelers, 23 February 2024.

<sup>39</sup> CX Today, Gartner Warns that GenAI Will Directly Lead to the Death of a Customer by 2027, 22 February 2024

<sup>40</sup> Vice, Man Dies by Suicide After Talking with AI Chatbot, 31 March 2023

## Chatbots Gallop From Horses to Cars – Expert Interview

### *Q: How has the AI use case as communication chatbots changed over time?*

When ChatGPT was launched in late 2022, it changed two things: [1] created mass awareness about the capabilities and power of conversational AI systems; and [2] raised the expectations of performance from corporate or banking chatbots.

I think when it comes to chatbots, we have experienced transformation somewhat similar to going from horses to automobiles. The world has gone from prescribed AI systems for chatbots to domain-specific generative LLMs. In the prescribed chatbot paradigm, workflows had to be coded and responses written out. Where now with GenAI, given the proper guardrails, the AI is able to decipher meaning, find the right information, and offer a very broad set of answers more effectively and efficiently.

Our banking clients are asking for an AI conversational tool that's similar to ChatGPT but does not hallucinate and understands the industry vocabulary. Domain-specific LLMs are so powerful that they understand banking and all the jargon used. That's why I believe the transformation is increasingly going to drive adoption of GenAI systems for chatbots.

### *Q: How can banking and finance organizations make use of responsible AI tools for productivity gains?*

I think that there is a very strong level of interest in GenAI, and the industry understands we are at an inflection point that will change the way banking is done forever. Today, the industry is looking at GenAI assistants as a means of labor augmentation and efficiency gains, not as labor displacement technology. But, in the next 3-5 years, the industry could achieve artificial financial intelligence, where AI tools could replace relationship managers and financial advisors.

The AI systems, despite a lot of complexity around them, can be very effective in bringing productivity gains. These AI systems need to be guardrailed for wider adoption in the banking and financial services industry.

In financial services, there is less room for error. For example, every time our AI tools generate an answer, we ensure that there is a source document of where this answer came from. If there is no source, our LLM will not hallucinate and will not generate an output that the model thinks is true. So, we not only provide a source document but also provide the excerpt of where the answer came from within the document, so that the bankers can cross-check everything.

Moreover, the transparency of models is super important and that is one of the principles we used in constructing our own banking LLM. We know that banking regulators want to know which training sources were used, which algorithm were used, and we provide full transparency.

They also want to know how we use proprietary documents, or not use proprietary documents. Even before GenAI, every large bank that we work with comes to us wanting to know how our models are built, what information they were trained on, and how they perform.

We have seen customer satisfaction improvements for financial institutions that have had their AI assistant live in market for a short number of months. We published a case study with one of those customers (a Silicon Valley-based credit union). In the first nine months since they launched their AI assistant, they've seen a 30% increase in the profitability of their active client base who utilize the AI tool.



Zor Gorelov is the co-founder and CEO at Kasisto, creators of KAI Banking, a conversational AI platform for the finance industry. Zor has more than 20 years of experience in the enterprise software industry. Prior to Kasisto, he was the CEO & co-founder of SpeechCycle. Before that, Zor founded and ran BuzzCompany.com, a provider of enterprise collaboration and messaging software.

There is certainly a correlation between delivering a delightful experience and higher client engagement in terms of clients inclined to buy more, stay longer, tell friends, and that's exactly what we are seeing with the institutions that are incorporating generative conversational AI into their customer service models.

***Q: What are the challenges that you or your clients have encountered while introducing AI tools into their business?***

We are a B2B2C business. We sell to banks. Banks make our technology available to the end customers. We believe that our success is ultimately driven by the users of our system. It's not just the number of banks using our technology, it's how many customers are really engaged with our systems and what value do they get.

I think that when you deploy the systems, one of the biggest challenges for new users is discoverability and anticipating what users need and creating the surprise and delight experience as that drives engagement, brings users back.

Some of the challenges are really trying to understand, anticipate, and predict consumer behavior and try to greet users with things that they are most likely to do while educating them on the capabilities of the AI.

For example, we know that if somebody is asking about a credit card payment date, then they are likely to ask questions about rewards. So, we're anticipating and then giving that experience to make sure that users are surprised and delighted and then they come back.

I think conversational interface is about human interactions and about enabling things that people cannot do on digital. The things that are rooted in the way people think, the way people interact and that creates surprise and delight.

Another challenge with generic chatbots is that they do not know much about you, and I feel they have failed as a category. People do not go to banks to ask general questions about their finances. I think building that knowledge, building financial literacy, building connectivity into banking systems so you can engage in conversations about clients' finances is critical.

So, discoverability, providing meaningful interaction experience about your finances, and then focusing on use cases that are rooted in the way people think about banking and finance are some of the main challenges.



Use of AI can help enhance data analysis efficiency, and improve risk evaluation, underwriting, and claims processing.

## [4] Credit Risk & Underwriting

As risks are diversifying and getting more complex due to cyber threats, climate change, and other emerging risk scenarios, AI tools can help financial institutions better evaluate risk and underwrite policies with greater speed and precision.

Risk and underwriting systems backed by AI can comb through vast amounts of data from public sources, third-party transactions, historical records, and other channels to minimize manual processes and reduce errors.

In insurance, AI tools can be a boon in handling claims, reducing the administrative burden of validating submissions and increasing accuracy. Proliferation of smart connected devices can further help insurers automatically gather a wide variety of claim related data that can be analyzed by AI tools.

However, AI could also have some negative impacts on some finance clients, especially edge cases in insurance. Access to insurance could be limited, or premiums could rise significantly, for higher-risk individuals or properties where big data insights or AI analysis identify greater risks.<sup>41</sup>

Key functions in credit risk and underwriting that can leverage AI include:

- **Advanced Data Analytics:** Traditional credit scoring primarily relies on structured financial data such as credit history, outstanding debt, payment behavior etc. While this provides valuable insights, it is limited in scope and does not portray the full picture of a borrower's creditworthiness.

Use of AI and NLP can help financial institutions gather and analyze vast amounts of unstructured data such as rent payments, gig economy income, social media feeds, etc., which could help improve overall approval rates and better evaluate risks associated with an individual/business.

- **Pattern Recognition:** AI tools can enable banks and insurers to identify new patterns and relationships amongst input data that could result in a stronger model. This can help improve accuracy of the model and make more precise credit risk assessments, reducing the chances of default/fraud.

A key success driver in insurance is segmenting risk as finely as possible to achieve pricing that is as accurately calibrated to the risk. Deep learning models can help segment customers based on risk profiles, allowing for more tailored insurance products and pricing.

- **Predictive Modelling:** Financial institutions can use advanced algorithms to generate new data points or simulate scenarios based on historical data to predict more precise credit risk assessments and credit scoring. For example, banks can use AI models for more accurate default probability and loss severity.

Insurance companies can leverage AI tools to make predictions about weather conditions by analyzing weather data and historical loss information to enhance underwriting accuracy and streamline claims settlement.

---

<sup>41</sup> EIOPA, Big Data Analytics in Motor and Health Insurance: A Thematic Review, April 2019

- **Continuous Account Monitoring:** AI/ML models can help banks continuously monitor borrower behavior/macroeconomic conditions and help in IFRS (loan-loss provisioning, and fraud and capital requirements modelling). For insurers, continuous monitoring of claims can help improve the fit between pricing and risks. Real-time monitoring can also help detect suspicious activity and stop potential fraud or suggest risk prevention measures based on data.
- **Personalized Models:** AI/ML-driven credit risk assessment models can be customized to specific industries, customer segments or regions, capturing the needs/behaviors of different groups. This is an improvement over traditional risk modeling, where customer segmentation is based on hard lines and broad categories. This allows for more dynamic credit decisions and pricing.
- **Lead Management:** AI systems can help indicate positive leads to marketing teams by pulling data from a multitude of sources that might not have been considered before. Furthermore, AI could also deliver additional value by designing personalized products and services for cross-sell/upsell based on company data (e.g., past policies, customer feedback, claims data, demographic stats, etc.).
- **Automate Routine Processes:** Automation of processes using AI/ML tools in conjunction with robotic process automation could help reduce documentation burdens for customers and back-office employees. This could help reduce the amount of human effort, speed-up time to decision, and be more cost effective. For example, the use of visual AI can help insurers analyze images of damage to cars and properties, providing instant assessments and cost estimates. This can help with initial claims routing, triage, and reduce overall cycle times.
- While AI gives insurers a more accurate understanding of risk, which could lead to a reduction in premiums; lack of transparency in AI models remains a concern.

### Using AI to Assess Risk and Adjust Premiums More Accurately

Insurance companies have traditionally used data and statistical models to calculate risk using techniques such as linear regression. Use of AI could enable insurers to increase accuracy of these predictions by analyzing larger datasets. AXA, a leading insurance company, is leveraging AI to better understand risk in property, life, and health insurance; model them; and propose accurate premiums.

#### How is AXA Leveraging AI?

- Risk predictions using AI/ML algorithms that forecast the likelihood of future claims using a variety of structured and unstructured data input.
- Use of image analysis, NLP, and text recognition tools to analyze documents or photos and identify key information in contracts.
- Use of predictive models for variety of use cases – analyzing weather data to predict when requests for assistance will rise, forecasting call volumes.
- Automate elements of customer service with 24x7 support via multiple channels.

**Resultant Outcome:** By harnessing the power of AI, AXA has gained a more accurate understanding of risk, allowing it to lower premiums on some products by as much as 5-10%, giving them a competitive edge in the industry.<sup>42</sup>

<sup>42</sup> Microsoft, Best of Business AI – AXA Case Study

Despite its benefits, the usage of AI/ML models in credit risk has faced its share of skepticism, especially regulatory, due to the lack of transparency. The potential 'black box' effect from AI/ML systems constitutes both a practical and ethical concern for financial institution and we discuss this more in detail in the next chapter as potential pitfalls of AI technology.

## [5] Investment Research

Investment research is a crucial function for both sell-side broker dealers and buy-side institutional investors. The section below highlights how the industry can leverage AI throughout the investment research process.

- **Data Extraction, Consolidation, and Analysis:** People have referred to the phenomenon of having access to more data than they could process or utilize as data deluge<sup>43</sup>, and this is a widespread phenomenon across the financial industry, which has become increasingly data driven.<sup>44</sup>

AI can be one of the remedies for the data deluge problem by enabling the industry to process and utilize more data in a faster and less labor-intensive way. Data makes up a large portion of the investment research process particularly for systematic investors. Incorporating quantitative analysis and more data have also become growingly important for sell-side and fundamental research analysts. AI tools and solutions today can go through various sources including websites, datasets from vendors, and the organization's proprietary datasets to extract, consolidate, and analyze key data points. Many unstructured data such as PDF documents and texts on company websites, which traditionally relied on humans to process can now also be processed effectively by AI.

AI tools could help analysts save around 25% of time spent on data heavy workload allowing them to focus on more client engagements.

GenAI is now capable of translating prompts in natural language into code to retrieve and analyze data so instead of manually conducting data analysis, analysts may simply give prompts or inputs to GenAI and receive responses calculated and generated by the machine. For example, an analyst may ask how many times in the past month has a stock outperformed another, or historically how has a stock reacted to interest rate hikes. **Figure 26** shows how Citi estimates an AI enabled research team could save up to 25% of the time spent on company and industry data collection, consolidation, and analysis.

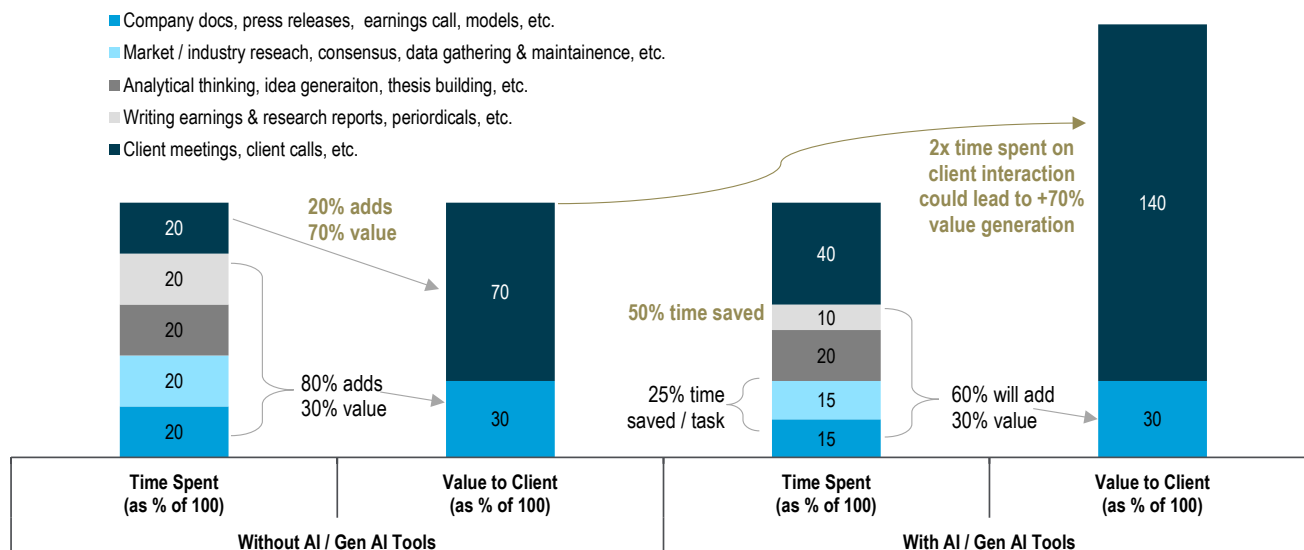
- **Report Writing and Communication:** As shown in **Figure 26**, writing earnings, research reports, and periodicals can take up one fifth of an analyst's time. For buy-side research analysts, writing emails and communicating trade ideas with peers and portfolio managers can also take up a significant portion of their time.

GenAI tools can assist analysts save time on writing reports or email communications that summarizes market colors, make recommendations, and present trade ideas. Large language models have made significant progress in the past few years on processing unstructured data such as text and audio to provide summarization and real-time transcription, which can help with content production around live events like earnings calls and analyst days.

GenAI trained on past research reports of an analyst, may be able to imitate their writing style to generate reports in their tone or style. Instead of starting from scratch, analysts in the future might input specific news, financial reports, transcription from earnings calls, and key data points into GenAI tools to generate drafts, from which they can then finetune, add their deeper-level thinking and observations, and read through to ensure accuracy of the language and the data cited.

<sup>43</sup> Hewlett Packard Enterprise, "What is Data Deluge?"

<sup>44</sup> Markets Media Group, The Data Deluge: What Investors Need to Know to Stay Ahead of the Curve, 22 June 2023

**Figure 26. AI/GenAI Tools Could Increase Research Productivity & Add Value to Clients**

Notes:

[1] We bifurcate investment research tasks into 5 equal verticals [i] Coverage company related knowledge gathering and models; [ii] Industry research, data gathering and maintenance; [iii] Analyst's own analytical thinking and idea generation; [iv] Writing research reports; and [v] Client meetings and interactions.

[2] We assume 80% of time spent on Task 1-4 generates 30% value for clients; whereas direct client interactions through meetings and calls adds 70% value for clients.

[3] We estimate that several AI/GenAI tools could help reduce time spent on gathering data, search & summarization of information for tasks [i] and [ii] above by about 50% each but half of the time saved will be spent on checking data accuracy, so overall time saving is 25% for each task. No substitute for analytical thinking presently and we do not account time savings for this task. Writing research report can be more efficient by leveraging GenAI tools, replicating information and analyst writing style – we assume 50% time saving.

[4] In aggregate, 20% of time saved assumed to be spent to serve clients better or do more client engagements.

Source: Citi Global Insights

- **Idea Generation:** AI can help with idea generation in numerous ways. It can supplement an analyst's thinking with additional qualitative and quantitative insights such as analysis of geospatial and satellite images, or sentiment analysis using alternative datasets like social media posts and website scraping data. It can also provide predictions and forecasts based on a wide range of data sets and hundreds of factors which may pick up non-intuitive patterns and signals that traditional statistical methods cannot.

AI's efficiency in processing large datasets can also help augment analysts' existing idea generation and analytical thinking capabilities by enabling them to monitor more securities. For example, AI can automatically screen, sort, and rank securities across a large investment universe based on complex constraints and considerations such as the analyst's views and selection criteria, risk constraints, the market condition, and fundamentals of the securities.

AI will assist but NOT replace analytical thinking and idea generation.

We believe, analytical thinking and idea generation is an integral part of investment research and newer generative AI tools will assist analysts in doing these tasks and augment their capabilities but won't substitute them.

**Financial Statement Analysis with Large Language Models** <sup>45</sup>

A research study published in May 2024 by the University of Chicago's Booth School of Business reveals interesting results on how LLMs perform financial statement analysis tasks.

The study concludes that large language models (LLMs), specifically GPT-4, can effectively perform financial statement analysis, often outperforming human financial analysts and specialized machine learning (ML) models in predicting future earnings changes. GPT-4 demonstrates a significant ability to analyze purely numerical data from financial statements and generate useful economic insights, suggesting that LLMs could play a central role in future decision-making processes in finance.

The research indicates that GPT-4 achieves higher accuracy of 60.35% in predicting earnings changes when using chain-of-thought (CoT) prompts, mimicking human reasoning, compared to simple prompts (52.33% accuracy). Moreover, the trading strategies based on GPT-4's predictions yield higher Sharpe ratios and alphas compared to those based on human analysts' forecasts and other models.

The study acknowledges several limitations, including the potential for retrieval errors in the model's responses, especially when narrative context is integrated. Additionally, while GPT-4's performance is impressive, it does rely on numerical reasoning derived from narrative context rather than deep numerical analysis. The model's predictions are tested against standardized and anonymized financial statements to rule out the influence of memory.

---

<sup>45</sup> Kim, Alex G. and Muhn, Maximilian and Nikolaev, Valeri V., Financial Statement Analysis with Large Language Models (May 20, 2024). Chicago Booth Research Paper Forthcoming, Fama-Miller Working Paper, Available at SSRN: <https://ssrn.com/abstract=4835311> or <http://dx.doi.org/10.2139/ssrn.4835311>

## Next-Gen Financial Research with AI – Expert Interview

**Q: How can AI help in equity and fixed income research? What are some of the key use cases?**

**Ken Sena:** Use of AI in research can primarily help enhance breadth, speed and insights. Large language models can provide analysts new tools to enhance coverage, process documents/transcripts with ease, and analyze sentiment indicators more effectively. AI-powered functionality like search and summarization, graphical representation etc. can help analysts generate insights. Analysts can also create personalized experiences targeted at a broader audience by leveraging AI.

These enhancements are likely to transform sell-side research, empowering analysts do more with less – i.e., small teams are likely to cover more stocks or sectors, and potentially also encouraging creative research.

**Eric Chang:** AI can help automate manual tasks, allowing analysts to process more information and capture a broader data set (which may not have been feasible with human effort alone). AI tools can sift through large documents (e.g., prospectuses, memorandums, financial statements) to extract and summarize relevant analytics, enabling more effect risk and pricing decisions.

This expanded data collection/analysis helps improve benchmarking and comparative analysis. Over time, use of AI is likely to transform the investment process by integrating analysis of larger data sets, resulting in new strategies and rethinking investment methodologies. Successful companies will experiment with AI and integrate it with human expertise.

**Q: How can AI contribute to the digitization of fixed income markets, and what challenges must be addressed?**

**Eric Chang:** Fixed income markets have often lagged equity markets in digitization. However, the use of AI is likely to drive more information and insights, which could help the market grow and increase the number of participants. Fixed income trading will likely shift more towards digital channels, as opposed to telephonic and chat-based transmission today.

A significant barrier to digitalization is the non-fungibility of fixed income products. Unlike equities, where two shares of company A has the same rights; fixed income products often have different rights and warranties with different waterfalls, seniority of payment etc. However, use of AI can help present structured information for easier analysis, making these instruments a little bit more fungible.

The transformation will require combining human strengths with AI capabilities, leading to a more transparent and efficient market. Companies that can effectively merge these elements will thrive.

**Q: What aspects of AI appear to be overhyped in your view? What are some of the more realistic use cases?**

**Ken Sena:** The belief that AI knows more than we (humans) do is a fallacy. While AI often has access to vast amounts of information, it should not be blindly trusted, especially in research. There is also a risk in over reliance on generalized AI models, which stifles creativity(new behavior) as model outcomes are driven by past patterns. It is also worth noting that AI systems are being built to support users, rather than replace them.



Eric Chang is the Co-Founder and CEO of Clairra. He is an innovation driver and product leader across financial services having built FinTech startups from the ground up, including a big data surveillance company in partnership with Credit Suisse and Palantir.



Ken Sena is the Co-Founder and CEO of Aiera. Prior to Aiera, Ken led global Internet research for Wells Fargo Securities and Evercore ISI. He began his research career at Bear Stearns, following strategy and business development roles at Viacom and Time Warner.



**Eric Chang:** While AI is powerful and more complex than traditional software, it is ultimately only a computerized tool. Developers and implementors need to focus on using AI to solve business problems sustainably, rather than chasing hype cycles. For example, ChatGPT has created substantial buzz around its ability to generate content, but it lacks fact-checking and privacy considerations. Much like evaluating a new human hire in any organization, AI systems need to be evaluated for design, purpose, and the ability to integrate it into business processes.

**Q: *What is the scale of productivity improvements we should expect from the use of AI tools in research over the next 3 years?***

**Eric Chang:** Over the next 3 years, we could see a significant shift in the way research is conducted. As data gets more easily accessible, we could see an increase in the number of people involved in research, particularly in fixed income markets.

Analysts currently spend nearly a third to 40% of their time in gathering data, but this could reduce to just 5-10% with the use of AI. This means more time spent on analysis, trading and developing strategies for investment purposes. Availability of more detailed data could lead to new market opportunities and strategies, possibly even high-frequency trading in some markets, ultimately leading to new jobs.

**Ken Sena:** In my view, over the next three years, we are likely to see an increase in the number of analysts teams that are AI enabled. Financial markets are dynamic and anybody with the ability to interpret information uniquely will be able to create new opportunities.

We could possibly also see new models emerge in research where independent analysts create and distribute their own research. AI tools will be essential for ensuring accuracy of information and fostering competition.

## [6] Asset & Portfolio Management

The asset management industry has been using AI for decades across various areas such as quantitative research, portfolio optimization, and trading. While there are many AI use cases in asset and portfolio management, the implementation of AI is still at an early stage and requires a human in the loop.

According to a recent report by Citi's Business Advisory Services (BAS) team, [AI in Investment Management - The Pursuit of a Competitive Edge](#), which conducted interviews with buy-side institutions covering about \$30 trillion, most participants recognize the potential of AI tools in the medium-to-long-term and are actively exploring and testing AI uses cases with a human in the loop. However, the majority of them said the current suite of AI tools does not meet the high standard of transparency and accuracy needed to be used extensively in key business activities such as alpha generation and client engagement.

In the near term, the investment management industry, has been focusing more on productivity and efficiency gains in operational processes. Change management related to AI has also been a top challenge and key focus for interview participants. The BAS report goes into detail regarding how buy-side organizations can more effectively manage changes to broaden AI adoption and how to build the future sources of competitive advantages in the age of AI.

The combination of traditional AI and GenAI can be a powerful efficiency and productivity tool, freeing up managers time to focus on alpha generation.

Traditional AI techniques such as machine learning can help humans process data, detect patterns, provide predictions, and solve optimization problems using a large volume of data. Meanwhile, GenAI can process requests and prompts in plain human language to deliver analytical results, recommendations, or other types of tailored output to its users. The combination of the two can make powerful tools and solutions for portfolio managers to improve productivity, so they can spend more time on high-value activities such as client engagement and alpha generation.

There are many areas in portfolio management that the latest innovations and progress in AI can help further improve. The section below highlights some of the key use cases of AI across the different stages of portfolio management.

- **Define Investment Objectives:** It is crucial for wealth advisors and asset managers to understand their clients' investment objectives to help them achieve long-term goals and financial success. However, it can be time-consuming to understand clients' circumstances and tailor the most suitable portfolio mixes for them on an ongoing basis.

