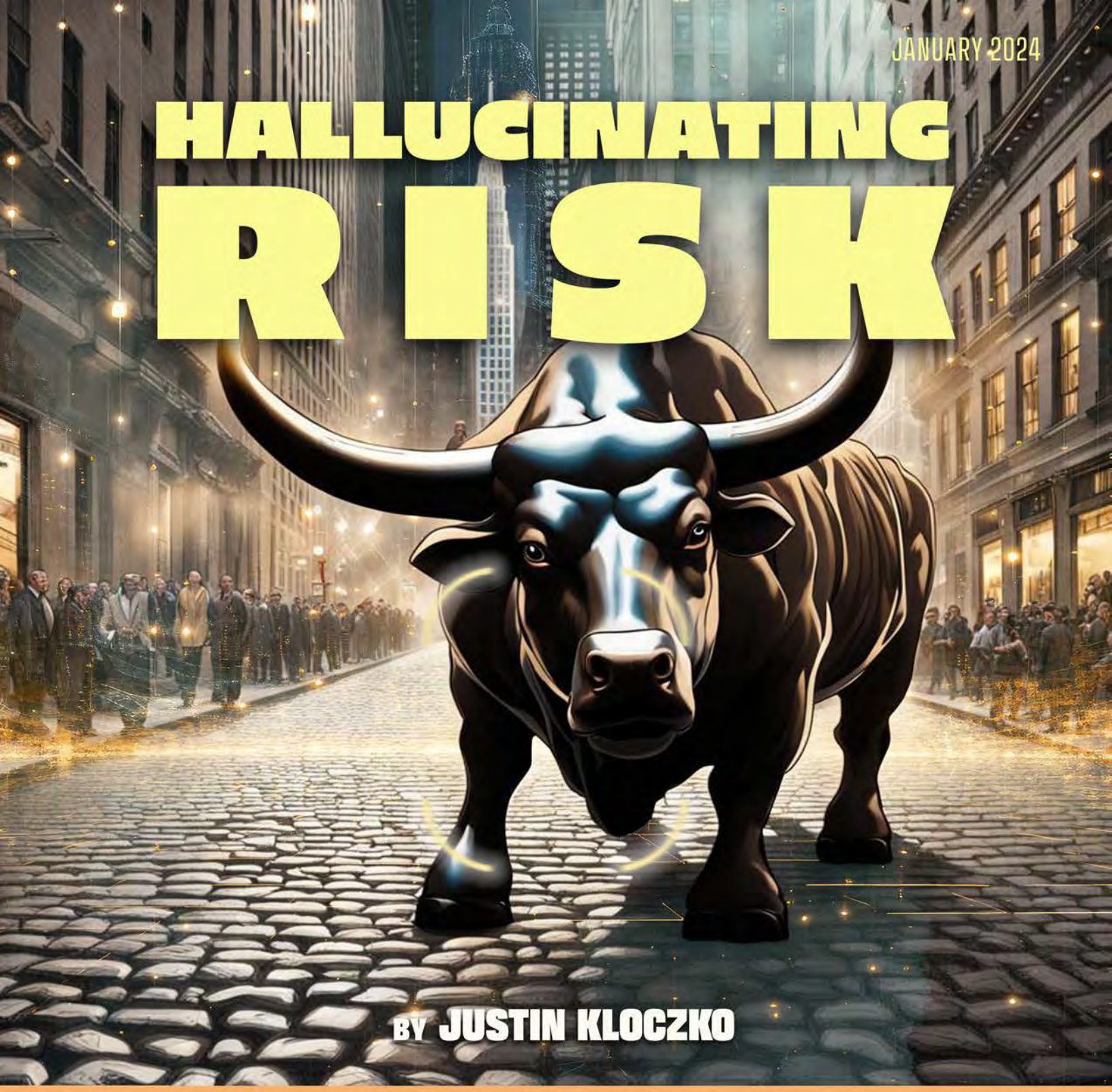
JANUARY 2024

HALLUCINATING RISK



BY JUSTIN KLOCZKO

AI IS RISKY AND MYSTERIOUS, BUT WALL STREET IS POURING BILLIONS INTO IT.

Consumer Watchdog

Table of Contents

Executive Summary 2

Introduction 7

Artificial Intelligence is a Mystery 9

How AI is Transforming Wall Street 12

Why Wall Street's Use of AI is Risky 15

What Can Be Done 20

Executive Summary

Major Wall Street investment banks are pouring billions of dollars into AI research, patents, and financing without adequate safeguards, according to an investigation by Consumer Watchdog.

Financial services spending on AI is larger than any other industry, exceeding even the tech industry.[1] A lot of attention is devoted to tech companies developing artificial intelligence, but Wall Street banks are just as interested in it. The unseen hand of AI is Wall Street financing.

According to a Consumer Watchdog analysis of patents filed with the United States Patent and Trademark Office (USPTO), Wall Street banks are seeking patents and trademarks for a range of uses, with JPMorgan Chase, Goldman Sachs and Morgan Stanley seeking AI for investment, analyzing securities, and predicting stock prices and portfolios. Every bank is seeking its own version of ChatGPT that will give financial advice to employees and customers. The lack of transparency with AI and its potential for bias means mysterious AI could push risky investments and loans, or hallucinate bad advice on managing debt without a consumer even knowing it was AI. Hallucinating occurs when an AI presents false information as fact.

"There is some systemic risk."

GERARD HOBERG,
USC PROFESSOR OF FINANCE

Without proper regulation, the next recession could be caused by AI, igniting in the mortgage or equity market due to a handful of powerful banks relying on a couple of algorithms.

"The more these companies use automation, and they start to use it on bigger markets, then there is some systemic risk," Gerard Hoberg, a University of Southern California professor of finance who studies artificial intelligence, told Consumer Watchdog.

Goldman Sachs is seeking to patent AI that will synthesize virtually all the data a trader would need to predict stock prices and another to predict a hedging portfolio.[2] JPMorgan has applied for a trademark called "Index GPT" that would dole out financial recommendations, and a patent that would match companies with investors[3]. Deutsche Bank is using artificial intelligence to scan the portfolios of clients, while ING Group is

---

[1] "Bridging the Gap Between BI and AI," *MIT Technology Review*, September 2022

[2] United States Patent and Trademark Office, (20230214923).pn, July 6, 2023.

[3] United States Patent and Trademark Office, (20230110458).pn., April 13, 2023.

screening for potential defaulters.[4] There is even AI to translate "Fedspeak," so banks can tell if statements by regulators such as the Federal Reserve are "dovish" or "hawkish." [5] And Wells Fargo is using generative AI to help