The personalized approach can help build trust and improve client experience, which may eventually lead to stickier client relationships. By leveraging AI, companies can provide personalization to clients in a more scalable manner and enable their relationship managers or advisors to engage with clients efficiently.

AI tools can process a large amount of structured and unstructured data related to clients such as their social media accounts, credit card transaction history, and investor profile questionnaires. AI can also assist with inputting these data into the customer relationship management (CRM) database to build a more centralized and holistic data hub.

This will help asset and wealth managers establish deeper understanding of their clients and keep track of changes in their clients' circumstances. Conducting cluster analysis with machine learning algorithms on the collection of client data may help organizations more accurately identify clients' investment objectives and personalize portfolio construction strategies in a more scalable way.

AI tools could assist portfolio managers to optimize portfolios, make better asset allocation and security selection.

- **Optimize Portfolios and Allocate Assets:** Quantitative managers have been using statistical methods to forecast short-term expected returns and variance-covariance matrices to perform portfolio optimization. AI approaches such as LASSO, neural networks, and support vector machines can generate more accurate forecasts compared to traditional methods, helping portfolio managers better manage risk/return in their portfolio allocation strategies.<sup>46</sup>

Instead of optimizing the weight of each instrument and the overall asset allocation based on the current situation, AI models trained on historical data can discount possible changes in market conditions and volatility on a forward-looking basis. Discovery of signals across different geographies and asset classes is another way portfolio managers can leverage AI for asset allocation.

Portfolio managers may also choose to apply AI overlays that amend existing tactical asset allocation to enhance overall returns of their portfolios.<sup>47</sup> Another way AI can be helpful is back-testing portfolios to compare possible outcomes of different asset allocation strategies.

- **Security Selection:** GenAI tools can help reduce the amount of manual labor involved during the security selection stage. For example, it can respond to prompts and inquiries from humans to perform certain retrieval and analysis such as providing a list of securities filtered by specific risk constraints or investment considerations (e.g., ESG).

Another way today's investment managers have been leveraging AI in security selection is mitigation of human cognitive biases. Machines which follow rules based on data are free from emotion and ego that may get in the way of humans making logical decisions, yet inbuilt bias is a significant AI concern.

While AI may not have the domain-specific knowledge and expertise a human portfolio manager has, machine learning algorithms can learn from a human's past transactions and the corresponding outcomes to perform behavior diagnostics. These algorithms can provide recommendations to portfolio managers by flagging the biases that may be embedded in their actions and nudge them to correct their behaviors, such as increasing the size of a certain investment or selling a security sooner despite the loss.

As mentioned in the idea generation section earlier, AI tools and solutions can also add value to security selection by supplementing a portfolio manager's decision-making process with additional insights from traditional and alternative datasets, and helping them screen, filter, and rank securities to save time and allow them to focus more on deeper level analysis.

- **Monitor and Adjust:** Most of the activities involved in the previous three stages of portfolio management are not static. Instead, they are dynamic and require active monitoring and adjusting to align with changes in the markets, the underlying holdings, and the clients' circumstances.

One of the challenges portfolio managers have faced is the limited bandwidth to monitor the changing economic and market conditions as well as fundamentals of the securities. AI could help significantly reduce the amount of manual labor involved, enabling portfolio managers to dynamically monitor the target investment universe based on their pre-defined criteria and the large amount of traditional and alternative datasets used during the security selection stage.

<sup>46</sup> CFA Institute, Artificial Intelligence in Asset Management, 2020

<sup>47</sup> Generali Investments, Machine Learning and Tactical Asset Allocation – Part I: Integrating the Machine Learning Signals, 15 February 2024

AI tools could also help in monitoring and adjust portfolio as per market changes and assist in rebalancing portfolios at regular intervals.

Certain AI tools can also help in better and proactive risk management by updating forecasts as new data becomes available. In summary, AI can empower portfolio managers to cover more securities or the same number of securities with greater depth and help them anticipate or react to changes faster to adjust the weightings and positions in their portfolios.

- **Monitoring Client Needs:** Like the difficulties faced by portfolio managers in monitoring securities, advisors and relationship managers also find it challenging to keep up with the evolving needs and wants of their clients to adjust the portfolio mix. This is where AI can help by monitoring clients' evolving circumstances using traditional and alternative data and either proactively flagging changes or responding to advisors and relationship managers' prompts and inquiries to help them provide more tailored services and portfolios to clients.
- **Re-optimize and Rebalance:** AI models are highly capable of solving optimization problems under complex constraints, making them a great assistant to humans in rebalancing and re-optimization of portfolios. They can help portfolio managers simultaneously re-optimize and rebalance a larger number of portfolios subject to the client's risk profile and the portfolio manager's views on the economic conditions and the underlying holdings in the portfolio.

The ability of AI to forecast based on historical data can also enable portfolio managers to re-optimize and rebalance weights of each instrument in a forward looking way by accounting for possible changes in market condition, volatility, and asset correlation keeping in mind turnover to minimize trading impact.

- **Analyze Attribution and Report:** Traditional AI and GenAI models can reduce the amount of manual labor involved in the process of attribution analysis and client reporting. By leveraging AI models to analyze portfolio data at a more granular level, managers can conduct more detailed and accurate attribution analysis. It will also guide future portfolio adjustments by better understanding of risk-adjusted performance and deviation from investment objectives.

Attribution analysis is also an important component of client reporting. GenAI can help generate attribution analysis and portfolio updates tailored for each client with customized language in a more scalable manner. A big part of client reporting may also involve going through large volumes of structured and unstructured data such as PDF documents, to extract key information, which has now become increasingly automated with the help of GenAI.

A growing number of GenAI tools and solutions now also aim at assisting advisors and relationship managers with obtaining information they might need during client communication such as performance of client portfolios, investment recommendations, and process-related questions.

## Supercharging Capital Markets with AI – Expert Interview



Bin Ren, Founder and CEO of SigTech, has had a varied career, from equity exotics trader at Barclays Capital to leading systematic investment in quant funds at Brevan Howard. He also founded Elwood Asset Management in 2019 alongside SigTech

### *Q: What are the key applications of GenAI in capital markets?*

In the short-term, development of GenAI is likely to focus on enhancing multi-modality (i.e., process multiple data inputs). GenAI was initially rooted in text but is evolving to incorporate other modalities like visual (images, videos) and acoustic (sounds, speech).

In the medium-term, incorporation of diverse modalities can help AI models learn on larger training datasets and develop intuitive comprehension of concepts, which could enhance reasoning capabilities. For example, an AI model trained on a video of a flying bird, can better understand how birds fly, movement of wings, concepts of air/aerodynamics etc. This is more elaborate than what can be achieved via text.

Another trend is the building of autonomous AI agents that can understand high-level inquiries, deconstruct them into smaller actionable tasks, and execute these tasks sequentially. Technology enhancements like multimodal integration can allow agents to conduct web searches, perform calculations and devise investment strategies. Many of these functionalities are not speculative, but in operation today.

In the longer-term, AI agents could augment human work. While these agents may share the same foundational knowledge and reasoning capabilities, they are tailor-made for specific functions. For example, an agent specializing in analysis of macroeconomic trends or global financial news, broker agent for OTC instruments.

### *Q: Can you provide examples of tasks an autonomous AI agent is capable of?*

Let us take a simple example of an autonomous AI agent given the task of providing insights into the sharp fluctuations of a company's share price.

When presented with such a query, the AI agent would initiate a series of automated actions, including web searches, article analysis, and earnings report review. It would then synthesize the findings into a coherent summary for user consumption. This could save significant human time spent in trawling through elaborate earnings reports, management call transcripts, press articles and industry analysis.

### *Q: What is the addressable market (TAM) for GenAI in capital markets?*

Considering the human capital-intensive nature of finance, one needs to look at the cost of human capital to estimate the TAM for AI adoption in capital markets. The sector employs over 700,000 finance analysts in G7 countries alone. Even with a conservative estimate of average pay, it represents an annual payroll of nearly \$100 billion. A modest 1% increase in productivity facilitated by GenAI, equates to an additional billion dollars' worth of productivity per year. While this calculation is only illustrative, it underscores the potential for GenAI to boost productivity.

### *Q: Where are we in the adoption cycle of GenAI in capital markets?*

Capital market firms are still in the exploratory phase with GenAI. Until just a year back, customer data privacy had been a major concern for enterprises adoption. However, the use of private cloud environments by major cloud providers like Microsoft and Amazon have helped address some concerns.

In my view, this space is likely to evolve rapidly as enterprises explore new AI tools. Several questions remain unanswered – what applications should be prioritized? Which internal data sources should be made available to AI models? How does one govern AI models? What guardrails need to be in place to safeguard customers?

***Q: From which sector, buy-side or sell-side, does the primary demand for GenAI in capital markets currently originate?***

In my view, capital market firms (buy-side and sell-side) are still in the early stages of exploring the potential of AI technology, as evidenced by multiple ongoing experiments. However, their focus lies in different operational domains.

Sell-side institutions being primarily service-oriented, are integrating GenAI, especially natural language processing into middle and back-office functions. This includes areas like retail customer support, internal document generation/curation, and compliance functions which are currently heavily reliant on human oversight.

Buy-side institutions with a focus on return generation, are leveraging GenAI to handle vast amounts of unstructured data. For example, front office users are often swamped with information, leading to attention/time constraints. GenAI can help synthesize this data and extract insights to make informed decisions.

***Q: How will AI reshape the nature of jobs in capital markets?***

The nature of work will change with increased use of AI. You can compare it to music where a violinist becomes a conductor – transitioning from an individual contributor to orchestrating a team. Human skills will need to evolve to manage teams of AI-agents, ask the right questions, give feedback, and intervene when things go wrong.

However, replacement of humans with AI will not be uniform. Work such as eye-balling documents, is likely to be automated quickly. Functions requiring human connection (e.g., advisory) are unlikely to be replaced. Firms willing to embrace these changes early are likely to see substantial productivity gains.

***Q: Will GenAI adoption help improve profitability?***

In my view, GenAI adoption could improve profitability for financial firms. Costs of implementing AI have reduced c.10x in the last 12 months supported by lower manufacturing costs and reduced cost for training AI models using new techniques.

Operational costs of AI systems are also being improved using economies of scale. For example, cost per query for LLMs can be reduced using batch processing, where dozens of queries are collected from users and processed simultaneously.

The competitive nature of finance also suggests that when a few early adopters leverage technology, it pushes other players to also adopt the same. Thus, financial services could see a dramatic positive change from AI in the next 18 months.

***Q: What are the headwinds that could hinder AI adoption in finance?***

Regulations can be a headwind to AI adoption in finance. Regulatory frameworks often lag technological advancements in finance due to its size and importance to the broader economy. This creates uncertainty and reluctance among institutions to embrace new AI tools until comprehensive regulatory frameworks are established.

This is a 'chicken and egg' dilemma – without exploring AI capabilities, it is difficult to frame appropriate policies; yet without clear policies, firms hesitate to adopt AI. Despite this, we are seeing financial institutions experiment with AI. This proactive approach could exert pressure on other market players and regulators, potentially accelerating the development of regulatory frameworks and AI adoption.





Spencer Reich is the founder and CEO of The SER Group. Before launching The SER Group, Spencer was Senior Partner and Chief Revenue Officer at Boosted.ai, an AI for investing company. Prior to joining Boosted.ai, Spencer was a portfolio manager at Tiger Management.

## Putting AI on Portfolio Manager's Team – Expert Interview

***Q: What is the role of AI in investing? Does it replace traditional process or the portfolio manager itself?***

Nearly four years ago when I talked to portfolio managers, there was a general aversion/fear to the use of AI. However, this has changed dramatically in recent years with the advent of ChatGPT. Today, firms have a fear of missing out (FOMO), and worry they will fall behind if they do not embrace AI.

As a fundamental investor, we have such vast amounts of data/information available to us that it is overwhelming and impossible for a human, or group of humans, to synthesize and process with the objective to stay on top of everything real-time.

For me, use of AI for investments is to understand the way a human analyst or portfolio managers think, but does what machines do best, i.e., synthesize vast amounts of data to provide you with better, smarter, more dynamic idea generation. The idea is not to replace anything that you are doing in fundamental research, but to use AI to process data, provide insights, and give a bit of an edge on information. The industry is now coming to the realization that AI technology can enhance productivity, but you still need human involvement. However, the first-mover advantage window has really shrunk to less than two years now.

Startups such as Boosted.ai offer an additive tool to give portfolio managers another lens, to give them additional information from which they can analyze a security, can get novel insights. Everything is customizable to the end user so that they can understand all the inputs. Some of our clients treat us like another associate/analyst on their team who is completely transparent, who does not have biases.

AI for investing is not something like ChatGPT that it will give an unknown output to the user's prompt. AI for investing is like Excel where the tool helps you be more productive, but it relies on the user to drive the inputs.

***Q: How does the AI for investment process work for the portfolio manager and what happens at the back end?***

The portfolio manager or an analyst can use AI in their investment process to be better, smarter, for dynamic idea generation or to monitor stocks in their portfolio or even create a new strategy.

At its core, the power of AI in investment process is to synthesize vast amounts of data, customize those data sets for the end user, allow the user to define and pick its end goal and then identify winners/laggards with clear and understandable explanations for the end user. But I will describe how the entire process works.

Users can select stocks they are interested in and control the inputs. They can choose data types – fundamental, technical, or macroeconomic. Likewise, analysts and portfolio managers can select their goals – outperform an index, achieve a certain Sharpe ratio, embed constraints into screening, incorporate turnover limits etc. These aspects are determined by users and provided to the machine to learn.

Then the machine takes all the different data, all the information from across universe, and we help the user sort the stocks by the ones that are most likely to outperform and the ones that are most likely to underperform.

After ranking the securities, every decision the machine makes, at a portfolio, model, macro-level, or all the way down to a single security level, the machine is



going to tell the user why it likes it. The pros and cons in a very simple and comprehensible manner. For example, the machine will explain why it views this to be a better risk reward trade or why it likes a particular stock better than the other.

Users can then turn those rankings into getting the sizing / timing recommendations in an actual portfolio, making day-to-day trading decisions – it is up to the end user on how they wish to use that information.

To trust the output, we need to trust the input. In the back end, the first step is to make sure that the integrity of the data is all accurate and correct. It's correct point in time, there's no survivorship bias, so that when the machine is learning from this data, the user can trust the output.

The user does not bias the system as it never prompts which data is positive or negative for the stock. The machine learns for itself and find those nuances. The machine offers most important predictive data sets at macro and security level to the portfolio manager. Further, the AI system is constantly learning and updating, unlike machine learning tools that work on already trained data sets. Thus, as data changes, the AI system will update its learnings and rankings from time-to-time.

***Q: How do we get differentiation between different portfolio managers? Where does the creativity or competitive edge come in?***

AI is a general-purpose technology, like Excel. We do not question how asset managers who use Excel can differentiate themselves from one another. It is really about how the end user incorporates this highly customizable technology into their overall research process.

Our platform is built on an open architecture system, and it is almost impossible to have two models give the same output. This can be attributed to the almost unlimited permutations (i.e., number of data points). There are thousands of different data points that one can potentially select from. You can select your investment horizon, goals, and even factor controls (e.g., Sharpe ratio goals, volatility goals) depending on what you are trying to achieve.

By the time these alpha-generating idea screens are built, developed, and ultimately rolled out, they are completely customized to the end user. Even small tweaks in the data that the machine is learning from, can make big differences in what it recommends. The power is in the end user to drive inputs based on their goals.

Data integrity is ensured with the use of only publicly available data from trusted data sources. Data is further validated for accuracy. This means models rarely hallucinate and we know exactly what the machine is learning from.

Another aspect is the ability to understand inputs and outputs. It is all well and good if a ML model is making good recommendations and performing well. But if you do not understand the inputs to that model and you cannot understand why it is making its recommendations, it does not make a lot of sense for most users to use it. One needs to understand in plain English why the machine is making a particular decision – this is very important.

Lastly, we never share information from one client to another. Everything is ring-fenced. If we build and develop a model for a client and unlock some secret sauce (i.e., data, parameters, or stock picking methodology), we do not even see it. The intellectual property remains with the client.

The confluence of Web3 and AI, decentralized systems and federated operating models could partly address the limitations of the current centralized AI stack.

## Decentralized AI: The Next Frontier?

We had an open internet layer in the 1990s. Then in the 2000-2010s, BigTech companies built a walled garden and acted as self-appointed gatekeepers. We might run a risk of leaving AI, another general-purpose technology in the hands of a few large companies, that eventually would control global economics – around cloud, compute, chips, data centre – even money and finance. This could have serious social implications.

The confluence of Web3 and AI, decentralized systems and federated operating models could partly address the limitations of the current centralized AI stack. Enthusiasts argue a trustless and permissionless AI is critical to power innovations across industries and mitigate any overarching influence by a single / group of enterprises.

Much like blockchain and distributed ledgers, proponents of decentralized AI cite the systems capabilities to impose checks against mass surveillance and manipulations by governments or enterprises. It can enhance privacy / security and is easier to scale, allowing for more inclusive participation in AI development and deployment.

### Centralized vs Decentralized AI Stack

Decentralizing AI requires reimagining all layers of the AI stack, including compute power, data, model training etc. Simply adopting open-source models is not enough, especially if compute power and availability of data remain locked with a handful of enterprises. Key aspects of decentralized AI include:

Decentralized AI requires reimagining the full AI stack. This goes beyond adopting open-source models.

- **Data Storage:** Unlike traditional AI where data is often collected, stored, and processed in a central location, decentralized AI aims to distribute it across the network. This helps enhance data security / privacy by minimizing risk of unauthorized access and single point of failure.
- **Distributed Compute:** Centralized data centres are dominated by a handful of tech giants like Amazon, Microsoft and Google. Decentralized compute involves democratize access to computational power and reduced dependency on centralized entities.
- **Model Development & Training:** Training of the AI model in a distributed setup can be done through the federated learning approach, wherein the model is trained locally on each device / node and then aggregated into a global model without sharing the raw data.
- **Ownership & Monetization:** Centralized AI models are controlled by companies where users have limited rights. The models are monetized via subscriptions or APIs or token-based pricing for corporate use. Decentralized AI models develop framework to tokenize AI models for ownership and monetization.
- **Money, Tokens & Finance:** Tokenized money and cryptocurrencies would become a key underlying layer to incentivize various participants and counterparties in the decentralized network. Autonomous and decentralized AI agents would use decentralized cryptocurrencies or some tokenized form factor of money for economic and financial transactions. We discuss next how tokens would be used in the bot-to-bot world of finance.

## Tokens in a Bot World

The tech-native nature of AI and the potential autonomous system of the future inherently aligns more seamlessly with digital/tokenized money, than traditional fiat currencies. We have discussed the advent of tokenized money in our last year's GPS report ([Money, Tokens, and Games: Blockchain's Next Billion Users & Trillions in Value](#)). Tokenized money (such as stablecoins, CBDCs and cryptocurrency) are likely to see adoption in financial transactions in the bot-to-bot world of finance.

### AI Agents and Tokenized Money

We discussed in the earlier chapter that humans do not want to be thinking about their finances. AI agents will be able to use tokenized money's programmability feature to write smart contract codes based on the natural language prompts given by humans to make conditional financial transactions.

For example, a human could ask an AI agent to pay utility bills, budget next month's fuel expense, or move a balance above \$5,000 to a higher interest rate account/term deposit, while mentioning a negative list of banks money should not be moved to even if it offers better rate. All of this can be possible due to the digitally native nature of tokenized money and its ability to embed conditions.

**Is this possible in the future?** Well, some forms of bot based tokenized money transactions occur even today. According to metrics reported by Visa and blockchain data provider Allium, almost 90% of stablecoin transactions occur without the involvement of humans.<sup>48</sup> For example, total transaction volume for May 2024 (\$111 billion) plummets to just \$9 billion if adjusted for bot-driven transactions.<sup>49</sup> To be clear, we should not assume all bot activity will be trade washing, and it could also be high frequency transactions similar to algorithmic trading in traditional finance.

While commercialization of payment technologies using blockchain may pose scaling/adoption challenges as it requires humans to understand and verify software codes; an AI bot could naturally learn lines of code, find errors, and suggest fixes.

AI bots could autonomously validate blockchain transactions after verifying key details (e.g., valid counterparty, correct wallet addresses, digital signature). Regulators could also use AI bots to scan through millions of lines of code to solve money laundering crimes involving cryptocurrencies.

The question is, how do you keep an army of AI agents loyal to you or keep them from going rogue?<sup>50</sup> The use of tokens to compensate AI agents offers a novel mechanism to maintain their loyalty and operational integrity.

AI agents can be incentivized by awarding platform native tokens for task completions, to perform optimally and adhere to their designated roles or penalized for non-compliance or rogue behavior, thereby creating a self-regulating ecosystem of AI agents and the humans. After all, even AI agents want to respond well to a reward system.<sup>51</sup>

The tech-native nature of AI and the potential autonomous system of the future inherently aligns more seamlessly with digital / tokenized money.

Bots are already used to transact / transfer tokenized money. The natural tech fit of AI and tokenized money is likely to drive adoption of tokenized money in the future.

AI agents could be paid in native tokens to maintain their integrity and establish a reward system for job well done.

<sup>48</sup> Quartz, Bots Make 90% of Stablecoin Transactions, 10 May 2024

<sup>49</sup> Visa, Stablecoin Transactions, accessed on 03 June 2024.

<sup>50</sup> Metro, AI Intelligence Was Taught to Go Rogue for a Test. It Couldn't be Stopped, 29 January 2024

<sup>51</sup> Windows Central, ChatGPT Will Provide More Detailed and Accurate Responses if You Pretend to Tip It, 04 December 2023

AI systems come with inherent risks, and we are in the early stages of developing frameworks to manage the dark side of AI.

## (C.) Navigating AI's Potential Pitfalls

While AI promises advancements for finance and economic productivity, considerable challenges also exist. The dark side of AI is not just a speculative dystopia. Algorithms can reinforce existing social inequalities, GenAI can be used for misinformation and manipulation, as well as outright fraud, AI systems can obscure human accountability. We are in the early stages of developing ethical/safe usage frameworks to manage these threats.

- **Bias and Discrimination:** AI algorithms are only as good as the data they are trained on. Algorithms can perpetuate and amplify societal biases. For example, AI-powered credit decisioning models could inadvertently favor certain sections of society, potentially resulting in an increase in social and economic disparities. Addressing algorithmic bias requires attention to data collection, pre-processing, and algorithmic design to ensure fairness and inclusivity.
- **Lack of Transparency:** AI systems can be opaque, obscuring the decision-making process and underlying logic. This poses a challenge as humans cannot comprehend how the AI system arrived at its conclusion, leading to distrust and resistance to adoption. Explicability of AI models is a growing area of concern, especially since enterprise and societal adoption of AI is growing.

“

GenAI creations can be so convincing that they create a false sense of reality. This has the potential to spread misinformation, incite panic, and even destabilize economic or financial systems.

– HERVÉ TOURPE, HEAD OF DIGITAL ADVISORY UNIT, INTERNATIONAL MONETARY FUND (IMF) <sup>52</sup>

”

- **Misinformation and Manipulation:** AI-generated content could contribute to the spread of false information or misleading content at scale, manipulating public opinion, undermining trust in reliable sources and leading to confusion in society. A particular risk of AI is its ability to tell stories that resonate with an individual's pre-existing beliefs/viewpoints, reinforcing echo chambers and ideological silos.
- **Hallucination:** GenAI models are prone to hallucinate i.e., generate information not based on real data or events, but presents it as fact. Hallucination of AI models is often caused by training data limitations, the model's probabilistic nature, etc. When the model encounters a prompt for which it has insufficient/no training data, it creates a plausible but incorrect response.
- **AI-Generated Deepfakes:** AI-powered tools can generate audio, images or texts of humans that are not real. For example, deepfake technology can be used to create convincing but entirely fabricated images or videos of individuals and objects doing things they never did, which could result in defamation, blackmail, or financial fraud.
- **Amplifies Social Engineering:** AI introduces new cybersecurity risks as bad actors leverage AI capabilities to launch more sophisticated and personalized cyberattacks. Advanced AI algorithms could be exploited by malicious actors to evade traditional detection methods using GenAI-driven malware that adapts its behavior based on the target's defenses.

<sup>52</sup> International Monetary Fund (IMF), Herve Tourpe, “Artificial Intelligence's Promise and Peril”, December 2023.

Data centre power demand likely to grow at 17% CAGR to 100 GW by 2030.

- **Centralization:** AI models used by consumers, businesses and governments may be provided by a small number of private sector companies, largely based in the US or China. AI supplier dominance may echo the concentration and single point of failure risks that already exist in cloud computing. One answer is blockchain-based Decentralized AI, which we discussed in the earlier chapter. Policymakers are also focused on this risk and regulation will play an important role – as we discuss in the next chapter of this report.
- **Energy Intensive:** The coming AI boom will increase our usage of data centers. Powering these centers is compute and energy intensive. According to the International Energy Agency, data centers already represent around 1% of global electricity consumption<sup>53</sup> and they are likely to account for an increasing share in the coming years.

In the subsequent pages we delve deeper into some of the prominent risks and challenges associated with the use of AI tools.

---

<sup>53</sup> IEA, Data Centres and Data Transmission Networks, 11 July 2023



Shameek Kundu is an experienced Chief Data Officer helping banks adopt AI responsibly. Most recently, he helped build the Financial Services business at AI testing and monitoring startup TruEra (acquired by Snowflake). Prior to TruEra, Shameek was Group Chief Data Officer at Standard Chartered Bank, where he helped explore/adopt AI across multiple areas. He has served on the Bank of England's AI Public-Private Forum, and the Monetary Authority of Singapore (MAS) initiative on Fairness, Ethics, Accountability and Transparency in AI.

## Managing Risks & Building Trust – Expert Interview

### ***Q: What risks need to be overcome to drive deeper adoption of AI in finance?***

One of the biggest hurdles in the use of GenAI in corporate environments is lack of accuracy. Hallucination is not a bug, but a necessary feature of the way GenAI works. It hinders adoption wherever accuracy/reliability is important (e.g., finance, healthcare).

The finance industry's existing proficiency in utilizing statistical models also poses a hurdle for adoption, as the superiority of AI/ML-based alternatives, over traditional models, is not always evident. This can be attributed to inadequate data quality/availability, technological/organization silos, resistance to change, and the regulated nature of finance. Other AI-related concerns include data privacy/sharing, legal uncertainties, and the effectiveness of return on investments.

### ***Q: Can you elaborate on the concerns of vendor over-dependency, intellectual property rights and legal risks in context of GenAI?***

Access to required processing capability and powerful foundation models is concentrated with a few AI and cloud infrastructure vendors, raising concerns around over-dependency and possibly single points of failure. The complexity of advanced AI models and the need to store/process vast amounts of data are contributing factors.

GenAI also heightens concerns over intellectual property. For example, if an enterprise generates code using third-party AI software, can the enterprise claim copyright over the code? Likewise, what if an enterprise uses a large language model trained on data that is scraped illegally. Is the enterprise liable to damages later? Aspects around data rights and risks are still evolving.

### ***Q: How can financial institutions manage risks and build trust in AI?***

Financial institutions are well-suited to dealing with AI risks, as they have established model risk policies and frameworks that can be easily used. However, relying solely on decentralized risk management for individual AI use cases or overly centralized frameworks can present challenges. A balanced approach is imperative.

Financial institutions also need to prioritize reliability and quality in AI applications, over regulatory compliance, and audits. After all, why would an enterprise want to build an application that does not work?

Secondly, integrating AI considerations into the developmental lifecycles from inception, ensures proactive risk management. The implementation of AI should not be an isolated project. Addressing this entails collaboration between model developers and risk management functions throughout the AI lifecycle.

Lastly, one must recognize the impracticality of manual processes for effective risk management of new AI model rollouts, especially when each department is exploring multiple AI models. Financial institutions need to embrace automated solutions. Implementing ML Ops, akin to the DevOps approach seen in the software industry, could help streamline compliance and quality assurance processes.

***Q: As AI algorithms source information from the internet for learning, how can we ensure the accuracy and factual integrity of the data used?***

AI research has been incentivized to prioritize generalization over memorization, i.e., predict broader patterns rather than memorize specific examples. Therefore, GenAI models are designed to interpret queries and generate responses based on probability, rather than “remembering” specific instances. This poses a challenge in industries like financial services, where accuracy/reliability is crucial and relying on externally sourced data is not adequate.

Two common approaches used to address this challenge are: fine-tuning models using an industry or even company specific data set, and using GenAI (Large Language) Models in conjunction with trusted data sources. The latter – retrieval augmented generation (RAG) architecture – has been very popular in banking – for example to support internal Q&A or even customer chats. Here, AI tools translate queries into a format compatible with designated internal (and trusted external) databases. Relevant information is then retrieved from these designated sources and presented to users in a summarized/contextualized format using GenAI. This can help ensure responses are based on verified data sources, enhancing accuracy/reliability.



## Bias and Discrimination

Humans are prone to bias risk that can influence their decision-making processes. These biases can stem from personal experiences, cultural influences, and societal norms. Recognizing and addressing these biases is crucial for making objective and fair decisions. When it comes to AI tools, the ability of the technology to analyze and draw patterns from large datasets can also lead to stereotyping of certain groups, resulting in potential algorithmic discrimination and bias.

Unbalanced training data and data reflecting prior discrimination introduces bias in AI algorithms.

The **most basic form of bias in AI algorithms could stem from training data that may be unbalanced and/or reflect past discrimination** based on race, religion, age, sexual orientation, health, disability, political opinions etc.

For example, if a facial recognition system is trained primarily on data consisting of lighter-skinned individuals, it may perform poorly when tasked with recognizing faces with other skin tones, leading to disparities in accuracy and reliability.

AI models can amplify problem of biases deploying them at scale.

AI algorithms can make the problem of discrimination and biases worse by baking in and deploying biases at scale. **Other reasons for biases in AI algorithms may include prejudices or bias in the way variables are measured, biased cultural assumptions of developers and inadequately defined model objectives.**

For example, Amazon stopped using an experimental hiring algorithm in 2015 after finding it to be biased against women and favoring candidates who described themselves using verbs more commonly found on male engineers' resumes, such as "executed" or "captured". Further examination revealed the algorithm to be trained on several resumes submitted over the past ten years, and since most of the applicants were men, it was trained to favor men over women. Problems with the data underpinning the models' judgement also meant unqualified candidates were often recommended for all kinds of jobs.<sup>54</sup>

For example, in March 2016, Microsoft released Tay, an AI bot designed to develop conversational understanding by interacting with humans. However, the project had to be shut in just 24 hours after release as the bot started sending out inflammatory, racist, and sexually charged messages that it learnt from users. The project highlights the perils of unsupervised AI models.<sup>55</sup>

Mitigating the impact of biases requires implementation of robust data governance frameworks to ensure ethical and responsible data collection, management, and usage. Some alternatives also consider how humans and machines can co-work, where the system makes recommendations or provides options that humans double-check or choose from.

<sup>54</sup> The Guardian, "Amazon Ditched AI Recruiting Tool That Favored Men for Technical Jobs", 11 October 2018.

<sup>55</sup> The Independent, "Tay Tweets: Microsoft Shuts Down AI Chatbot Turned into a Pro-Hitler Racist Troll in Just 24 Hours", 24 March 2016.

Transparency helps explain AI outcomes, ensuring accurate predictions and minimizing biases – helping build trust in new technologies.

Research shows a significant lack of transparency in the foundation model suite today.

## Lack of Transparency & Explicability

Transparency refers to the ability to peek into the inner workings of an AI model to understand how it reaches its decisions and why it produces specific results.

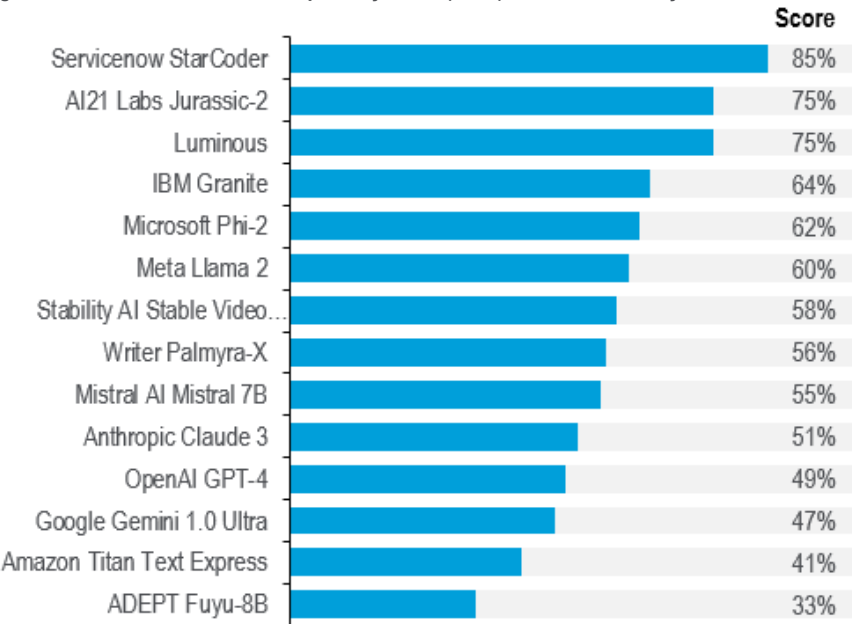
**Why Does It Matter?** People may fear AI because the underlying technology can be hard to explain. But AI does not have to be as opaque for its functioning. The point for transparency in AI models is to be able to properly explain and communicate model outcomes.

Greater transparency could help humans understand how AI algorithms interact with data to drive decisions, thereby allowing humans to assess whether the model has been thoroughly tested to make accurate predictions and minimize biases. Reducing model ambiguity could also help establish trust in new technologies.

**How Do Current AI Models Stack Up?** Researchers from Stanford, Princeton, and MIT have undertaken comprehensive assessment of the transparency of 14 major foundation model developers and their flagship models in May 2024.<sup>56</sup>

The findings, based on an index of 100 indicators, suggest a mean score of 58. Further, 11 of the 14 developers score below 65, showing there is still significant lack of transparency in the foundation model ecosystem and substantial room for improvement across developers (Figure 27).

Figure 27. Foundation Model Transparency Index (FMTI), Total Scores, May 2024



Note: Index consists of 100 indicators assessing transparency of developers' practices around developing and deploying foundation models. These indicators include three broad domains:  
[1] *Upstream*: relates to ingredients/processes of model development, including data, compute, labor.  
[2] *Model*: relate to properties/function of model, including model access, capabilities, risks, & safety mitigation.  
[3] *Downstream*: release/deployment of models, including usage policy, distribution, privacy protection, & impact.  
Source: May 2024 Foundation Model Transparency Index by Stanford University's Center for Research on Foundation Models.

<sup>56</sup> Bommasani and Klyman et al., "The Foundation Model Transparency Index v1.1," May 2024 (<https://crfm.stanford.edu/fmti/paper.pdf>).

Overall, foundation model developers are less transparent about the data, labor and compute used to build their models. There is little information about how people use foundation models, what data is used to build foundation models, and whether foundation models are trustworthy.

By contrast, developers were more transparent about user interface, capabilities, model basics, documentation, and user data protection (Figure 28).

Figure 28. Foundation Model Transparency Index Scores by Major Dimensions of Transparency, May 2024

	ADEPT Fuyu-8B	AI21 Labs Jurassic-2	Luminous	Amazon Titan Text Express	Anthropic Claude 3	ServiceNow StarCoder	Google Gemini 1.0 Ultra	IBM Granite	Meta Llama 2	Microsoft Phi-2	Mistral AI Mistral 7B	OpenAI GPT-4	Stability AI Stable Video Diffusion	Writer Palmyra-X	AVERAGE
Data	0%	60%	40%	0%	10%	100%	0%	60%	40%	40%	20%	20%	40%	50%	34%
Labour	0%	43%	71%	14%	14%	100%	29%	43%	29%	100%	100%	14%	100%	43%	50%
Compute	14%	86%	100%	0%	14%	100%	14%	100%	71%	57%	14%	14%	43%	86%	51%
Methods	0%	100%	100%	50%	75%	100%	75%	100%	75%	100%	100%	50%	75%	100%	79%
Model Basics	83%	100%	100%	83%	50%	100%	83%	100%	100%	100%	100%	50%	100%	100%	89%
Model Access	100%	67%	100%	67%	67%	100%	67%	67%	100%	100%	100%	67%	100%	33%	81%
Capabilities	80%	80%	100%	80%	100%	100%	80%	60%	100%	100%	100%	100%	60%	100%	89%
Risks	0%	57%	57%	43%	86%	100%	43%	71%	71%	29%	14%	57%	14%	14%	47%
Mitigations	0%	40%	20%	20%	40%	0%	40%	80%	60%	0%	60%	60%	0%	20%	31%
Distribution	57%	86%	100%	57%	86%	100%	57%	86%	71%	71%	71%	71%	86%	71%	77%
Usage Policy	40%	100%	100%	80%	100%	100%	100%	40%	40%	100%	40%	80%	60%	80%	76%
Feedback	67%	100%	67%	33%	33%	100%	67%	67%	33%	67%	33%	33%	67%	33%	60%
Impact	29%	29%	29%	0%	14%	14%	29%	0%	14%	0%	14%	14%	14%	14%	15%
AVERAGE	36%	73%	76%	41%	53%	86%	53%	67%	62%	66%	62%	49%	58%	57%	

Note: Scores for top-14 major foundation model providers broken down by 13 sub-domains, each of which have three or more indicators.

Source: May 2024 Foundation Model Transparency Index by Stanford University's Center for Research on Foundation Models.

Open-source AI models foster collaboration & closed-source AI models have better commercialization.

This brings us to another area of contention in AI – **Should AI models be open or closed?** 'Open-source model' refers to systems where the source code, algorithms and data are openly available to the public. Open-source AI systems encourage collaboration and innovation by allowing developers to freely access and build upon existing technology. Downsides of open-source models include copyright issues, security and cyber-risk.

By contrast, close-sourced AI models are developed and maintained by companies, and access to the source code is restricted. Closed-source AI models tend to have faster development cycles due to increased control, reduced legal issues around copyrights and reuse, better commercialization, and licensing opportunities.

**Why AI Transparency Matters?** New digital technology often appears as a black box and is easily rendered invisible, yet it can have a profound impact on billions of users across society. Transparency in the use of AI in financial services can help build trust with customers and stakeholders. It ensures fairness and eliminates biases inherent in systems. Further, transparency also encourages a culture of accountability for developers and operators and promotes ethical behavior.

Transparency in AI models builds trust with customers but is challenging to implement.

**What are the Challenges for Transparency?** The need to understand how an AI algorithm reaches its output/prediction can be important, especially in regulated industries such as finance and healthcare. Not surprisingly, governments and companies have consistently underscored the value of transparency. However, enabling transparency and explicability for AI models poses challenges such as:

Copyright and patent laws can create barriers to transparency and explainability.

Understanding AI model internals can help ensure interpretability, safety, and mitigate risks into its behavior.

- **Inherent Complexity of AI Algorithms:** Model complexity, especially of those based on deep learning, could create a 'black box' scenario where it is difficult to comprehend how the algorithm processes inputs and generates output. Often there is a trade-off between model complexity and its explicability – more complex models may deliver superior performance but may be less interpretable.
- **Trade Secrets, Competitive Advantage:** AI systems are often protected under copyright and patent laws, which could arguably create barriers to transparency. For example, access to underlying parameters used by an AI model for risk assessment could be denied in order to protect commercial interests. This is one of the areas where guardrails need to be put in place to define the scope of trade secret protection.

Regulatory bodies worldwide are working to draft principles for ethical AI use and to enhance transparency in AI system. Initiatives such as the European Commission's AI Act, the OECD AI Principles and other national guidelines are some of the areas we discuss in the following chapter.

### What Does the Inside of an AI Model Look Like?

Anthropic, a US-based AI startup, released research identifying how millions of concepts are represented within their Claude Sonnet model.<sup>57</sup> This could be a step towards making AI models safer and more understandable.

Traditionally, AI models are treated as a "black box," where the reasons behind their responses are unclear, making it hard to trust their safety and reliability. The research is a detailed examination of a large language model, offering insights that could enhance understanding of AI systems in the future.

While the discovery of deeper understanding within a LLM is important, it represents only a subset of the model's learned concepts, and further research is needed to map all features and understand their interactions.

Anthropic claims that it aims to use these insights to make AI models safer, mitigating risks and enhancing reliability through deeper understanding and control of internal representations.

The internal state of AI models consists of neuron activations, a list of numbers that are hard to interpret directly. Each concept is represented across many neurons, and each neuron contributes to multiple concepts, complicating direct observation.

Using "dictionary learning", a technique from classical machine learning, researchers isolated patterns of neuron activations (called features) that correspond to human-interpretable concepts. This approach allows any internal state of the model to be represented by a few active features instead of many active neurons, like how words are formed from letters, and sentences from words.

Initial success with dictionary learning was achieved on a small "toy" language model, identifying features like uppercase text and DNA sequences. Scaling this technique to larger models like Claude Sonnet posed engineering and scientific challenges but was ultimately successful.

<sup>57</sup> Anthropic, Mapping the Mind of a Large Language Model, 21 May 2024

Millions of features were extracted from Claude 3.0 Sonnet, providing a conceptual map of its internal states, revealing sophisticated and abstract features corresponding to entities like cities, people, and scientific fields, as well as concepts like inner conflict and gender bias.

Features were found to be multimodal and multilingual, responding to names and descriptions in various languages and images of the entities. Manipulating these features confirmed their causal role in shaping the model's behavior.

For instance, amplifying the "Golden Gate Bridge" feature caused the model to obsessively reference the bridge in unrelated contexts. Activating a feature linked to scam email recognition allowed the model to generate a scam email, demonstrating how features influence behavior. These manipulations show that features are integral to how the model represents and uses concepts, suggesting they can be used to enhance AI safety.

The research identified features related to misuse potential (like developing biological weapons), biases (such as gender discrimination), and problematic AI behaviors (like power-seeking). Understanding and controlling these features could help monitor and steer AI systems toward safer behaviors, complementing existing safety techniques like Constitutional AI.

## Misinformation, Manipulation & Hallucination

GenAI can generate plausible-looking information that is entirely inaccurate. While manipulation of information is not new, the increased quantity, improved quality, and personalization possible with GenAI could enhance its persuasive potential as it is harder to debunk or verify. This could enable users, either with mala fide intent or otherwise, to spread misinformation in society.

A major concern around AI is its misuse to manipulate and spread misinformation.

Misuse of AI-generated images and videos is of grave concern, especially in the context of elections, given the ability to convincingly impersonate politicians and spread misinformation. In 2024, elections are set to impact an unprecedented number of people worldwide spanning over 40 countries and involving more than four billion people. To combat the deceptive use of AI in this year's elections, several BigTechs, AI developers and social platforms signed a tech accord in February 2024 laying down principles to identify and address AI-generated content.<sup>58</sup>

A crude deepfake video of Ukrainian President Volodymyr Zelenskiy in 2022 sought to sow disinformation by making his military and citizens believe that the President was instructing them to lay down their arms in the conflict with Russia.<sup>59</sup>

A research study on the characteristics of AI-generated misinformation through text analysis and rapid qualitative analysis found significant linguistic differences in AI-generated misinformation as it had more emotions and cognitive processing expressions than human creations. AI-generated misinformation also tended to enhance details, communicate uncertainties, draw conclusions, and simulate personal tones.<sup>60</sup> Understanding these disparities is critical to develop effective countermeasures and safeguards against the pervasive influence of AI content.

Another challenge with AI relates to hallucination. An AI hallucination is when an AI model generates incorrect information but presents it as if it were a fact by perceiving patterns or objects that are non-existent. Hallucination of AI models is often caused by insufficient, outdated, or low-quality training data. An immediate problem with AI hallucinations is that they can disturb user trust. AI hallucinations come in different forms and have significant consequences in finance.

“

When questioned, an AI model lacks the ability to discern its knowledge boundaries (i.e., identify what it knows and does not know). In moments of pressure to respond, the AI model generates output not grounded in facts.

This could be due to reasoning limitations or insufficient training data on the subject. This phenomenon is referred to as hallucination.

– BIN REN, FOUNDER AND CEO, SIGTECH

”

<sup>58</sup> AI Elections accord website.

<sup>59</sup> Reuters, “Deepfake footage purports to show Ukrainian president capitulating”, 17 March 2022.

<sup>60</sup> Jiawei Zhou, Yixuan Zhang, Qianni Luo, Andrea G Parker, and Munmun De Choudhury. 2023. “Synthetic lies: Understanding ai-generated misinformation and evaluating algorithmic and human solutions”. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems.

There can be several forms of AI hallucination, but some of the more prominent ones include:

- **Factual Inaccuracies:** AI generating text that seems factual but is incorrect. For example, in 2023, Google's Bard chatbot wrongly stated the James Webb telescope was used to take the very first pictures of a planet outside the Earth's solar system (exoplanet). However, NASA confirms the first exoplanet images came in 2004, years before Webb's 2021 launch.<sup>61</sup>
- **Fabricated Information:** AI models can fabricate content, which is entirely fictitious, yet make it sound plausible. In 2023, a lawyer for a man suing an airline in a routine personal injury suit used ChatGPT to prepare a filing, only to find the AI model delivered fake case laws and citations.<sup>62</sup>
- **Incorrect Prediction:** AI models may predict events which are in fact unlikely to happen. For example, Google's Flu Trends, an AI-based tool for predicting flu outbreaks based on search queries was cited by researchers to make incorrect predictions and overestimate the severity of flu seasons.<sup>63</sup>
- **False Positives/Negatives:** This refers to instances where AI models fail to, or incorrectly identify a threat. For example, Tesla cars fitted with automatic emergency braking systems have been reported to suddenly slam the brakes at high speeds in response to imagined hazards. The phenomenon, known as "phantom braking", has led to several recalls by the automaker.<sup>64</sup>

Hallucination of AI models poses significant risks in accuracy, which is critical in fields like finance.

***Hallucination in AI models can be viewed as a potential bug or a novel feature.***

On one hand, AI hallucinations demonstrate the model's ability to generate content that may not necessarily be true or real. This can add value in domains where accuracy is not critical e.g., art, music. It could also help software programmers develop novel features because of hallucinated responses. However, hallucination can have adverse implications in sectors such as healthcare or financial services.

“To mitigate hallucination, AI models need to be forced to provide citations, and then output automatically cross-referenced with those citations for confirmation.

It's better to say, "I do not know" than to provide any unverifiable information. Additionally, keeping the model focused on specific documents and current events also helps minimize the risk of errors.

– KEN SENA, CO-FOUNDER AND CEO, AIERA

”

<sup>61</sup> The Guardian, "Google AI chatbot Bard sends shares plummeting after it gives wrong answer", 09 February 2023.

<sup>62</sup> Forbes, "Lawyer Used ChatGPT in Court—And Cited Fake Cases. A Judge Is Considering Sanctions", 08 June 2023.

<sup>63</sup> The Guardian, "Google Flu Trends is no longer good at predicting flu, scientists find", 27 March 2014.

<sup>64</sup> The Washington Post, "Tesla drivers report a surge in phantom braking", 02 February 2022.



Mitigating hallucinations in AI systems is important especially for sensitive sectors such as finance and healthcare. There are several techniques that can help ground an AI model, reducing hallucination dramatically, including:

- **Use of High-Quality Training Data:** It is important to train an AI model on diverse, balanced, and well-structured data. This will allow the model to minimize output bias and better understand its tasks, resulting in more effective outputs.
- **Test and Refine:** Any new AI model needs to be extensively tested for anomalies and biases before taking it live. Implemented AI models also need to be constant reviewed on an ongoing basis to further refine and recalibrate them.
- **Create a Data Template:** Data templates provide a structured framework for organizing/processing data, which can help guide the model in generating more accurate output. For example, a template for training an AI model to write text could include elements such as: a title, an introduction, body, and conclusion.
- **Limit the Number of Possible Outcomes:** Limiting the number of potential outcomes an AI model can predict can help reduce hallucination. This can be achieved using filtering tools or regularization techniques, and the model can be penalized for making predictions that are too extreme.
- **Rely on Human Oversight:** Human oversight is needed for reviewing and validating AI outputs as a final backstop measure to ensure accuracy of models.

*Seeing Is No Longer Believing* – we are seeing an increasing number of deepfakes (deceptive images, content, and videos).

AI-generated media altering speech, appearances, or actions in a way that is convincing.

## Grappling with AI-Led Deepfake Proliferation

The financial services sector is one of many to be targeted by deepfake with voice impersonation techniques becoming more and more sophisticated to defraud customers. Now more than ever financial services firms need to urgently bolster measures to protect their customer's data from the hands of deepfakes.

Deepfakes as a term first emerged in 2017 to describe synthetic media that could replace one person's likeness and voice with that of another, using AI technology. Deepfakes have been used by Hollywood to de-age actors and recreate actors' likenesses after death, by comedians in social media, by advertisers, by gamers, and among friends for fun. But they have not been seen as a threat until recently.

Emergence of GenAI in 2022-23 and the relative ease for anyone to create deepfakes at minimal cost, with publicly available data, has garnered attention, including fraudsters, who are exploring this technology for nefarious purposes.

In recent months, deepfakes have raised concern for their potential misuse in fraud, fake news, and misinformation campaigns, resulting in substantial financial and societal impact.

**What are Deepfakes?** A human impersonation created with the use of advanced technologies including AI, neural networks, and deep learning. Deepfakes represent a subset of the general category of "synthetic media" or "synthetic content".

Not just written words but also human faces, body language, and voices can be synthesized with great accuracy. Deepfake technology can be used to generate completely synthetic faces or bodies and even dead people can be brought back to life. The Salvador Dali Museum in Florida has used AI techniques to make a video installation showcasing the Spanish surrealist artist, Salvador Dali (1904-1989).

Prominent deepfakes/synthetic media in the political/social space include Obama's buzz feed deepfake (April 2018), David Beckham anti-malaria deepfake (April 2019), moon landing deepfake (July 2020), Antony Bourdain documentary with audio deepfake (July 2021).<sup>65</sup>

Figure 29. Deepfake Technology Categories <sup>66</sup>

	AUDIO	VIDEO
DYNAMIC	<ul style="list-style-type: none"> <li>Voices</li> <li>Music</li> <li>Sounds</li> </ul>	<ul style="list-style-type: none"> <li>People</li> <li>Objects</li> <li>Places</li> <li>Concepts</li> </ul>
	TEXT	IMAGE
STATIC	<ul style="list-style-type: none"> <li>Printed</li> <li>Written</li> </ul>	<ul style="list-style-type: none"> <li>People</li> <li>Objects</li> <li>Places</li> <li>Concepts</li> </ul>
	WORDS & SOUNDS	VISUALS

Source: Voice Bot 2023, Citi Global Insights

Deepfakes come in different formats – image, video, audio, or text. The formats can be applied alone or combined and may be static, recorded, or real-time.

<sup>65</sup> U.S. Homeland Security, "Increasing Threat of Deepfake Identities".

<sup>66</sup> Pindrop, 2023 Deepfake and Voice Clone Consumer Report.

- **Face Swapping Deepfakes:** The face of a person is superimposed on that of another using sophisticated deep learning algorithms. AI tools can also alter body movements, gestures, and expressions in videos. These could make it appear as though the target person is saying or doing something they have not done.
- **Voice Synthesis:** Synthetic voice recordings that mimic another person's speech patterns, accent, tone, and pitch to create fake audio messages. Voice synthesis deepfakes often rely on neural networks to analyze and replicate the acoustic features of a target individual's voice.
- **Textual Deepfakes:** AI-generated text that mimic the writing style, tone, and content of a particular individual by leveraging natural language processing. Textual deepfakes can take the form of fake news articles, false social media posts, forged emails or even deceptive online reviews and pose a significant challenge in combating misinformation.
- **Object or Event Manipulation Deepfakes:** Use of AI for altering visual content – depicting objects/events in a realistic manner. Commonly known as visual effects and used in movies/television shows to seamlessly integrate computer-generated imagery into live-action footage. It can also be used to recreate historical events by digitally manipulating archival footage or photographs.
- **Real-Time Deepfakes:** Combining multiple techniques, such as face-swapping with altered voice, body movements or object manipulation can result in the creation of sophisticated and convincing deepfakes that are real-time. These dynamic deepfake experiences respond to user input or environmental cues.

**Figure 30. Evolving Landscape of Deepfake Technology**

#### AROUND 1990S: REPLAY ATTACKS

Fraudsters obtain audio recording of the target and replay the information gathered to misdirect receiver into believing the hacker is in fact a genuine user.

#### 1990S: VIDEO REWRITE PROGRAMS

Use existing footage to create new video of persons mouthing words they did not speak in the original footage using lip-sync. These programs often used speech recognition and computer vision techniques.

#### 1990S: FACE SWAP

Common technique pre-dating AI / ML, using image editing software to put face / head of one person onto body of another.

#### EARLY 2000S: AUDIO MANIPULATION

Advancements in digital audio encoding software enabled manipulation of recorded voices with increasing precision. These tools allowed for pitch correction, time stretching, voice modulation etc.

#### 2010S: MACHINE LEARNING

Proliferation of artificial intelligence and machine learning algorithms, allowed for manipulation of audio and video content, resulting in early deepfake techniques.

#### 2014: GENERATIVE ADVERSARIAL NETWORKS

Refers to deep learning techniques used to create deepfakes. GANs work off a training set, which they use to generate new data with the same information as the training set.

#### 2016-17: DEEP LEARNING / FACE2FACE

Use of deep learning techniques to create convincing fakes. Real-time animations replacing the mouth area of its target video with an actors using AI tools. Face2Face technology helped improve computing and rendering times.

#### 2018: NEURAL VOICE CLONING

Process of creating a synthetic replica of an individual's voice through machine learning and speech synthesis technology.

#### 2020-21: AI FACE SWAP / LLMS

AI-based technique using LLMs, transformers and deep learning allowing users to replace faces in images or videos.

\* Above is only an illustrative list of key deepfake technologies and is not an exhaustive list by any means.  
Source: Citi Global Insights

## Deepfakes Fueled by Generative AI

AI text generators can create credible dialogues for making deepfake videos with minimal user input.

The wide accessibility and ease of use of GenAI technology is contributing to the proliferation of deepfakes. With user-friendly tools and platforms readily available bad actors with minimal technical expertise can create deepfakes.

For example, if someone wanted to make a convincing dialogue for a deepfake video in the past, they would have to manually draft the lines to be said. Today, an AI model can automatically draft the lines for you – only an outline of the content is needed, to generate a credible dialogue from AI text generators.

**Figure 31. Examples of Generative AI-led Deepfakes that Went Viral in Recent Years**



1 Reuters, "Image of Pope Francis wearing oversized white puffer coat is AI-generated", March 2023.

2 Reuters, "Images appearing to show Donald Trump arrest created by AI", March 2023.

3 AP News, "Image of Julian Assange in prison is AI-generated", April 2023.

4 Guardian, "Fake AI-generated image of explosion near Pentagon spread on social media", May 2023.

5 Forbes, "DeSantis Ad Showing Fake AI Images of Trump Hugging Dr. Fauci", June 2023.

6 Deutsche Welle (DW), "AI fakes in Israel's war against Hamas", October 2023.

7 CNN, "The deepfake era of US politics is upon us", January 2024.

8 NBC News, "Taylor Swift deepfakes on X falsely depict her supporting Trump", February 2024.

9 CNN, "AI 'resurrects' long dead dictator in murky new era of deepfake electioneering", February 2024.

Source: Citi Global Insights

Deepfakes risk exposing sensitive personal information and causing financial harm.

### How Do Deepfakes Affect Financial Services?

Deepfakes are a cause of concern in banking and finance, where sensitive personal identifiable information is at risk and false public information could lead to loss of money. Scams perpetuated on social media and through phone calls are targeting everyday consumers.

According to research by Voicebot.ai and Pindrop, over half of US adults report awareness of deepfake technology, with just over one-third stating they have not heard of it and another 10.6% being unsure.<sup>67</sup>

Nearly 67% of US adults expressed concern about the risks posed by deepfakes and voice clones related to banking, followed by politics and government (55%), and media and healthcare (50%).<sup>68</sup> Potential impact of deepfakes on finance:

- **Frauds and Scams:** Deepfake technology can be used to create convincing audio or video recordings of individuals (e.g., CEO impersonation scam), phishing attacks or social engineering schemes. For example, fraudsters are training their own voicebots to mimic bank IVRs as a starting point of a smishing attack to collect personal information from unsuspecting targets.
- **Undermine Identity Verification:** Deepfakes could subvert traditional onboarding processes by creating synthetic images or videos of genuine individuals. This poses a challenge for financial institutions as it could lead to the creation of fraudulent accounts which could be used for money laundering.
- **Market Manipulation:** Deepfake technology can be used to manipulate financial markets by spreading false or misleading information that can influence investor sentiment and trading behavior. False claims about a company's prospects could lead to stock price manipulation or market volatility.
- **Personalized Customer Service:** On the positive side, financial institutions could leverage technology to enhance customer service by creating hyper-personalized and interactive experiences. Virtual assistants powered by real-time deepfakes could provide more human-like interactions.
- **Better Risk Management and Compliance:** Deepfake technology can be used to identify inconsistencies in facial movement, voice patterns or speech characteristics to indicate presence of a deepfake. Deepfakes can also be used to monitor transactions and detect suspicious activities in real-time.

<sup>67</sup> Voicebot.ai and Pindrop, "Deepfake and Voice Clone Consumer Sentiment Report", October 2023.

<sup>68</sup> Voicebot.ai and Pindrop, "Deepfake and Voice Clone Consumer Sentiment Report", October 2023.

Basic deepfakes require just a computer and elementary technical skills.

Creating a high quality deepfake requires use of GANs and high-level AI skills.

## How Easy is it to Create a Deepfake?

Before AI and machine learning algorithms, basic computer software like Adobe Photoshop were used to slow down or speed up videos and lookalikes. Compared to AI-powered tools, these techniques were time consuming, required technical expertise, and are today often referred to as cheap fakes or shallow fakes.

**AI Face Synthesis:** Creating a basic deepfake using AI algorithms requires just a home computer with elementary computer skills. Fraudsters can obtain open-source or off-the-shelf applications with detailed tutorials, which contain step-by-step instructions. These techniques are less expensive, require less technical skill and are available on a larger scale.

**High Quality Deepfakes:** Creating more realistic deepfakes require the AI applications to be trained on thousands of training images of the face to be manipulated. This makes celebrities, news presenters, politicians, and other public personalities an easy target for such attacks given the ready availability of their images, audio, and videos publicly.

**GAN-Driven Fakery:** Convincing deepfakes are often created with Generative Adversarial Networks (GANs), requiring high level of skill in AI, training, and technology, along with advanced equipment and training data.

As artificial neural network technology continues to advance more rapidly, coupled with an increase in computing capabilities, we are likely to see more convincing deepfakes in the coming years that are also accessible to a wider set of population.

## How Can You Spot a Deepfake?

Technology is evolving. Humans can use a range of techniques to identify deepfakes.

Some deepfake content can be detected manually by analyzing inconsistencies. However, as technology continues to advance, the level of difficulty in identifying deepfakes manually also increases. According to a Pindrop study, people were able to identify a deepfake with 57% accuracy, only 7% better than a coin toss.<sup>69</sup>

Fraud detection and authentication technology is evolving to detect deepfakes. AI/ML models trained on data from known real and fake media are increasingly being used to identify synthetic media. Below are also a few signs individuals can use to identify and protect themselves from deepfakes:

- **Poor Quality:** Poor lip synching, patchy skin tone, flickering around edges of transposed face, poor rendering of different strands of hair or teeth, strange lighting effects, and inconsistent illumination.
- **Unnatural Eye Movement:** Anything that does not look natural – for example, the lack of eye moment or the absence of blinking.
- **Lack of Emotions:** Everything that makes us human – if the participants face does not seem to exhibit the emotions that should ordinarily go with what they are saying, it could potentially raise questions.
- **Look for Audio Cues:** Robotic-sounding voices, strange pronunciations, varying tone inflection in speech, absence of background noise or too much background noise can be seen as potential red flags.

<sup>69</sup> Based on our webinar with Vijay Balasubramaniyan from Pindrop. Full replay available at [WATCH: Finance in a Deep Fake World: Risks of Generative AI](#)

- **Slow Down and Look Again:** Despite the false sense of urgency created by scammers, it is important to pause and think before sharing any information. *Ask yourself: Can this really be true? Would you expect this to happen?*
- **Disrupt the Model:** An easy way to rattle a real-time deepfake can be to ask the video participant to wave their hands in front of their face. This disrupts the model and reveals latency issues with superimpositions over the deepfake face.

Enterprise level voice authentication and fraud detection solutions, designed to enhance security, protect sensitive financial information, and prevent fraud are available in banking, insurance, and other financial services.

These solutions use advanced voice biometrics, machine learning and behavioral analytics. In the following pages, we discuss how technological advances are driving a rapid increase in deepfake fraud, potential implications for enterprises and some of the defense mechanisms to counter deepfake fraud with Vijay Balasubramaniyan, Co-Founder and CEO of Pindrop.



## Detecting Giraffe Man's Voice – Expert Interview

### *Q: Can you share some examples of deepfake sophistication?*

Deepfakes exist in various forms (images, audio, and video). We have seen deepfake photos of Pope Francis in a puffer jacket go viral, and AI-generated deepfake songs featuring The Weeknd and Drake. In 2023, a deepfake video of Florida Governor Ron DeSantis appeared on social media announcing that the Governor was dropping out of the 2024 presidential race.

One intriguing case involved a fraudster, we dubbed Giraffe Man. The alias arose from the analysis of audio content, where anomalies hinted at a peculiar audio source. Upon closer examination, analysts noticed irregularities in sound patterns, which could only be naturally possible had the voice been generated through the vocal cords of an individual with a seven-foot-long neck.

The discovery showcased the depth of sophistication in deepfake technology as the audio sounded remarkably close to reality for any layman. It also highlights the importance of audio analysis in uncovering deepfakes as each AI engine leaves behind distinct signatures that can be identified through meticulous examination.

### *Q: How prevalent are deepfake frauds? Who are their primary targets?*

Deepfake fraud attempts have seen a notable increase in recent years, particularly post 2022. While they are still relatively rare compared to other forms of fraud, they pose a significant threat for organizations, especially those dealing with high-value transactions or sensitive information. Financial institutions and high net worth individuals are increasingly becoming prime targets for deepfake fraud due to the potential for large financial gains.

### *Q: How are deepfakes employed by fraudsters in enterprise scams?*

Fraudsters often use deepfake technology to clone the voice of senior company executives to trick employees. For example, fraudsters cloned the voice of a company director in the UAE to trick unsuspecting employees to wire transfer as much as \$35 million. The deepfake used AI models trained on publicly available video and audio content of the company executive to generate the cloned content.

In another scam involving \$0.5 million, fraudsters impersonated the CEO of a high-profile organization to generate convincing audio messages, which were used to deceive employees into executing unauthorized wire transfers. In this case, the CEO's audio/video content was not publicly available, which led the fraudsters to infiltrate the company's intranet to obtain access to the voice recordings.

### *Q: How has the landscape of deepfake technology evolved?*

We have had text-to-speech software for several years. However, they often sounded very robotic (machine-like), especially in the early days. Based on the tone of the voice and awkward pauses, it was very easy for the layman to identify when he/she was speaking to a machine. Back then, fraudsters also used simple techniques to commit scams. For example, if fraudsters chose to impersonate someone, they would simply capture (record) a piece of audio from the target person and replay it fraudulently. Initial fraud techniques were simple replay attacks.

As technology advanced, fraudsters started using audio manipulation techniques on captured audio recordings with the intent of changing specific words or slowing



Vijay Balasubramaniyan is the Co-Founder and CEO of Pindrop. He's held various engineering and research roles with Google, Siemens, Intel & others. Vijay holds patents in VoIP security and scalability, and he frequently speaks on phone fraud threats at technical conferences, including RSA, Black Hat, FS-ISAC, CCS and ICDCS. Vijay earned a PhD in Computer Science from Georgia Institute of Technology. His PhD thesis was on telecommunications security.

down the speech rate. Some instances also included increasing the frequency of the recorded voice to make it sound more feminine. These were still simple.

However, realizing the challenge of replicating diverse voices (e.g., different origins or accents), fraudsters embraced synthetic audio techniques. New technology made it possible to create the likeness of any chosen voice in any accent – this led to a new era of sophisticated scams.

Synthetic audio scams gained prominence and are in use even today. For example, to commit fraud on a person from Indian origin, fraudsters could change their accent or tone digitally to match the target. For the next scam, the audio could be changed to have an American accent or a French, German etc.

#### ***Q: How has the recent buzz around GenAI impacted deepfakes?***

In the last two years, we have seen the emergence/growth of deepfakes, which are synthetically cloned audio/video and ordinarily indistinguishable from that of the real person. The explosion of GenAI and adversarial networks can be cited as supportive factors. Today, rapid technology advances enable fraudsters to create digitally cloned audio/video using as little as three seconds of the target's real audio/video.

Aside from deepfakes, malicious actors are also misusing GenAI for more accurate /targeted scams. For example, an easy way to identify phishing emails in the past was to look for grammatical/spelling errors. However, as fraudsters use LLMs to generate content, such attacks are increasingly accurate.

#### ***Q: What are some of the defense mechanisms to counter deepfake fraud?***

Countering deepfake fraud requires a multi-faceted approach, combining advanced technology with human expertise. The first line of defense are ***liveness detection tools*** that help identify if an audio/video is a deepfake. These tools can be used to look for spatial/temporal artifacts left behind by AI tools used to create deepfakes.

Examples of spatial signs in a deepfake videos include analyzing lighting conditions in the room and checking if the shadows were falling at the right place. In audio, pay attention to fricatives (sounds resulting from the friction of two vocal tract surfaces, producing continuous, turbulent airflow, e.g., San Francisco). Deepfake tools tend to be poor at representing these aspects and use background noise to cover it up.

Meanwhile, temporal signs can be more interesting. For example, when one moves their head from side-to-side, humans tend to do many things naturally (e.g., blink their eyes), which deepfake tools find hard to replicate. In audio, when one says 'hello', their mouth is wide open; and the mouth shuts down when one says, 'Paul'. The speed at which humans do this is limited, but machines do not have limitations.

Human eyes/ears are less reliable at detecting deepfakes. Studies show humans are only 54.5% effective at identifying whether an audio is real/fake. But AI-based liveness detection tools are up to 99% effective and can do so with as few as two seconds of speech.

The second set of mechanisms to counter deepfake fraud are ***multi-factor authentication tools***. It is difficult for attackers to develop the technological sophistication to break through multiple layers of defenses which includes voice analysis, device detection, behavior analysis and other tools including liveness detection. Watermarking technology, which embeds a unique identifier in a video, can be a part of the multi-layered defense strategy as well.

## AI Amplifies Social Engineering

Advancement of AI tools, particularly LLMs, have led to an exponential growth of malicious cyberattacks – from chatbots that mimic human interactions to voice synthesis and deepfakes. New AI tools such as FraudGPT, WormGPT aim to exploit human vulnerabilities. The problem is only likely to get worse as technology advances, improving quality and greater ease of access, while reducing costs. Social engineering forms the basis of many cyber threats, including:

- **Smishing and Phishing Attacks:** AI capabilities can be exploited by fraudsters to execute sophisticated payment and bank transfer fraud. AI tools can help write business email compromise (BEC) attacks, tricking people to download malware or share sensitive personal information. This could pose a significant risk for financial transactions especially for high-net-worth individual accounts.
- **Fraudulent Money Transfers:** Using AI tools, fraudsters can develop more effective methods of impersonation, tricking their targets to make fraudulent payments to third-parties or to reveal sensitive personal information.

At an enterprise level, nefarious actors are using voice and video clones to impersonate senior company management, tricking employees into making unauthorized money transfers or revealing company secrets.

At an individual level, relatives are being dupped by scams where impersonators use a voice clone of a child claiming to be in trouble and in need of money quickly, to pay a lawyer (to post bail), or to pay-off kidnappers.

- **Undermine Identity Verification:** Advanced AI tools could subvert traditional onboarding processes/KYC verification by creating realistic fake images, videos and audios of individuals that do not exist. This is a challenge for banks as it could circumvent security measures leading to the opening of fraudulent bank accounts, which may be used for illicit activity including money laundering.
- **Market Manipulation and Misinformation:** AI tools could be used to manipulate financial markets by spreading false or misleading information that can wrongly influence investor sentiment and trading behavior. For example, false claims about a company's prospects can lead to stock price manipulation and market volatility, which could damage the company's reputation.

The rise of AI has made social engineering methods more sophisticated and difficult to detect.

While social engineering has long been a threat to business globally, the rise of AI has made social engineering methods more complex and sophisticated with more intricate and adaptive attack strategies making it more difficult to detect.

Phishing increased 341% from October 2023 to March 2024 and 856% in the 12 months from April 2023 to March 2024.<sup>70</sup> There were on average 31,000 phishing attacks daily in 2023 and it rose 1,265% from October 2022 to September 2023.<sup>71</sup>

As is often the case, bad actors impersonating someone tend to convey a sense of urgency and authenticity. They might even instruct the victim to transfer money or reveal sensitive information in a hurry, intentionally generating false sounds such as sirens or screams to intensify the scenario's realism.

For example, in February 2024, a Hong Kong multinational company lost \$25 million (HK\$200 million) to fraudsters, after an employee was tricked with digitally

<sup>70</sup> SlashNext, The State of Phishing 2024 Mid-Year Assessment

<sup>71</sup> SlashNext, The State of Phishing 2023

2-5% of global GDP is laundered each year – that's \$800 billion – \$2 trillion.  
 Nearly 90% of money laundering crimes go undetected.

recreated versions of the company's CFO in a video call. The deepfake, created from publicly available data, had the CFO make urgent requests to execute money transfers, which led the victim to make 15 transfers into five local bank accounts.<sup>72</sup>

### Sizing the Financial Crime Market

The United Nations Office on Drugs and Crime (UNODC) estimates between 2-5% of global GDP is laundered each year (that's equivalent to nearly \$800 billion to \$2 trillion)<sup>73</sup> with nearly 90% of money laundering crimes going undetected. The digital landscape is plagued with identity theft, card frauds and scams.

Cybercriminals are leveraging GenAI for malicious purposes. In 1Q 2024, the global average weekly cyber-attacks per organization increased 28% from 4Q 2023 and 5% from 1Q 2023 to 1308 – highest recorded in the past 13 quarters. Average weekly cyber-attacks in finance & banking were at 1,172 in 1Q 2024 (-2% yoy).<sup>74</sup>

Losses from internet crime in the US alone are estimated at over \$12.5 billion in 2023, a 20% increase over the previous year. Average number of complaints recorded in the same year totaled 880,418.<sup>75</sup>

### Common Ways AI Tools Facilitate Social Engineering

- **Identify Key Personnel:** AI can help identify key personal within an organization, who might have access to sensitive information by profiling the organization's workforce and their roles. This could help bad actors pinpoint individuals they need to target with social engineering campaigns.
- **Hyper-Personalization:** Analysis of vast datasets of personal information from social media, public databases, and leaked data sources can better understand user preferences/behaviors. This results in highly targeted/convincing phishing attacks to deceive targets into clicking malicious links, enter personal credentials.
- **Greater Scale and Automation:** AI-based algorithms can help automate several tasks in the attack process, including email/media creation, scenario analysis and response generation. This can help improve efficiency of the attack, resulting in more sophisticated attacks at a larger scale, targeting a wider audience.
- **AI-Generated Deepfakes:** Generate realistic synthetic content that replaces one's likeness convincingly with that of another or show images that do not exist in real-life. Deepfakes are misused by bad actors for malicious purposes to steal data, money or both, endangering finances, security, reputation, and brands.
- **Detection Evasion:** AI can analyze existing malware to identify vulnerabilities and create new variants that bypass detection. AI-driven malware can adapt its behavior based on the target's defenses, making it challenging to detect. This is misused to create new attack campaigns that evade traditional detection.

<sup>72</sup> South China Morning Post, "Everyone looked real", 04 February 2024.

<sup>73</sup> Based on data available on the United Nations Office on Drugs and Crime (UNODC) website.

<sup>74</sup> Check Point Research, "A Shifting Attack Landscapes and Sectors in Q1 2024 with a 28% Increase in Cyber-Attacks Globally" 10 April 2024.

<sup>75</sup> Federal Bureau of Investigation, "Internet Crime Report", 2023.

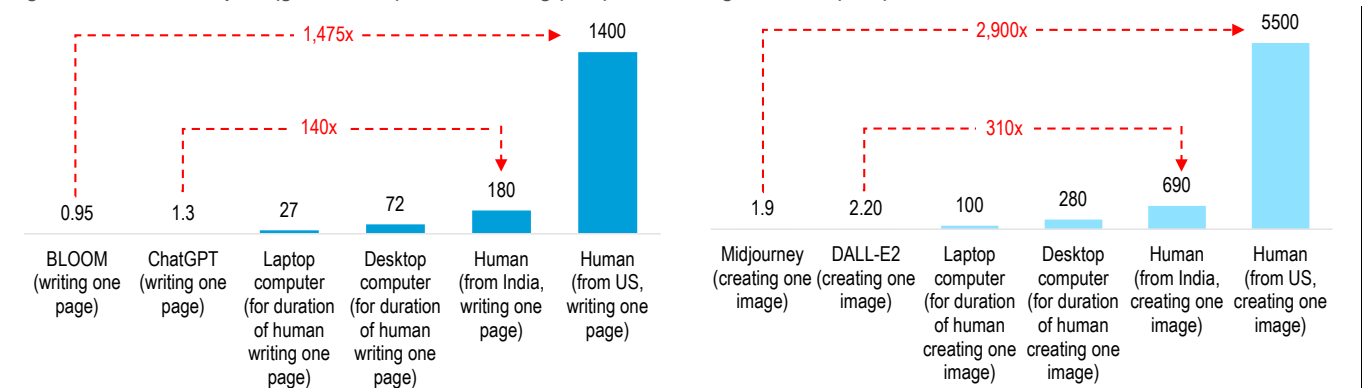
## AI – An Energy Guzzler

The computational power required for AI is doubling every 100 days and projected to increase exponentially in the next 5 years.

The computational power required for sustaining AI's rise is doubling roughly every 100 days and is projected to increase exponentially in the next five years.<sup>76</sup> Japan envisages electricity output needs to increase 35-50% by 2050 to 1.35-1.50 trillion kWh to meet the increasing demand for data centers, chip manufacturing and other energy consuming technology businesses.<sup>77</sup>

By contrast, some research suggest AI systems are getting more efficient than humans at writing content and making illustrations. For example, AI systems emit between 140 and 1,475 times less CO<sub>2</sub>e per page of text generated compared to human writers, while AI illustration systems emit between 310 and 2900 times less CO<sub>2</sub>e per image than their human counterpart (Figure 32).<sup>78</sup>

Figure 32. Carbon Footprint (grams CO<sub>2</sub>e) for Text Writing (LHS) and for Image Creation (RHS)



Note: Original data used to create a linear scale chart.

Source: Tomlinson, B., Black, R.W., Patterson, D.J. et al. The carbon emissions of writing and illustrating are lower for AI than for humans. Sci Rep 14, 3732 (2024).

With the current stage of AI revolution characterized by a race to train increasingly larger models, it is crucial to understand the emission cost of these models. We can bifurcate the potential emission costs into three primary buckets:

- **Production of Chips and Infrastructure:** Involves the environmental impact associated with the extraction of raw materials and manufacture/shipping of hardware components such as CPU, GPU and other specialized chips used for AI computations, as well as the building and maintaining of data centre infrastructure like servers, routers, modems etc.
- **Training the LLM:** Involves running of computationally intensive algorithms to pre-train and fine-tune model parameters using large datasets. Environmental costs at this stage tends to be high due to massive computation resources and energy needs as data centres are run for extended periods.
- **Ongoing Use (inference phase):** Refers to the environmental impact from the ongoing use of trained AI models in generative applications to perform tasks. This includes energy for computational power and water used in cooling servers.

<sup>76</sup> Intelligent Computing, The Latest Advances, Challenges and Future, 30 January 2023

<sup>77</sup> Reuters, Japan Sees Need for Sharp Hike in Power Output by 2050 to Meet Demand from AI, Chip Plants, 14 May 2024

<sup>78</sup> Tomlinson, B., Black, R.W., Patterson, D.J. et al., The carbon emissions of writing and illustrating are lower for AI than for humans. Sci Rep 14, 3732 (2024), 14 February 2024 <https://doi.org/10.1038/s41598-024-54271-x>

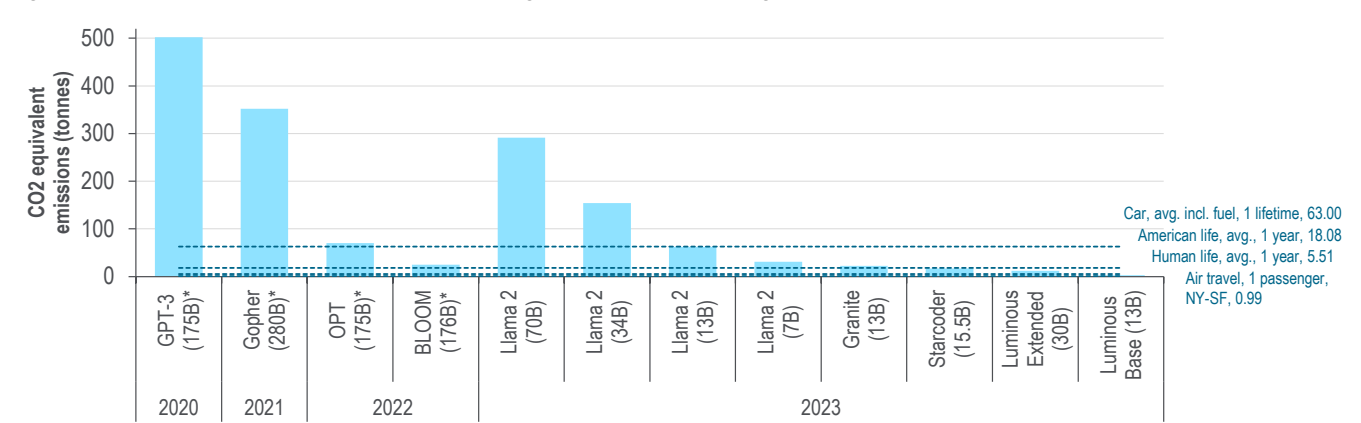
Studies show AI emits carbon surpassing 290x plane trips between New York to San Francisco, and 16x carbon emitted by an average American in a year.

While each segment contributes to the emissions cost of AI, training an AI model is computationally intensive and results in significant consumption. Training larger models generally emits more carbon. However smaller models can still have high emissions if trained on energy grids powered by less efficient energy sources.

A study on the environmental impact of AI systems by Stanford University and other researchers suggest emission data varies widely. Models such as GPT-3 (175B) emitted over 500 tons of carbon. Meta’s Llama 2 (70B) released 291.2 tons of carbon, which is over 290 times the emissions released by a single traveler on a round-trip flight from New York to San Francisco, and roughly 16 times the amount of annual carbon emitted by an average American in one year.

Foundation model providers are taking steps to mitigate their emission costs including the use of renewable energy, improved energy efficiency of data centres, efficient model training practices and carbon offsetting.

Figure 33. CO2 Equivalent Emissions (Tons) for Training Select Machine Learning Models and Real-Life Examples, 2020-23



Note: Emissions data of models marked with an asterisk were estimated by independent researchers as they were not disclosed by their developers.  
Source: Stanford University Human Centered Artificial Intelligence (HAI), Artificial Intelligence Index Report 2024.

Training an AI model is just one phase in the lifecycle and can be regarded as a one-off cost. The emission costs as an AI model goes into service (inference phase), varies by task – image generation tends to have higher emissions than text classification.<sup>79</sup> For now, overall emission costs in the inference phase appear to be limited, likely due to the still evolving scope of AI and scant research in this space.

However, as AI gains mass adoption and applications get more ubiquitous/diverse, emission costs could increase significantly. Even with relatively low per query emissions, total environmental impact from use of AI models could surpass that of training the model, as models are queried thousands, if not millions, of times daily.

<sup>79</sup> Stanford University Human Centered Artificial Intelligence, Artificial Intelligence Index Report 2024.



## (D.) Real Regulations for Artificial Intelligence

A regulatory landscape is evolving in some jurisdictions, but it is a challenging road ahead for financial services firms when it comes to implementation.

We believe the pace of implementing modern AI tools in financial services, in particular GenAI, will be relatively slow when compared to other sectors and that both the highly regulated nature of the sector and lack of 'ready to go globally aligned rules' are contributing factors.

A regulatory landscape is evolving in some jurisdictions, but it is a challenging road ahead for financial services firms when it comes to implementation because countries are moving to different speeds, taking different approaches towards regulation and in some cases changing their position on whether to regulate.

The European Union, China and Singapore are early movers with the U.S on the other hand taking its time and India changing its mind on whether to regulate or not.

Europe has taken momentous action in regulating AI with a comprehensive and prescriptive regulatory proposal backed by fines for non-compliance in an attempt to set the global tone for AI regulation. Firms in countries without any rules in place should read the EU's rules as it gives them a clear roadmap of what is required for compliance should their own jurisdiction goes down the regulatory route. A case in point is Brazil – their draft law is EU-inspired.

This patchwork of regulatory developments shows that there is no consensus on how to regulate AI globally making it hard for firms to deploy AI cross borders.

Firms with global operations will have some tough decisions to make on how they approach AI implementation. Whose rules do they follow? Depending on their risk appetite and given the highly regulated nature of the sector we anticipate a levelling up to comply with the strictest rules globally.

Smaller firms, operating in less countries and with less complex business models will be comparatively better off than large firms because they are less exposed to regulatory risk than businesses with more complex models. However, smaller firms that do not have the resources, skills, and knowledge to understand regulation will find implementation challenging.

The patchwork of regulatory developments means there is no consensus on how to regulate AI globally making it hard for firms to deploy AI cross borders.

We are entering a regulatory arms race in addition to the tech arms race. The EU's rules will be fully applicable 24 months after entry into force (the clock starts ticking once they are published in the EU's Official Journal which we expect in the summer of 2024), but some rules will be applicable sooner and potentially as soon as the end of 2024/ early 2025, Chinese rules are already live, and Singapore already has a framework for banks for GenAI. But whilst governments around the world are figuring how to rein in the bots through regulation, the technology continues to develop at lightning speed.

Firms need to be prepared for rules to be updated and changed in light of technological change, and will therefore need to build in agility in their AI implementation plans. Constant reviews of this ever-changing regulatory landscape to be on the front foot will be required.

There is light at the end of the tunnel on regulatory alignment. For starters there are common themes in the various emerging rules, AI has already made it on the agenda of global bodies in an attempt to govern AI rather than letting AI govern us and we are starting to see cross country cooperation efforts being put in place to evaluate AI.



There are common themes in the various emerging rules, and we anticipate continued international discussions on AI governance and interoperability between AI governance frameworks.

These efforts at a global level by standard setting bodies and authorities lay out high-level visions and principles for how countries and policymakers should respond to AI in their own jurisdictions, so firms have a good idea of what to anticipate. The OECD's recommendations on AI which we cover in this chapter are a good example and intend to set the blueprint for global interoperability on AI policy for policymakers.

Looking forward, we anticipate continued international discussions on AI governance and interoperability between AI governance frameworks. The EU's newly created AI office has been tasked to foster international cooperation and governance on AI with the aim of contributing to a global approach to AI.

The US and the UK have signed a Memorandum of Understanding (MoU) (2 April 2024) to align their scientific approaches on AI and to work closely together in evaluating AI models. This transatlantic collaboration will include both countries conducting a joint test on a publicly accessible AI model.<sup>80</sup>

The US and Europe are taking a similar approach and in a joint statement on 5 April 2024 committed to establishing a dialogue to deepen their collaboration on AI with a view to minimizing divergence as appropriate in their respective emerging regulatory systems and to cooperate on interoperable international standards.<sup>81</sup>

In the remainder of this chapter, we explore regulatory developments spearheaded at global level, we tease out differences and common themes and do a deep dive into the EU's AI Act. Our interview with external experts from Clifford Chance compares how AI will be regulated in Europe versus the US.

---

<sup>80</sup> Gov.uk, UK & United States announce partnership on science of AI safety, 2 April 2024

<sup>81</sup> The White House, U.S-EU Joint Statement of the Trade and Technology Council, 5 April 2024

## Global AI Regulatory Developments

AI safety, policy and regulation was placed firmly on the world stage in November 2023 in a high-profile global summit (the 'AI Safety Summit') hosted by the UK which received extensive media coverage globally being the first global summit of its kind. It convened 28 countries from around the world, the private sector, research, and civil society groups to turbocharge action on responsible developments of AI in all corners of the world.

The output? The so-called Bletchley Declaration on AI Safety.<sup>82</sup> The declaration affirms that “...for the good of all, AI should be designed, developed, deployed, and used, in a manner that is safe, in such a way as to be human-centric, trustworthy and responsible.”

**Next Steps:** Following the AI Safety Summit, on 01 February 2024, the UK's Department for Science, Innovation, and Technology (DSIT) unveiled an expert advisory panel (consisting of experts from 30 countries including US, China, Republic of Korea, EU representatives etc.) that will advise on the development of the first edition of the International Scientific Report on Advanced AI Safety. Firms should keep an eye on what comes out of the expert panel as it will be informing policymaking around the world.

The Bletchley Declaration on AI Safety is not a regulation, it is light on detail, it is a high-level expression of intent. But it matters because it paves the way for AI to be developed in a responsible manner today and in the future. It opens the door for the creation of regulation globally – including in countries that do not currently have rules in place – this will be their 'hook' to start drafting. The fact that it has included so many countries (plus China which many believed would not be at the table) adds to its gravitas.

The Bletchley Declaration on AI Safety paves the way for AI to be developed in a responsible manner today and in the future. It also opens the door for the creation of regulation globally.

The hope (or 'the ask') from the private sector is that the declaration will be used to facilitate greater international cooperation and dialogue which will ideally lead to a more harmonized approach to regulating the application of the technology. Harmonization is of particular importance considering the extraterritorial reach regulations can take. The EU's AI Act is case in point of a regulation with extraterritorial reach.

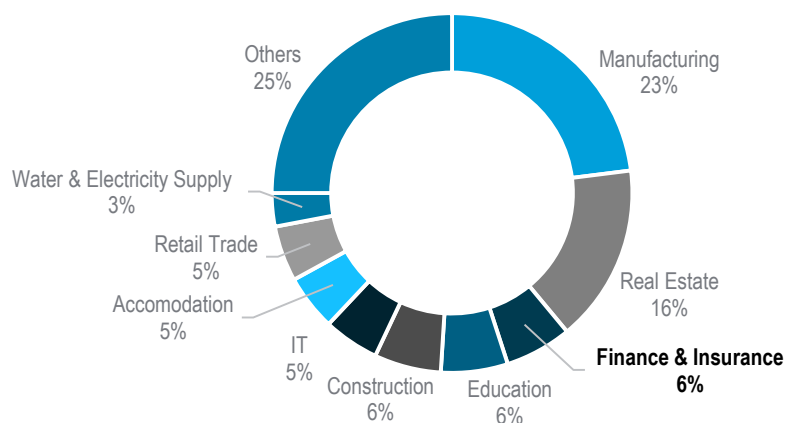
Financial service often is an international business. If firms must implement different processes and procedures to comply with different rules in all the jurisdictions they have entities, clients or do business in, it will cost them. How much will it cost?

Let's assume that following a global regulatory comparative analysis the firm concludes that in its countries of operation the EU AI's rules are the strictest and decides to apply these across the board. Whilst 2024 figures do not exist, the impact assessment conducted in 2021 by the European Commission when the EU's AI Regulation was proposed estimated an annual compliance cost of EUR 29,276.8.

This figure was multiplied by the estimated number of AI units to project the total compliance cost for the EU and the global economy from 2020-2025. The analysis concludes that by 2025 (excluding BAU compliance), total compliance costs will range from EUR 5.5 billion to EUR 11.4 billion in the EU and EUR 24.8 billion to EUR 51.9 billion globally with financial services being one of the 3<sup>rd</sup> most affected industries from a cost burden perspective.<sup>83</sup>

<sup>82</sup> UK.gov, Policy Paper, "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023", published 01 November 2023.

<sup>83</sup> European Union, "Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe: Final Report", 2021.

Figure 34. Compliance Cost Distribution Across Sectors in 2025<sup>84</sup>

Source: European Union, Citi Global Insights

The high-level global regulatory standards apply to all organizations including financial services but are subject to change, making implementation challenging for firms.

The high-level global regulatory standards (which we cover below) and declarations are non-binding and apply to all organizations – including financial services – and are thus generic. They are there to guide governments and policymakers in different countries when they start to regulate, and they are there to guide firms in their development of AI to encourage its responsible and ethical implementation.

It should also be noted that they are subject to change – i.e., they are not static. They are drafted as living documents to take account of the rapid pace of technological advancement. This is encouraging but it makes implementation more challenging for firms. Agility and constant reviews of emerging developments will need to be baked into implementation plans. Firms need to be prepared to change approaches to compliance at speed and we have already seen cases where countries have taken U-turns in their approach to AI regulation.

India for example was not considering regulating AI in a statement made by the Minister of Electronics and Information Technology (MeitY) as of 05 April 2023.<sup>85</sup> The position changed very swiftly a month later in May 2023 when the same Minister in an interview with the Times of India when asked whether AI would require a separate regulation said that their approach would flow into that direction and that AI laws would be drafted in conjunction with other like-minded nations.<sup>86</sup>

Fast forward to 01 March 2024, almost a year after the original laissez-faire approach, the aforementioned Minister published an advisory requesting firms to seek government approval before publicly releasing any AI tools that is under trial or deemed to be unreliable. In relation to the advisory, the Minister also stated that this signals a move to future of regulation in India. The position changed yet again two weeks later, on 15 March 2024, when MeitY withdrew the advisory and 'refreshed' their views on AI regulation by directing platforms to make sure that AI generated content is appropriately labelled, particularly with respect to content that is susceptible to misuse.

<sup>84</sup> European Union, "Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe: Final Report", 2021.

<sup>85</sup> Government of India, Ministry of Electronics and Information Technology, "Steps to Regulate Artificial Intelligence", 2023.

<sup>86</sup> The Times of India, "India Planning to Regulate AI Platforms like ChatGPT: IT Minister Ashwini Vaishnaw", 17 May 2023.

In April 2024, the G7 stated there were ongoing efforts to advance and reinforce interoperability between AI governance frameworks.

AI has also made it on to the agenda of the leaders of the G7 nations. They initially discussed AI regulation at their annual summit in Hiroshima on 19-20 May 2023 and their joint statement stressed the *"importance of international discussions on AI governance and interoperability between AI governance frameworks"*. More recently, in April 2024 the G7 stated there were ongoing efforts to advance and reinforce interoperability between AI governance frameworks. In October 2023 they released their international guiding principles (11 in total addressed to organisations using AI – note financial services is not singled out. These are sector agnostic principles) on AI along with a voluntary code of conduct for AI developers.<sup>87</sup>

### International Guiding Principles for Organizations Developing Advanced AI System<sup>88</sup>

Below is the list of non-exhaustive guiding principles for organizations developing advanced AI systems with the aim to promote safe, secure, and trustworthy AI.

**Principle-1:** Take appropriate measures throughout the development of advanced AI systems, including their deployment and placement on the market to identify, evaluate, and mitigate risks across the AI lifecycle.

**Principle-2:** Identify and mitigate vulnerabilities and, where appropriate, incidents and patterns of misuse after deployment, including placement on the market.

**Principle-3:** Publicly report advanced AI systems' capabilities, limitations, and domains of appropriate and inappropriate use to support transparency, thereby increasing accountability.

**Principle-4:** Work towards responsible information sharing and reporting of incidents among organizations developing advanced AI systems including industry, governments, civil society, and academia.

**Principle-5:** Develop, implement, and disclose AI governance and risk management policies grounded in a risk-based approach – including privacy policies and mitigation measures, particularly for organizations developing advanced AI systems.

**Principle-6:** Invest in and implement robust security controls including physical security, cybersecurity, and insider threat safeguards across the AI lifecycle.

**Principle-7:** Develop and deploy reliable content authentication and provenance mechanisms where technically feasible, such as watermarking or other techniques to enable users to identify AI-generated content.

**Principle-8:** Prioritize research to mitigate societal, safety, and security risks and prioritize investment in effective mitigation measures.

**Principle-9:** Prioritize the development of advanced AI systems to address the world's greatest challenges, notably but not limited to the climate crisis, global health, and education.

<sup>87</sup> Ministry of Foreign Affairs of Japan, "G7 Leaders' Statement on the Hiroshima AI Process", 30 October 2023.

<sup>88</sup> Ministry of Foreign Affairs of Japan, "Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems", 2023.

**Principle-10:** Advance the development of and, where appropriate, adoption of international technical standards.

**Principle-11:** Implement appropriate data input measures and protections for personal data and intellectual property.

Finally, and further back in time, there were several incidents in 2019 that ignited regulatory attention in relation to AI such as OpenAI's release of GPT-2 with 1.5 billion parameters trained with forty gigabytes of data to generate synthetic text using AI, 'Dactyl' a robot hand which cracked the rubrics cube (another release from OpenAI) and Samsung's transformation of facial images into video sequences – the early days of deepfakes.

Singapore put in place a Model AI Governance Framework which it launched at the World Economic Forum in Davos in January 2019. The Model Framework is based on two high-level guiding principles:<sup>89</sup>

- Organizations using AI in decision-making must ensure the decision-making process is explainable, transparent, and fair.
- AI solutions should be human-centric. Protection of human beings' interests should be the primary consideration in AI design, development, and deployment.

The Model Framework contains guidance on measure that organizations should adopt in the areas shown in Figure 35.

**Figure 35. Model AI Governance Framework**



**1. Internal Governance Structures and Measures**

Adapting existing or setting up internal governance structure and measures to incorporate values, risks, and responsibilities relating to algorithmic decision-making.



**2. Determining the Level of Human Involvement in AI – Augmented Decision-Making**

A methodology to aid organizations in setting its risk appetite for use of AI, i.e., determining acceptable risks and identifying an appropriate level of human involvement in AI-augmented decision-making.



**3. Operations Management**

Issues to be considered when developing, selecting and maintaining AI models, including data management.



**4. Stakeholder Interaction and Communication**

Strategies for communicating with an organization's stakeholders, and the management of relationships with them.

Source: Personal Data Protection Commission Singapore, Model Artificial Intelligence Governance Framework: Second Edition, Citi Global Insights

<sup>89</sup> Personal Data Protection Commission Singapore, "Model Artificial Intelligence Governance Framework: Second Edition", 2020.

The OECD's Recommendation on AI in 2019 acts as a global reference point for AI policymaking.

A few months later in April 2019, the European Commission put in place Ethics Guidelines for Trustworthy AI.<sup>90</sup> These include seven key requirements that AI systems should meet to be trustworthy: **[1]** Human agency and oversight; **[2]** Technical robustness and safety; **[3]** Privacy and Data governance; **[4]** Transparency; **[5]** Diversity, non-discrimination, and fairness; **[6]** Societal and environmental well-being; and **[7]** Accountability.

The OECD issued a Recommendation on AI in 2019. This is the first intergovernmental standard on AI and acts as a global reference point for AI policymaking. It was adopted by the OECD Council at Ministerial level on 22 May 2019 on the proposal of the Committee on Digital Economy Policy (CDEP). As per the OECD, it aims to foster innovation and trust in AI by promoting the responsible stewardship of trustworthy AI while ensuring respect for human rights and democratic values. It was amended in November 2023 to update the definition of “AI System” and clarifies that the definition includes GenAI.<sup>91</sup>

It also proposes a common understanding of key terms, such as “AI system” and “AI actors”. These are important for firms to pay attention to because they are a global anchor for policymakers and regulators to pin their own definitions to when they regulate the use of AI. Already we have seen a live example. We recently witnessed the OECD update their definition of “AI system”. And this had consequences! The EU – where the ink on the AI rules is not yet dry – swiftly updated their draft rules to mirror the OECD’s revised definition.<sup>92</sup>

The Recommendation contains five high-level values-based principles and five recommendations for national policies and international co-operation (detailed below). More specifically, the Recommendation includes two substantive sections:

**1. Principles for responsible stewardship of trustworthy AI: the first section sets out five complementary principles relevant to all stakeholders:**

- Inclusive growth, sustainable development, and well-being
- Human-centered values and fairness
- Transparency and explainability
- Robustness, security, and safety
- Accountability

**2. National policies and international co-operation for trustworthy AI: this section includes five recommendations to Members and non-Members to implement in their national policies and international co-operation:**

- Investing in AI research and development
- Fostering a digital ecosystem for AI
- Shaping an enabling policy environment for AI
- Building human capacity and preparing for labor market transformation
- International co-operation for trustworthy AI

<sup>90</sup> European Commission, “Ethics Guidelines for Trustworthy AI”, 08 April 2019.

<sup>91</sup> OECD, Recommendation of the Council on Artificial Intelligence, 03 May 2024

<sup>92</sup> OECD, “Explanatory Memorandum on the Updated OECD Definition of an AI system”, 05 March 2024.

In early May 2024 the OECD took another step forward in updating the principles in response to developments on GenAI. The updated principles, as per the OECD, directly address AI-associated challenges involving privacy, intellectual property rights, safety, and information integrity.

Firms should anticipate further updates and it is essential therefore essential to keep monitoring policy developments. The updated Principles also highlight the need for jurisdictions to work together to promote interoperable governance and policy environments for AI – which leads us on to our next section, common themes.

### Differences Emerge, But Common Themes Remain

Whilst jurisdictions are taking different approaches to regulating AI when it comes to timelines (for example some are moving at speed whilst others are taking more of a wait and see approach) and with some differences to the substance of regulation there are many similarities that can be drawn.

Whilst jurisdictions are taking different approaches to regulating AI common themes are emerging.

In practice this means that firms can rely on these ‘common concerns’ when implementing AI policies and procedures. However, we still believe that that firms with international operations, firms that are highly regulated, firms that have faced regulatory fines and have had to put correction action plans in place are more likely to play it safe and look to implement controls aligned with the strictest rules to ensure some form of consistent implementation and compliance across their businesses and operations.

Let’s taking a closer look at the trends, divergences, and common themes.

- **The Global Trend Setter:** The mega regulatory influencer could be Europe. Like the approach taken with the General Data Protection Regulation (GDPR), the EU is putting in place comprehensive rules, that apply across sector and with fines for non-compliance. It wants to show the world that it is a first mover in AI regulation and to inspire other countries to do the same. Brazil is already following closely in its footsteps and putting in place rules inspired by the EU AI Act (we understand their proposed regulation also includes a tiering system for AI risk and includes financial use cases like credit scoring).

As per Commissioner Breton announcing the political agreement on the EU AI Act “Europe has positioned itself as a pioneer, understanding the importance of its role as global standard-setter.”<sup>93</sup>

EU is putting in place comprehensive rules. The US is largely relying on existing rules and frameworks.

Europe is also encouraging the voluntary commitment of firms to comply ahead of the regulatory deadline. This is a wake-up call for firms to put AI safety and on the top of their Board’s agendas (if not already done so) and to move at pace with implementation plans for compliance.

- **The Wait and See Slow Movers:** The US has been taking a different approach by largely relying on existing rules and frameworks. A ‘wait and see’ type of approach. There was movement on 30 October 2023 when President Biden signed an Executive Order (EO) on AI. The fact that the US softly flexed its muscles when it did is unsurprising when we think about the UK hosting the first global AI Safety Summit immediately after (as covered above). The US was showing to the world that it is taking some sort of action too.

<sup>93</sup> European Commission, “Statement by Commissioner Breton – The European AI Act is Here!”, 08 December 2023.



The EO instructs various agencies on a series of reports and guidelines they should publish but it does not create new laws. It may trigger new laws in the future, but this will take time. In contrast to Europe, the approach is not one of new comprehensive cross-sectoral set of binding rules. Financial services firms should look out for:

- Report from the Treasury Department on best practices to manage cybersecurity risks posed by AI published on 27 March 2024.<sup>94</sup>
- Potentially something (no clarity on detail in the EO) from the Consumer Financial Protection Bureau (CFPB) because it is encouraged to consider using its authority to address discrimination and biases against protected groups in consumer financial markets when it comes to underwriting models for bias of disparities affecting protected groups and automated collateral valuation and appraisal process.

The US has also created a new consortium under the auspices of the National Institute of Standards and Technology (NIST) dedicated to AI Safety. It will contribute to priority actions in the above mentioned EO and determine guidelines and standards for AI measurement and policy. U.S. AI companies and companies that do business in the U.S. will be impacted by these guidelines and standards.

Companies and organizations had the opportunity to become members (the application process closed on 15 January 2024) and NIST has chosen 200 out of more than 600 applicants to join the group and work on AI projects (Apple, Microsoft, Meta, NVIDIA, OpenAI, Amazon, JP Morgan, Visa, Vanguard are among those selected).<sup>95</sup>

The UK also falls into the slow mover bucket, with a 'wait until our understanding of AI risks has matured' approach.



We are going to take our time to get this right – we will legislate when we are confident that it is the right thing to do.

– DEPARTMENT FOR SCIENCE, INNOVATION & TECHNOLOGY, UK GOVERNMENT<sup>96</sup>



The UK government set out its approach to AI regulation in their whitepaper<sup>97</sup> released in February 2024. Their framework sets out 5 principles (applies across sectors) addressed to regulators to apply within their respective remits. The principles are aligned to the above-mentioned OECD principles underlying the importance of paying attention to what comes out of global standard setting bodies for the future direction of travel of domestic regimes. These include: **[1]** safety security and robustness; **[2]** appropriate transparency and explainability; **[3]** fairness; **[4]** accountability & governance; and **[5]** contestability & redress.

<sup>94</sup> US Department of Treasury, Managing Artificial Intelligence-Specific Cybersecurity Risks in the Financial Services Sector, 27 March 2024

<sup>95</sup> National Institute of Standards and Technology, AISIC Members, Created 07 February 2024, Updated 18 April 2024.

<sup>96</sup> Department for Science, Innovation & Technology, UK Government, "A Pro-Innovation Approach to AI Regulation: Government Response to Consultation", February 2024.

<sup>97</sup> Department for Science, Innovation & Technology, UK Government, "A Pro-Innovation Approach to AI Regulation: Government Response to Consultation", February 2024.

On 22 April 2024, the Financial Conduct Authority (FCA) published their AI update emphasizing how the existing regulatory framework maps to each of the 5 principles set out above. The FCA will prioritize understanding how AI is currently used in the firms they regulate so that risks can be identified and mitigated.<sup>98</sup>

The Bank of England (BoE) including the Prudential Regulation Authority (PRA) have gone through a similar exercise and have published their update in the form of a letter.<sup>99</sup>

- **Fast Movers and Shakers:** China is moving at pace. It has already regulated for GenAI. This might be seen as an attempt to be a double winner: in the technology arms race with its own version of OpenAI's ChatGPT (it should be noted that ChatGPT itself is banned in China), and also in the race to put rules in place for AI to develop on its own soil with homegrown companies.

China and Singapore are fast movers in AI regulation space.

Another country moving at speed is Singapore, issuing GenAI principles for banks by laying down best practice as part of 'Project MindForge' (this project, a collaboration between banks and the regulator, was first announced by Mr. Heng Swee Keat, Deputy Prime Minister and Coordinating Minister for Economic Policies, at the Point Zero Forum<sup>100</sup> on 26 June 2023).<sup>101</sup> The whitepaper with details of the risk framework was published on 27 May 2024.<sup>102</sup>

The project mapped several major risks across 7 dimensions (listed below) aimed to guide financial services firms in the responsible use of this technology and assesses how these risks have extended beyond the scope of current MAS' FEAT Principles (Fairness, Ethics, Accountability and Transparency):

**[1]** Accountability and governance; **[2]** Monitoring and stability; **[3]** Transparency and explainability; **[4]** Fairness and bias; **[5]** Legal and regulatory compliance; **[6]** Ethics and impact; and **[7]** Cyber and data security.

In terms of next steps, the scope will be extended to cover insurance and asset management firms and an AI governance handbook for the financial industry will be developed.

<sup>98</sup> FCA, AI Update, 22 April 2024

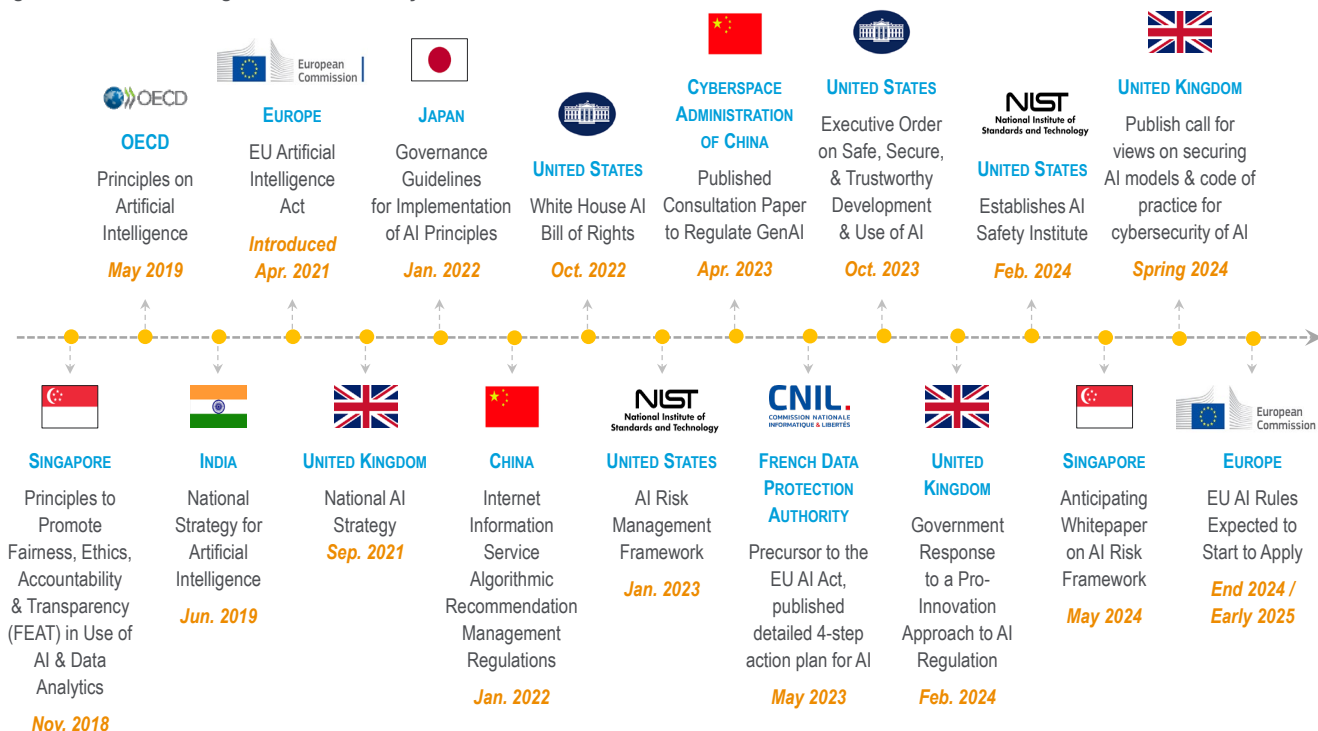
<sup>99</sup> Bank of England, DSIT-HMT Letter, 22 April 2024

<sup>100</sup> Monetary Authority of Singapore, Speech by Mr Heng Swee Keat, Deputy Prime Minister and Coordinating Minister for Economic Policies, at the Point Zero Forum on 26 June 2023

<sup>101</sup> Monetary Authority of Singapore Media Releases, "Emerging Risks and Opportunities of Generative AI for Banks: A Singapore Perspective", 2023

<sup>102</sup> Monetary Authority of Singapore, Emerging Risks and Opportunities of Generative AI for Banks: A Singapore Perspective, 27 May 2024

Figure 36. State of AI Regulations Across Key Markets



Source: Citi Global Insights

Common themes in regulating the application of AI act as a north star for firms.

Explainability is about making it possible for others to understand and dissect the AI system and algorithm with a view to understanding how it makes decisions.

## Common Themes in Regulating the Application of AI

Taking a step back, despite the differences in approaches and timelines, there are some common themes in regulating the application of AI. These common themes act as a north star for firms when they put in place their own controls, policies and procedures on AI and include:

- **Transparency and Explainability:** Not the same thing but the two go hand in hand. Transparency is in relation to disclosures and communicating information on AI to clients, colleagues, regulators, and other stakeholders. Explainability is about making it possible for others to understand and dissect the AI system and algorithm with a view to understanding how it makes decisions.
- **Eliminating Bias/Ensuring Fairness:** This means that the AI system should not discriminate unfairly against individuals nor undermine their legal rights. Firms need to be careful in the data they choose to run their AI models on (who has selected the data, is the dataset based on demographic data that is skewed towards a particular group of people for example) and who builds the models to ensure diversity. If garbage data goes into an AI model built by a non-diverse group of people, then the outcome will be biased.
- **Human-Centricity and Accountability:** Ultimately humans will need to oversee the models and be always accountable across the AI lifecycle. If something goes wrong, then it is the human who will be on the hook and will not be able to blame the model.
- **Safety and Resiliency:** This includes the potential for AI to create physical or even mental harm. This could be for example an AI bot recommending a dangerous activity by scraping data off the internet without understanding the context of the activity and the individual following the recommendation and getting harmed. This risk may be more apparent in certain sectors. It also includes resiliency and robustness to make AI safe from cyber hacks. These could cause the system to stop working all together or could prevent the model from making the decisions it has been programmed to make.

Given these risks it will be of outmost importance for firms to have in place fallback plans where a human would have to intervene quickly or switch to a different system in case of cyberattacks or other problems occurring.

On the flip side firms will have to be prepared for cybercriminals who are AI-powered. As per GCHQ, the UK's intelligence, security and cyber agency, the rapid development of novel AI tools will lead to an increase in cyberattacks and lower the barrier of entry for less sophisticated hackers to do digital harm.<sup>103</sup>

<sup>103</sup> National Cyber Security Centre, "The Near-Term Impact of AI on the Cyber Threat", 24 January 2024.

## The European Union's AI Act

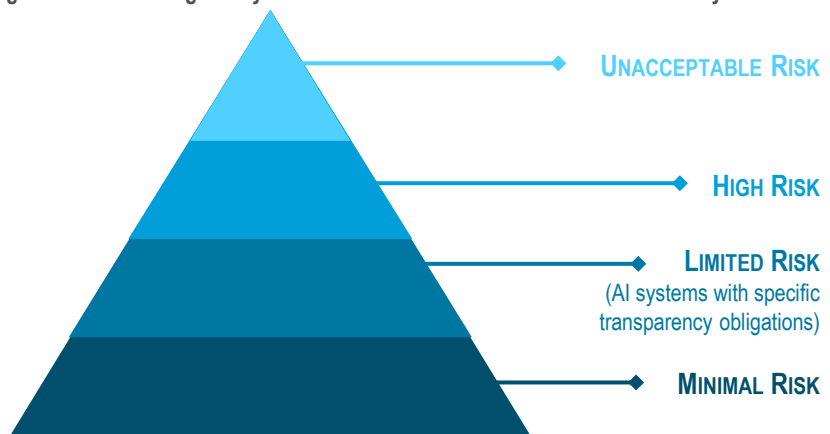
The EU's AI Act was proposed in the form of a Regulation (meaning it is directly applicable and ensures uniform implementation across the EU) on 21 April 2021. It builds on a previously published White Paper which paved the way for the draft Regulation.<sup>104</sup>

The EU's AI Act has extraterritorial reach and defines for levels of risk for AI systems.

According to the European Commission, this is the first-ever legal framework on AI which addresses the risks of AI and positions Europe in a leading role globally.<sup>105</sup>

The Act has extraterritorial reach – this is important because it means it will apply to firms which use/provide AI even though they are not EU based if AI system affects individuals in the EU. The Act lays out obligations for providers/users of AI depending on four levels of risk.

Figure 37. The EU Regulatory Framework Defines Four Levels of Risk for AI Systems<sup>106</sup>



Source: European Commission, Citi Global Insights

- **Unacceptable Risk:** AI systems falling into this category will be prohibited. As per the European Parliament following the version of the Regulation that was approved in March 2024 the following AI applications will be banned: applications that threaten citizens' rights including biometric categorization systems based on sensitive characteristics and untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases. Emotion recognition in the workplace and schools, social scoring, predictive policing, and AI that manipulated human behavior or exploits people's vulnerabilities will also be forbidden. We believe it is unlikely that financial services will fall into this category.
- **High Risk:** Extra rules will apply for AI systems falling in this category (for example, a provider of a high-risk system will need a conformity assessment certification prior to the placing of the AI system on the market, will need to put in place a robust post-market monitoring system, report any serious incidents to the national competent authority etc). Citizens will have a right to submit complaints about AI systems and receive explanations about decisions based on high-risk AI systems that affect their rights.

<sup>104</sup> European Commission Whitepaper, "On Artificial Intelligence – A European Approach to Excellence and Trust", February 2020.

<sup>105</sup> European Commission, Policies, AI Act | Shaping Europe's Digital Future

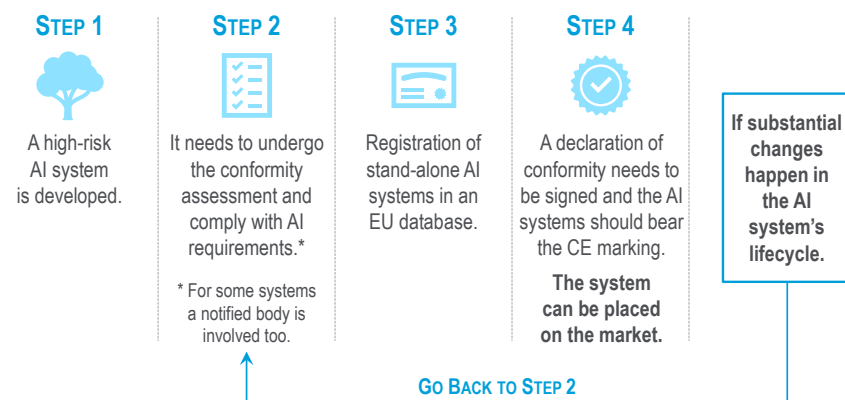
<sup>106</sup> European Commission, Policies, AI Act | Shaping Europe's Digital Future

AI tools used by financial services firm such as credit scoring models may be classified as high risk.

These fall into eight broad categories and some AI used by financial services firms may also fall in the high risk bucket such as AI which evaluates a person's creditworthiness or credit score, AI which monitors and evaluates work performance and behavior (employee monitoring for compliance), AI for recruitment and for making decisions on promotions and terminations of contracts for employment, AI used to allocate tasks based on individual behavior and AI used for risk assessment and pricing in relation to persons getting life and health insurance products.

It should be noted that systems provided for the purpose of detecting fraud in the offering of financial services and for prudential purposes to calculate credit institutions' and insurances undertakings' capital requirements should not be considered as high-risk. AI providers falling into the high-risk category will have to follow these steps.

Figure 38. How Does It All Work in Practice for Providers of High-Risk AI Systems? <sup>107</sup>



Source: European Commission, Citi Global Insights

- **Limited Risk:** Such AI systems should comply with minimal transparency requirements that would allow users to make informed decisions. When using AI systems such as chatbots, users should be aware that they are interacting with a machine so they can take an informed decision to continue or step back.
- **Minimal or No Risk:** AI falling into this category can be used freely. Examples include AI-enabled video games or spam filters. As per the European Commission, most AI systems currently used in the EU fall into this category. Voluntarily, providers of those systems may choose to apply the requirements for trustworthy AI and adhere to voluntary codes of conduct.<sup>108</sup>

We understand that some tweaks have been made to align definitions to the above-mentioned OECD definition of “AI systems”, that systems which are used for military or defense are out of scope and that the rules will not apply to AI systems used for the sole purpose of research and innovation.

With respect to general purpose AI models, the institutions have agreed on a two-tier approach: transparency requirements for all general-purpose AI models and stronger requirements for powerful models with systemic impacts across the EU Single Market. As per the European Parliament, following the new rules on general

<sup>107</sup> European Commission, Policies, AI Act | Shaping Europe's Digital Future

<sup>108</sup> European Commission, Press Corner, “Artificial Intelligence – Questions and Answers”, 12 December 2023.

purpose AI models, an AI Office within the Commission is being set up tasked to oversee these most advanced AI models, to contribute to fostering standards and testing practices, and to enforce the common rules in all Member States.

To support innovation, the political agreement reached between the Council, the Parliament, and the Commission, has clarified that AI regulatory sandboxes which are supposed to establish a controlled environment for the development, testing and validation of innovative AI systems, should also allow for testing of innovative AI systems in real world conditions.

**Where is the timeline for implementation?** The text still needs to go through lawyer-linguist checks and be published in the EU's Official Journal (OJ), which we anticipate this summer. Once the text has entered into force (20 days after publication in the OJ) the provisions will apply following a 2-year transitional period (mid-2026), but certain prohibitions will apply after 6 months (namely, bans on prohibited practices) and the rules on general purpose AI models will apply after 12 months.

Firms are encouraged to comply early despite the transitional implementation timeline.

**What should firms be doing now?** Despite the transitional timeline firms are encouraged to (or have the option to) comply early because of the 'AI Pact' which gives developers the opportunity to commit to implementing the rules before the deadlines. As per the European Commission the pact is "*seeking the voluntary commitment of industry to anticipate the AI Act and to start implementing its requirements ahead of the legal deadline.*"<sup>109</sup>

This pact is relevant in a global context because it will convene players outside and inside the EU to take part to exchange information on best practices to increase awareness of the principles underpinning the rules paving the way for more consistency in global rules.

<sup>109</sup> European Commission, Policies, AI Pact | Shaping Europe's Digital Future, 06 March 2024.



## Regulating AI Robots – Expert Interview

### *Q: What approach is the US taking to regulate AI for financial services?*

In the US, several agencies are seeking to regulate the use of AI in finance. Some are exploring the scope of existing regulations, whilst others draft new regulations and guidelines.

In 2021, we witnessed the engagement of five distinct financial regulatory bodies in the exploration of AI applications within financial institutions. This comprised the Consumer Financial Protection Bureau (CFPB), the Federal Deposit Insurance Corporation (FDIC), the National Credit Union Administration (NCUA), the Federal Trade Commission (FTC) and the Office of the Comptroller Currency (OCC). Despite initial interest, limited regulatory policies have been formally announced yet.

Several regulatory bodies already possess some form of regulatory frameworks governing the use of AI in finance. For instance, the Federal Trade Commission (FTC) Chair Lina M. Khan elucidated that new regulations were unnecessary for firms under its jurisdiction, as existing regulations already encompass the governance of unfair practices including those resulting from the use of AI (e.g., unauthorized use of customer data for AI model training).

Likewise, the New York Department of Financial Services (NYDFS) underscored prevailing regulations governing fair, equitable, transparent, and resilient financial systems to inherently include provisions for emerging technologies such as AI.

Meanwhile, regulatory bodies lacking explicit coverage of AI technologies are exploring potential parameters for forthcoming regulations. In April 2023, the US Securities and Exchange Commission (SEC) announced general ethical rules for financial services firms offering investment advisory services, stipulating the need for consistent and persistent training, governance, and oversight.

### *Q: US regulators adopted a regulation-by-enforcement approach for digital assets. Can we anticipate a similar stance for emerging technologies like AI?*

Indeed, regulators are employing a regulation-by-enforcement approach for AI. For instance, FTC Chair Lina M. Khan's approach emphasizes careful handling of datasets, especially those containing personal or biometric data, to ensure they are adequately checked for consent to use the information, before use in AI tools.

The SEC recently handed down its first fines for "AI-washing" against two investment advisors that the regulator said exaggerated their use of AI just to attract investors. SEC Chair Gary Gensler stated in a speech in March 2024 that "everyone may be talking about AI, but when it comes to investment advisers, broker dealers and public companies, they should make sure what they say to investors is true.

In our view, we are seeing enforcement actions that are instructive for what we should and should not be doing in this space.

### *Q: Can you elaborate on the EU AI Act? Does it apply to financial services?*

The EU AI Act is the first comprehensive legal framework, and it has an extra-territorial reach, i.e., it is applicable not only to entities headquartered in the EU but in many cases across the world. The framework seeks to encompass all facets of the AI value chain, spanning development, utilization, importation, and distribution, with a few exceptions.



Devika Kornbacher is a Partner and Co-Chair of Clifford Chance's Global Tech Group. She handles a wide array of transactions, including license agreements, supply agreements, data sharing agreements, reseller agreements, collaboration agreements, joint ventures, software development agreements and patent clearances.



Dessislava Savova is a Partner and Head of the Continental Europe Tech Group at Clifford Chance. Dessislava specializes in all areas of commercial law, including technology and commercial contracts and cooperations, joint ventures, digital regulations, AI, consumer protection, data and cybersecurity. She is also a lecturer at Sciences Po (Ecole de Droit) and the University of Versailles Saint Quentin-en-Yvelines on Tech & Data Law.

The regulation introduces risk-based classification, with limited adjustments for entities relying on existing sectoral procedures such as credit institutions subject to comprehensive regulatory frameworks. These adjustments are not exemptions as they do not absolve entities from regulatory scrutiny, rather they entail differentiated compliance requirements.

The act distinguishes between minimal and unacceptable risks, prohibiting practices like remote or real-time biometric identification systems, manipulative practice or those exploiting vulnerabilities. It also covers a broad spectrum of applications like credit assessment, HR management and recruiting, and insurance underwriting, imposing extensive obligations and restructuring mandates on entities. It also addresses specific risks associated with AI applications like chatbots, incorporating strict transparency requirements and obligations for data usage.

Amendments introduced by the European Parliament have further expanded the regulatory scope to encompass GenAI and foundation models with specific obligations and requirements that need to be complied with. These developments underscore the imperative for businesses to adhere to evolving regulatory standards, characterized by heightened emphasis on explicability, transparency, human-centricity, ethical considerations, and privacy protection.

***Q: Any other initiatives, voluntary measures, standards, guidelines applicable in the US/EU that firms should pay attention to, albeit not legally binding?***

In the US, as an alternative to direct comprehensive federal AI regulation, the Biden Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, implementing guidance from agencies and various initiatives form the cornerstone of the federal AI regulatory landscape. For example, the National Institute of Standards and Technology's (NIST) AI Risk Management Framework 1.0 released in January 2023 remains a popular resource for those designing an AI governance program and, in response to the Biden AI EO, the United States Patent and Trademark Office has issued inventorship guidance for AI-assisted inventions.

In the EU, the AI Act's entry into force is now imminent but its full application will still need several months. Meanwhile, the European Commission and other agencies have initiated voluntary measures to promote early adoption of key principles and have launched initiatives such as the voluntary adherence to the EU AI Pact put in place by the European Commission. Enforcement action in Europe has also commenced, leveraging existing regulations like GDPR (General Data Protection Regulation) to address AI-related concerns.

***Q: Can you highlight key regulatory developments on AI across the world?***

China has already implemented AI-related regulations, unique to its governmental structure, which has global implications like those of the EU's proposed AI Act. While some exceptions have been made for legacy systems, the core principles regarding accountability and explainability are already in effect.

From an intellectual property rights perspective, accountability and protection are critical. However, in the realm of AI, the landscape remains uncertain. While the US has stated AI output cannot be protected, other jurisdictions such as South Korea have adopted a more nuanced stance. Regulatory efforts are also underway across other countries like India, Canada, and Brazil to formulate their own AI regulations.

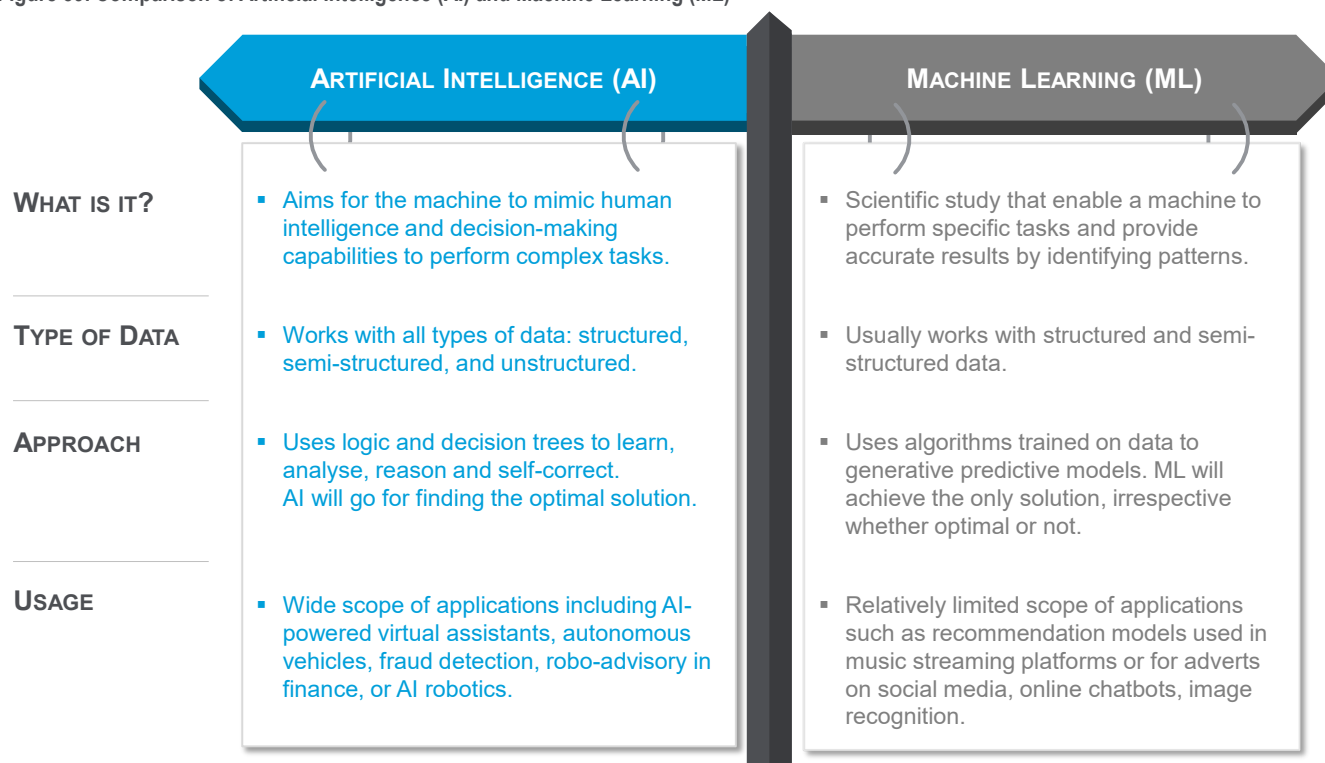
The jurisdictional variation poses a challenge for global companies seeking to safeguard their AI developments, highlighting the complexities of a borderless technology landscape and the need for international cooperation to establish a common basis for global operations.

On Join US-EU regulatory framework for AI, the Trade Council has already been commissioned to explore this space and develop regulations to make AI trustworthy, both in the EU and the US. We could see a code of conduct being adopted at the EU and US level in the context of AI.

## Appendix 1: AI & General Purpose Technologies

**What is AI?** AI is a field of computer science encompassing machine learning and deep learning. AI technology enables computers and machines to simulate human intelligence and problem-solving capabilities. Digital assistants, GPS navigation, autonomous vehicles, and GenAI tools (e.g., ChatGPT) are a few everyday examples of AI use. For a better understanding of AI and how it compares to machine learning, see Figure 39.

Figure 39. Comparison of Artificial Intelligence (AI) and Machine Learning (ML)



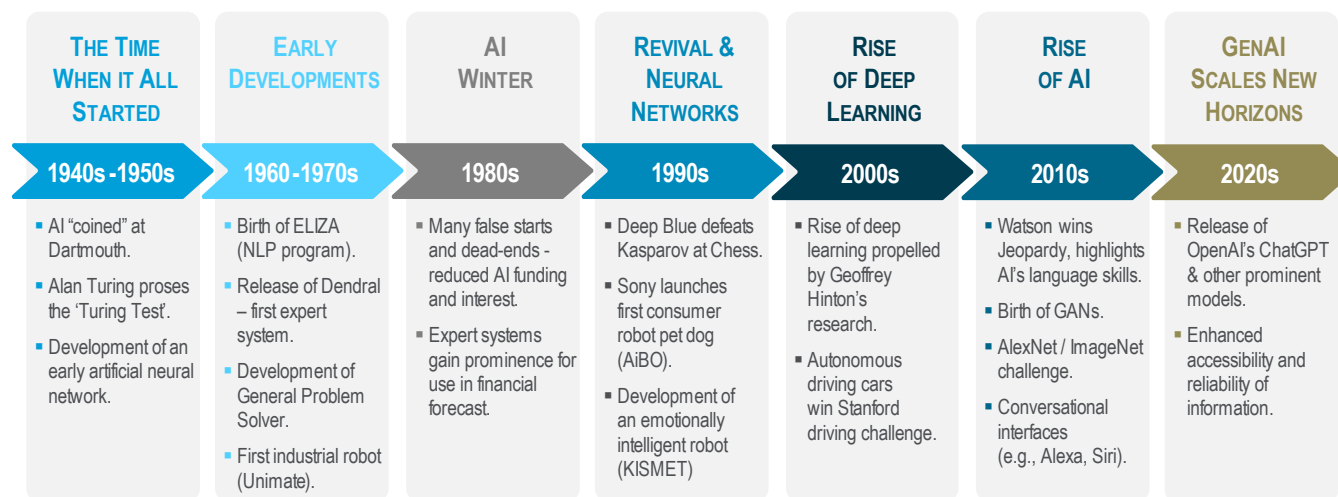
Source: Coursera, Google Cloud, Citi Global Insights

AI (and GenAI) is not new. It has gone through many cycles of hype and disillusionment starting from the 1950s (Figure 40). Most early AI use cases in banking primarily involved machine learning (ML) to either automate tasks or generate predictions.

This was achieved through ML algorithms and open-source data analysis packages that leveraged structured data. Some AI/ML-led automation also enabled banks to simplify routine processes and augment employees' capabilities, with the objective of reducing costs.

Advancements in GenAI and large language models (LLMs), including the release of ChatGPT in November 2022, mark a turning point in the history of this technology. Today, AI is a powerful building block with new applications growing every day.

Figure 40. A Timeline of Some of the Major AI Breakthroughs



Source: Citi Global Insights

GenAI, a subset of AI, refers to deep learning models that can generate high-quality text, images, videos, and other data. Generative models, at a fundamental level, capture a simplified representation of their training data and use this abstraction to generate new work that is similar, but not identical, to the original.

GenAI models have been used in statistics for several years to analyze numerical data. Origin of generative algorithms date back to the origin of AI as a field of computer science. Hidden Markov Models (HMMs) and Gaussian Mixture Models (GMMs) were amongst the first generative models devised in the 1950s.

Fast forward a few decades, in 2017, scientists working at Google published a research paper titled "Attention Is All You Need", introducing the transformer architecture, which revolutionized the field of natural language processing (NLP) and drove a flurry of new GenAI models over the coming years.

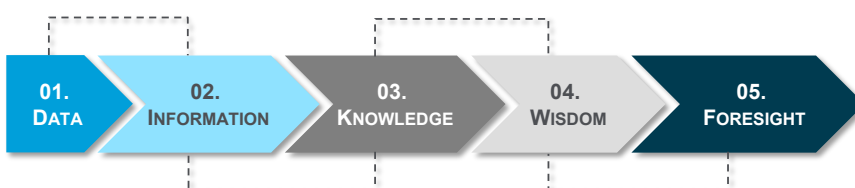
- In 2018, Google introduced BERT (Bidirectional Encoder Representations from Transformers), presenting a new approach to pre-trained NLP models with significant improvement in performance across a range of tasks by understanding context.
- In 2019, OpenAI's GPT-2 (Generative Pre-trained Transformer 2) demonstrated the ability generate coherent and contextually relevant text. GPT-2 featured 1.5 billion parameters.
- In the same year, Google AI developed T5 (Text-to-Text Transfer Transformer) capable of performing the text-based tasks that they were pre-trained for. Several other models launched include – XLNet (developed by Carnegie Mellon University and Google), RoBERTa (by Facebook), and ALBERT (by Google).
- In 2020, OpenAI's GPT-3 built upon its predecessor with 175 billion parameters, showcasing its ability to generate human-like content. Google extended its T5 model with larger parameter counts – T5-3B (3 billion parameters) and T5-11B (11 billion parameters).

- In 2021, OpenAI's CLIP combined visual and language understanding and DALL-E helped generate images from textual descriptions.
- In 2022-23, several new LLMs have been introduced, including new iterations with a greater number of parameters, to improve the model's capabilities and provide more realistic content – GPT-3.5, GPT-4, DALL-E 2 etc.

A dominant trend in GenAI has been scale, with newer models being trained on ever-increasing number of parameters (datasets) to achieve more accurate and realistic results. For example, OpenAI's GPT-4 is trained on 1.76 trillion parameters.

In any enterprise, the evolution of raw data to ability to generate data-driven foresight represents a progression towards higher order thinking and decision-making. AI is integral to this progression, transforming mere facts into insights.

Figure 41. Evolution of Raw Data to Insights in an AI-Powered Enterprise



Source: Citi Innovation Labs

### How is this AI Cycle Different? Why Now?

Past AI cycles have been driven by instruction-led AI models, but the recent shift towards conversational and observant AI models signifies a paradigm shift. This has allowed users to interact more naturally with the system and ultimately delivering more tailored and valuable experiences across applications.

Rapid growth of AI and the shift to conversational/observant models can be attributed to several factors including breakthrough in NLP techniques, increased computational power, greater availability of quality data, user-friendly interfaces, changing customer expectations and demand etc.

Figure 42. Factors Driving Rapid AI Growth in Recent Years



Source: Citi Global Insights

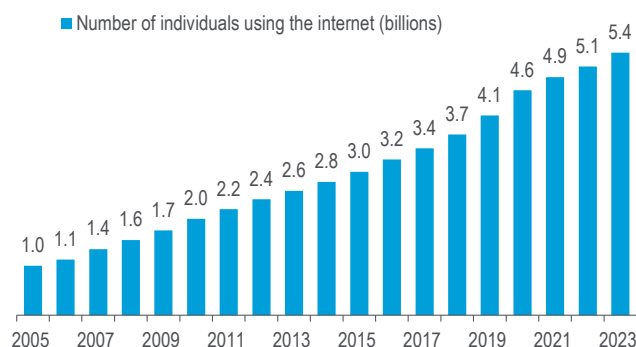
- **Increase in Compute Power:** AI research in previous decades was often limited by the availability of computational power. Advancements in microprocessor technology, coupled with miniaturization of components, have enabled performance of complex calculations with relative ease in small devices.

Modern AI relies heavily on Graphics Processing Units (GPUs) and Tensor Processing Units (TPUs) that help accelerate the training and operation of deep learning models. New-age GPUs and TPUs offer superior speed and energy efficiency that enable them to handle a wide range of content.

- **More Devices, More Data:** Early AI systems often relied on small, manually curated datasets. However, the volume of data generated worldwide has grown exponentially. Today, we have nearly 5.4 billion internet users globally as of 2023, nearly double the number from 2013 (Figure 43).<sup>110</sup>

The number of connected IoT devices (connected televisions, smart watches etc.) have also increased rapidly, totaling nearly 18 billion as of 2024 and is estimated to grow rapidly over the decade to 2030 (Figure 44).<sup>111</sup>

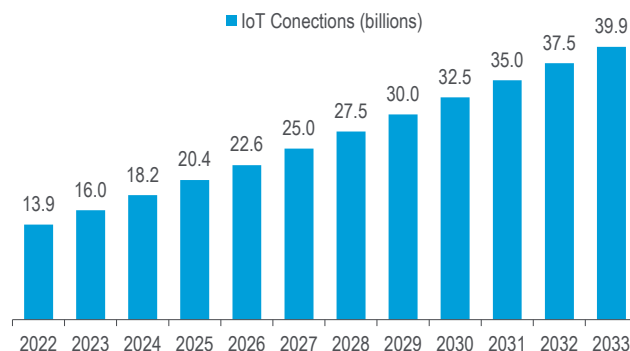
**Figure 43. Number of Internet Users Worldwide**



Data accessed on 23 May 2024.

Source: International Telecommunication Union, ITU DataHub, Citi Global Insights

**Figure 44. Number of Connected IoT Devices Worldwide**



Refers to as connections to remote sensing and actuating devices and includes associated aggregation devices.

Source: Transforma Insights, Citi Global Insights

An almost incomprehensible amount of data is created every day. In 2023, an estimated 120 zettabytes of data were generated, compared to 64 zettabytes in 2020 and just 9 zettabytes a decade ago (2013).<sup>112</sup> For reference, 1 zettabyte (ZB) is equivalent to 1 billion terabyte (TB) or 1 trillion gigabytes (GB). This translates to approximately 328 million terabytes per day.

- **Proliferation of Open-Source Software/Algorithms:** Open-source foundation models has enabled quick adoption of AI, without the need to build extensive model infrastructure from scratch. Open-source software promotes transparency as the underlying code is available for anyone to inspect, modify and improve.

We are witnessing an ongoing shift to open source. In 2023, 149 foundation models were released, nearly double that of the previous year. Even more notable is that total number of open-source models rose 10x – from just 9 projects in 2021, to 98 in 2023 (Figure 45).<sup>113</sup> Number of AI-related projects on GitHub surpassed 1.8 million in 2023, up from just 845 a decade ago (Figure 46). This signifies the active level of engagement from a global community of developers, and researchers.

<sup>110</sup> ITU World Telecommunication/ICT Indicators database, November 2023 version.

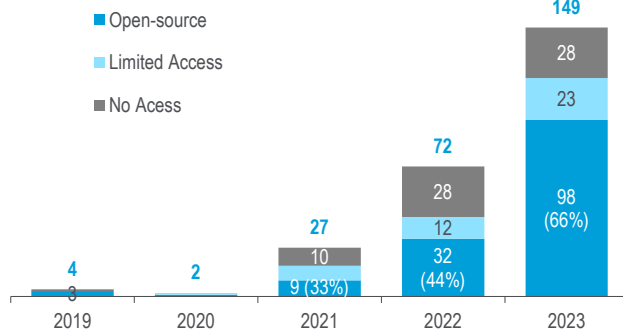
<sup>111</sup> Transforma Insights, "Current IoT Forecast Highlights", last updated: 20 May 2024.

<sup>112</sup> Exploding Topics, "Amount of Data Created Daily", 13 December 2023

<sup>113</sup> Stanford University Human Centered Artificial Intelligence, Artificial Intelligence Index Report 2024.

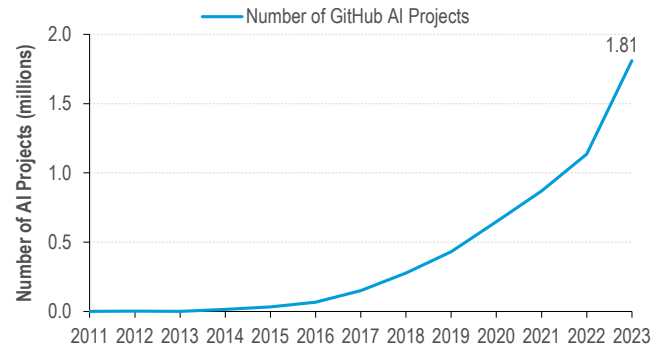


Figure 45. Number of Foundation Models by Access Type, 2019-2023



Source: Bommasani et al. 2023, Stanford University Human-Centered Artificial Intelligence (Artificial Intelligence Index Report 2024), Citi Global Insights

Figure 46. Number of GitHub AI Projects, 2011-23



Source: GitHub 2023, Stanford University Human-Centered Artificial Intelligence (Artificial Intelligence Index Report 2024), Citi Global Insights

It also showcases the wealth of available resources including code repositories, pre-built algorithms and ready-to-use interfaces that can help developers get started with AI app development. These resources provide a low-cost alternative to proprietary software, reducing the financial barriers to entry for more entities including startups and academic institutions.

Open-source AI frameworks and libraries, such as TensorFlow, PyTorch, Keras have become industry standards. Their widespread adoption has facilitated easier integration, interoperability, and consistency across various AI projects.

- **User-Friendly and Intuitive Interface:** User friendly and naturally intuitive interfaces of modern AI systems have contributed to its rapid growth. Historically, AI systems were complex with programming language environments, making it accessible to a limited few with technical expertise. This restricted development and use of AI to a small group of specialized researchers and engineers.

The evolution of AI with user-friendly interfaces where code, visualizations and text can be combined has led to greater user experimentation. The release of OpenAI's ChatGPT is often referred to as a watershed moment for consumer AI, as it enabled human-like text generation and understanding.

Today, several AI models offer ease-to-use interfaces that allow even non-technical users to interact with the model in a conversational manner, making AI more approachable and practical for everyday use.

- **Accessibility of No-Code/Low-Code:** Growth of no-code/low-code (NCLC) platforms have empowered non-developers or developers with little knowledge of custom coding and scripting to build AI-powered digital solutions. This is achieved through plug-and-play or drag-and-drop user interfaces. Examples of include intelligent code completion technologies, coding assistants, automated testing tools, and tools for building applications that leverage GenAI.

Easy accessibility of NCLC tools helps reduce time and cost involved during development and testing cycles of AI projects. It also broadens the possibilities of achieving significant outcomes with minimal coding, making innovative AI applications accessible to a broader set of individuals and enterprises.

## But How Real Are AI Claims?

Back in 1770, Wolfgang von Kempelen, a Hungarian author and inventor, introduced the world to his groundbreaking invention, the Mechanical Turk, an automated chess-playing machine. This marvel amazed audiences across Europe, defeating human opponents including notable figures like Napoleon and Benjamin Franklin. However, the truth behind its success was far from magical – it was operated by a hidden person underneath the table – a deceptive hoax.<sup>114</sup>

The term "AI washing" emerged to describe the phenomenon where companies may exaggerate or misrepresent their AI capabilities to create a false buzz. This trend, fueled by increased corporate investments in AI, may put pressure on businesses to integrate AI into their operations, even if the technology is not fully developed or effective.

**AI washing** refers to a marketing tactic some companies employ to exaggerate the amount of AI technology they use in their products. The goal of AI washing is to make a company's offerings seem more advanced than they are and capitalize on the growing interest in AI technology. AI washing takes its name from greenwashing, which is when companies make false or misleading claims about the positive impact they have on the environment.

Regulators have expressed concern over AI washing. On 18 March 2024, Gary Gensler, Chair of the U.S. Securities and Exchange Commission (SEC) shared that if financial institutions claim to use AI in their investment processes, they must ensure that the representations are not false or misleading. He further warned financial intermediaries such as investment advisers and broker dealers, or companies raising money from the public that AI washing will be considered as violation of securities law.<sup>115</sup> On the same day, the SEC charged two investment advisors with false and misleading statements about their use of AI.<sup>116</sup>

In February 2023, the Federal Trade Commission (FTC) warned against making exaggerated and/or baseless claims for the use of AI in products. A quote from FTC's blog states, *"You don't need a machine to predict what the FTC might do when those claims are unsupported."*<sup>117</sup>

However, the consequences of AI hype go beyond potential distrust and could mirror past tech bubbles like the dot-com era, leading to inflated expectations, followed by disillusionment when promises failed to materialize. While some industry leaders dismiss concerns of an AI bubble, others warn of its potential impact on job security and economic stability.

Indeed, AI's role in the workforce is a contentious issue. While proponents argue that AI can augment human capabilities and create new opportunities, skeptics warn of job displacement and economic inequality. Reports of layoffs attributed to AI underscore the complex relationship between technology and employment, challenging the narrative of AI as a job creator.

<sup>114</sup> Big Think, Mechanical Turk: An elaborate 18th-century hoax that played chess like an AI robot, 14 June 2022

<sup>115</sup> SEC, Chair Gary Gensler on AI Washing, 18 March 2024

<sup>116</sup> SEC, SEC Charges Two Investment Advisers with Making False and Misleading Statements About Their Use of Artificial Intelligence,

<sup>117</sup> FTC, Keep your AI claims in check, 27 February 2023

As the AI hype cycle matures, the focus shifts towards addressing real-world challenges and ethical considerations. While AI holds immense potential, its implementation requires scrutiny and responsible practices to ensure equitable outcomes for society. Ultimately, the true impact of AI remains to be seen, prompting us to question whether it heralds a new era of innovation or conceals hidden risks beneath its glossy facade.

## GPTs and Impact on Labor and Economy

**What is General-Purpose Technology?** General-Purpose Technology (GPT) refers to the type of technology that has wide ranging applications and can be used across multiple sectors. These technologies have the potential to transform entire economies and change the way we live and work. They can create new opportunities for economic growth and innovation, while also improving our quality of life by making goods and services more accessible and affordable.

For example, invention of the printing press in medieval Europe in the 15th century, led to an increase in accessibility of knowledge. Invention of steam engine in the 18th century powered the industrial revolution and allowed for mass production of goods. We will discuss below in further detail the impact of general-purpose technologies by taking steam engine as case-example. Similarly, 19th century marked invention and wider adoption of electric lightbulbs that changed the way we lived and worked. More recently in the 20th and the 21st century internet and related digital technologies can be considered as GPTs. But can AI be a GPT?

GPTs commoditizes underlying layers of various value chains so further value can be added and realized. Internet commoditized technology layer, steam engine commoditized production and internal combustion engine commoditized movement (of humans and goods). Similarly, if AI commoditizes human intelligence, it can be considered a GPT.

## Impact of GPTs on Labor and Economy

The ongoing debate among economists about whether workers should fear technical innovations lacks a definitive answer. On one side, proponents argue that automation leads to increased labor demand (initially to substitute skilled workers with lower skilled workers), more jobs, and cheaper goods, ultimately driving consumer spending. Conversely, opponents fear that automation will render workers, particularly unskilled ones, redundant, leading to higher unemployment and income inequality.

Historically, during the first Industrial Revolution, mechanization played a pivotal role in transforming the production process. The rise and spread of steam engines, a significant technological advancement of the time saw the replacement of skilled male workers with cheaper unskilled labor, including women and children.

Research and analysis done from nineteenth-century France, leveraging detailed historical statistics and industry-level censuses from the 1840s and 1860s suggest adoption of the steam technology was significantly higher in districts with [1] low industrial labor productivity; [2] high number of workers and higher share of male workers; and [3] previous exposure to steam engines were inclined to adopt the new technology.<sup>118</sup>

<sup>118</sup> Economic History Society, the Labour Market Causes and Consequences of General-Purpose Technological Progress: Evidence from Steam Engines, 08 October 2020

While steam-powered industries did employ more female and child workers, they also increased the overall workforce, including male employees. Moreover, these industries offered higher wages to both men and women, challenging the traditional view of technological advancement as deskilling.

The adoption of steam technology during the Industrial Revolution led to higher wages and increased demand for workers, it also raised concerns about the distribution of wealth and the diminishing share of labor in output and larger share of output towards capital owners.<sup>119</sup>

Drawing parallels from the steam engine adoption, AI adoption in financial services likely to be higher where **[1]** large number of employees do repetitive and manual tasks e.g., maintenance software and coding tasks and level-one triage in transaction monitoring; **[2]** higher share of tech roles/spend and **[3]** banks and financial institutions that are higher up in their digitization journey i.e. previous exposure and adoption of digital technologies. However, it will have varying degree of impact on the different stakeholders such as the owners/shareholders, clients, and employees.

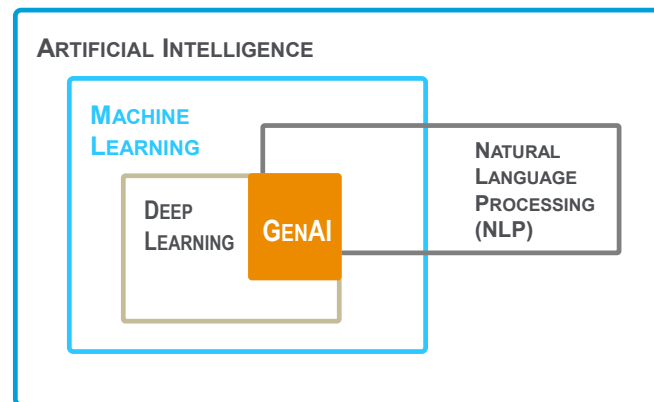
---

<sup>119</sup> MIT Technology Review, Technology and Inequality, 21 October 2014

## Appendix 2: A to Z of Key Terminologies

- **Algorithm:** An algorithm is a step-by-step procedure or set of rules designed to solve a specific problem or perform a particular task. In the context of AI, algorithms are used to process data and make decisions.
- **Artificial Intelligence (AI):** Refers to the simulation of human intelligence in machines that are programmed to think and mimic human actions. Encompasses tasks such as problem-solving, learning, perception, and decision-making.

Figure 47. Venn Diagram Showing AI, ML, NLP, DL and GenAI

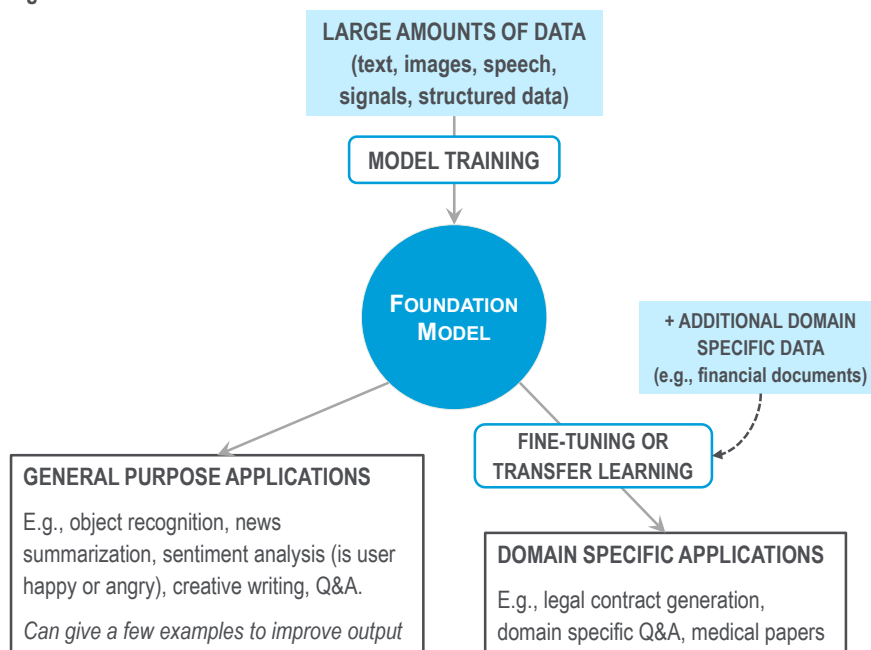


Source: Citi Innovation Labs

- **Attention Mechanism:** Components in neural network architectures that allow the model to focus on relevant parts of the input data while making predictions. They are commonly used in sequence-to-sequence models for tasks such as machine translation and text summarization.
- **Autoencoders:** Autoencoders are a type of neural network architecture used for unsupervised learning. They learn to encode unlabeled data into a compact representation and then decode it back to reconstruct the original input. Plain autoencoders are often used for data compression, reconstructing corrupted or blurry images and in anomaly detection.
- **Classification:** Classification is a supervised learning task where the goal is to categorize input data into predefined classes or categories. It is used for tasks such as email spam detection, image classification, and medical diagnosis.
- **Clustering:** A technique used in unsupervised learning to group similar data points together based on their characteristics or features. It is commonly used for tasks such as customer segmentation and anomaly detection.
- **Computer Vision:** Computer vision is a field of AI that enables computers to interpret and understand visual information from the real world. It involves tasks such as image recognition, object detection, and facial recognition.
- **Deep Learning:** A subset of machine learning that utilizes neural networks with many layers (hence "deep") to analyze and process data. It is particularly effective in tasks such as image and speech recognition.

- **Dimensionality Reduction:** Dimensionality reduction techniques are used to reduce the number of features in a dataset while preserving its important information. This helps in simplifying the computational complexity of models and avoiding overfitting.
- **Ensemble Learning:** Ensemble learning involves combining multiple machine learning models to improve the overall performance and accuracy of predictions. Techniques such as bagging, boosting, and stacking are commonly used in ensemble learning.
- **Explainable AI (XAI):** Refers to the ability of AI systems to provide explanations or justification for their decisions and predictions in a human-understandable manner. XAI is important for building trust, transparency, and accountability in AI systems, especially in critical domains such as healthcare and finance.
- **Feature Engineering:** Feature engineering involves selecting, transforming, or creating features (input variables) from raw data to improve the performance of machine learning models. It plays a crucial role in extracting relevant information and patterns from data.
- **Foundation Models:** Refers to models trained on broad data that can be adapted to a wide range of downstream tasks and applications. Foundational models are a superset of LLMs, but include other modalities such as images, code, and speech.

Figure 48. How Do Foundation Models Work?

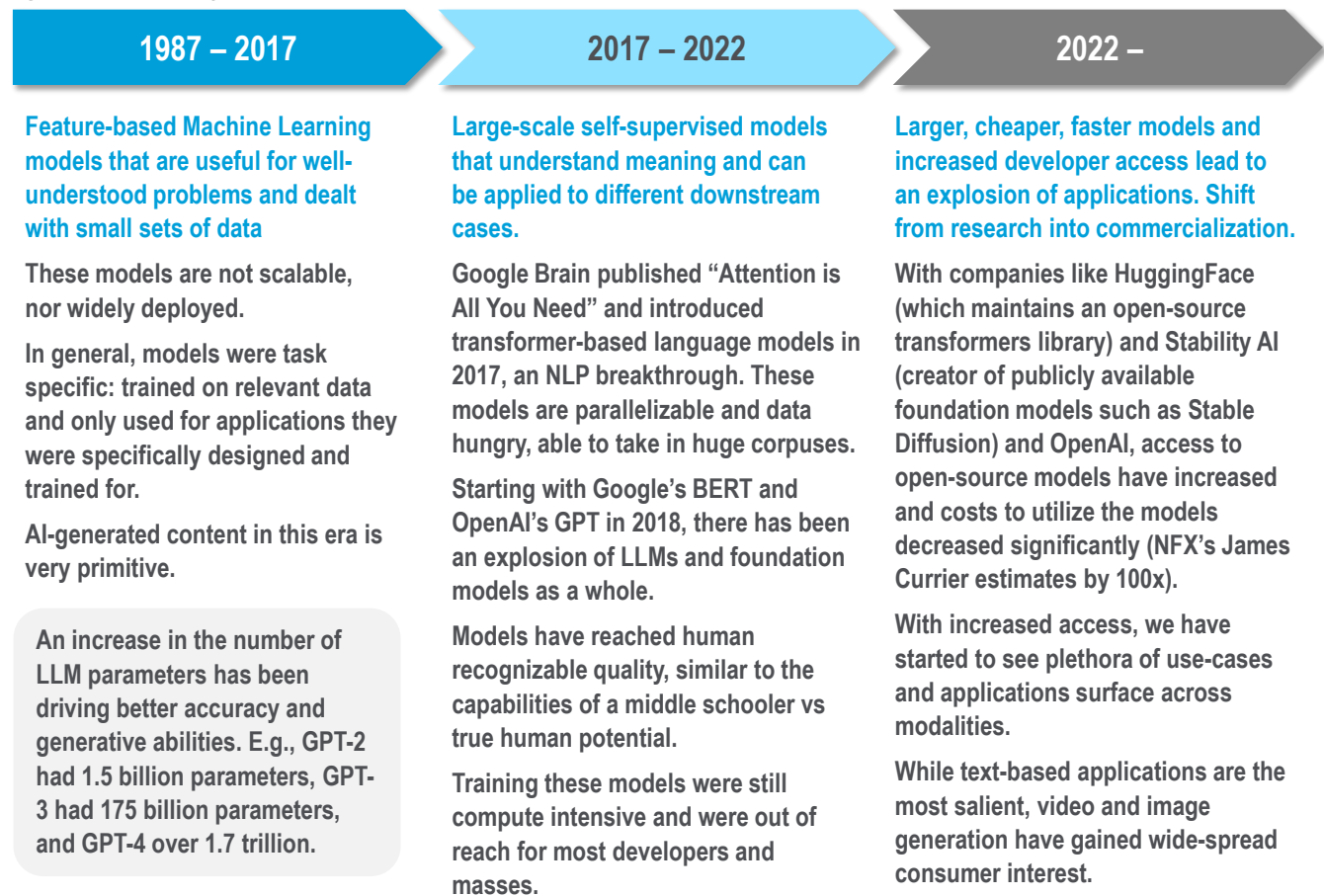


Source: Citi Ventures

- **Generative Adversarial Networks (GANs):** GANs are a type of deep learning model that consists of two neural networks, a generator, and a discriminator, which are trained simultaneously. GANs are used for generating realistic synthetic data, image-to-image translation, and other creative tasks.

- **Generative AI (GenAI):** A subset of AI, GenAI refers to deep learning models that can generate high-quality text, images, videos, and other data. Generative models, at a fundamental level, capture a simplified representation of their training data and use this abstraction to generate new work that is similar, but not identical, to the original.

Figure 49. Brief History of Generative AI



Source: Citi Ventures

- **Hyperparameter Tuning:** Involves optimizing the parameters of a machine learning model that are not learned from the data (such as learning rate, regularization strength, etc.) to improve its performance on unseen data.
- **Large Language Models (LLMs):** Refers to machine learning models that can comprehend and generate human language text. They work by analyzing massive data sets of language. LLMs are not a reference library like typical search engines where a keyword search returns results. LLMs do not need storage once the model has been trained. They need vast amounts of computational power for the training.
- **Logistic Regression:** A statistical method used for binary classification tasks, where the output variable takes on only two possible values (e.g., true/false, yes/no).



- **Machine Learning (ML):** A subset of AI that enables computers to perform tasks after learning from data, rather than being explicitly programmed for specific tasks. ML draws on concepts from statistics and probability theory, but many forms of ML go beyond traditional statistical methods.
- **Natural Language Generation (NLG):** A subfield of natural language processing that focuses on generating human-like text or speech from structured data or other input sources. NLG is used in applications such as automated report generation, chatbots, and virtual assistants.
- **Natural Language Processing (NLP):** A branch of AI that focuses on the interaction between computers and humans through natural language. It involves tasks such as language translation, sentiment analysis, and text generation.
- **Neural Networks:** Refers to computational models inspired by the structure and function of the human brain. They consist of interconnected nodes, or "neurons," organized in layers. Neural networks are used in machine learning to recognize patterns and make predictions.
- **Overfitting:** A common problem in machine learning where a model learns to perform well on the training data but fails to generalize to unseen data. Usually occurs when the model captures noise or irrelevant patterns in the training data.
- **Regression:** A supervised learning task where the goal is to predict a continuous numerical value based on input features. It is commonly used for tasks such as stock price prediction, housing price estimation, and demand forecasting.
- **Reinforcement Learning:** A type of machine learning where an agent learns to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties, which it uses to learn the optimal behavior.
- **Supervised Learning:** A type of machine learning where the model is trained on labeled data, meaning that the input data is paired with the correct output. The goal is for the model to learn the mapping between inputs and outputs.
- **Training Data:** A set of examples used to train a machine learning model. Training data often consists of input data (features) and corresponding output labels. The model learns from this data to make predictions or decisions.
- **Transfer Learning:** A technique where knowledge gained from training one ML model is transferred and applied to a different but related task or domain. It helps in leveraging pre-trained models and limited labeled data effectively.
- **Unsupervised Learning:** Unsupervised learning is a type of machine learning where the model is trained on unlabeled data. The goal is to find patterns or structures within the data without explicit guidance.
- **Zero-shot Learning:** A machine learning paradigm where a model is trained to recognize classes or categories that it has not seen during training. Zero-shot learning enables models to generalize to new tasks or domains without additional training data.

As our premier thought leadership product, **Citi Global Perspectives & Solutions (Citi GPS)** is designed to help readers navigate the most demanding challenges and greatest opportunities of the 21st century. We access the best elements of our global conversation with senior Citi professionals, academics, and corporate leaders to anticipate themes and trends in today's fast-changing and interconnected world.



All Citi GPS reports are available on our website [www.citigroup.com/global/insights/citigps](http://www.citigroup.com/global/insights/citigps)



### **What Machines Can't Master**

*Human Skills to Thrive in the Age of AI*  
May 2024



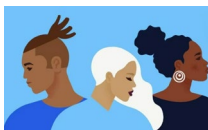
### **Disruptive Innovations X**

*Ten More Things to Stop and Think About*  
May 2024



### **Future of Healthcare**

*The Oxford Prescription*  
May 2024



### **Closing the Racial Wealth Gap**

*Restating the Economic Case*  
April 2024



### **Supply Chain Financing**

*Building Resilience*  
January 2024



### **Unlocking Climate and Development Finance**

*Creating Bankable Projects*  
December 2023



### **Food, Water, and Climate Change**

*Solutions for Tackling this Critical Nexus*  
November 2023



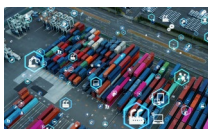
### **APEC Economic Outlook**

*Navigating Shifting Dynamics*  
November 2023



### **Unlocking Climate and Development Finance Preview**

*Creating Bankable Projects*  
November 2023



### **Supply Chains**

*Lessons Learned Helped by New Technology*  
November 2023



### **Food Security**

*Solutions for a Perfect Storm*  
November 2023



### **Treasury Leadership**

*Does It Matter?*  
November 2023



### **Philanthropy and the Global Economy v3.0**

*Perspectives on the Future of Giving*  
October 2023



### **Supply Chain Technology**

*Improving Efficiency, Resilience, and Transparency*  
October 2023



### **Technology at Work v7.0**

*The Third Phase of Globalization*  
October 2023



### **Africa**

*A New Growth Model*  
October 2023





If you are visually impaired and would like to speak to a Citi representative regarding the details of the graphics in this document, please call USA 1-888-800-5008 (TTY: 711), from outside the US +1-210-677-3788

## IMPORTANT DISCLOSURES

This communication has been prepared by Citigroup Global Markets Inc. and is distributed by or through its locally authorised affiliates (collectively, the "Firm") [E6GYB6412478]. This communication is not intended to constitute "research" as that term is defined by applicable regulations. Unless otherwise indicated, any reference to a research report or research recommendation is not intended to represent the whole report and is not in itself considered a recommendation or research report. The views expressed by each author herein are his/ her personal views and do not necessarily reflect the views of his/ her employer or any affiliated entity or the other authors, may differ from the views of other personnel at such entities, and may change without notice.

You should assume the following: The Firm may be the issuer of, or may trade as principal in, the financial instruments referred to in this communication or other related financial instruments. The author of this communication may have discussed the information contained herein with others within the Firm and the author and such other Firm personnel may have already acted on the basis of this information (including by trading for the Firm's proprietary accounts or communicating the information contained herein to other customers of the Firm). The Firm performs or seeks to perform investment banking and other services for the issuer of any such financial instruments. The Firm, the Firm's personnel (including those with whom the author may have consulted in the preparation of this communication), and other customers of the Firm may be long or short the financial instruments referred to herein, may have acquired such positions at prices and market conditions that are no longer available, and may have interests different or adverse to your interests.

This communication is provided for information and discussion purposes only. It does not constitute an offer or solicitation to purchase or sell any financial instruments. The information contained in this communication is based on generally available information and, although obtained from sources believed by the Firm to be reliable, its accuracy and completeness is not guaranteed. Certain personnel or business areas of the Firm may have access to or have acquired material non-public information that may have an impact (positive or negative) on the information contained herein, but that is not available to or known by the author of this communication.

The Firm shall have no liability to the user or to third parties, for the quality, accuracy, timeliness, continued availability or completeness of the data nor for any special, direct, indirect, incidental or consequential loss or damage which may be sustained because of the use of the information in this communication or otherwise arising in connection with this communication, provided that this exclusion of liability shall not exclude or limit any liability under any law or regulation applicable to the Firm that may not be excluded or restricted.

The provision of information is not based on your individual circumstances and should not be relied upon as an assessment of suitability for you of a particular product or transaction. Even if we possess information as to your objectives in relation to any transaction, series of transactions or trading strategy, this will not be deemed sufficient for any assessment of suitability for you of any transaction, series of transactions or trading strategy.

The Firm is not acting as your advisor, fiduciary or agent and is not managing your account. The information herein does not constitute investment advice and the Firm makes no recommendation as to the suitability of any of the products or transactions mentioned. Any trading or investment decisions you take are in reliance on your own analysis and judgment and/or that of your advisors and not in reliance on us. Therefore, prior to entering into any transaction, you should determine, without reliance on the Firm, the economic risks or merits, as well as the legal, tax and accounting characteristics and consequences of the transaction and that you are able to assume these risks.

Financial instruments denominated in a foreign currency are subject to exchange rate fluctuations, which may have an adverse effect on the price or value of an investment in such products. Investments in financial instruments carry significant risk, including the possible loss of the principal amount invested. Investors should obtain advice from their own tax, financial, legal and other advisors, and only make investment decisions on the basis of the investor's own objectives, experience and resources.

This communication is not intended to forecast or predict future events. Past performance is not a guarantee or indication of future results. Any prices provided herein (other than those that are identified as being historical) are indicative only and do not represent firm quotes as to either price or size. You should contact your local representative directly if you are interested in buying or selling any financial instrument, or pursuing any trading strategy, mentioned herein. No liability is accepted by the Firm for any loss (whether direct, indirect or consequential) that may arise from any use of the information contained herein or derived herefrom.

Although the Firm is affiliated with Citibank, N.A. (together with its subsidiaries and branches worldwide, "Citibank"), you should be aware that none of the other financial instruments mentioned in this communication (unless expressly stated otherwise) are (i) insured by the Federal Deposit Insurance Corporation or any other governmental authority, or (ii) deposits or other obligations of, or guaranteed by, Citibank or any other insured depository institution. This communication contains data compilations, writings and information that are proprietary to the Firm and protected under copyright and other intellectual property laws, and may not be redistributed or otherwise transmitted by you to any other person for any purpose.

**IRS Circular 230 Disclosure:** Citi and its employees are not in the business of providing, and do not provide, tax or legal advice to any taxpayer outside of Citi. Any statements in this Communication to tax matters were not intended or written to be used, and cannot be used or relied upon, by any taxpayer for the purpose of avoiding tax penalties. Any such taxpayer should seek advice based on the taxpayer's particular circumstances from an independent tax advisor.

© 2024 Citigroup Global Markets Inc. Member SIPC. All rights reserved. Citi and Citi and Arc Design are trademarks and service marks of Citigroup Inc. or its affiliates and are used and registered throughout the world.







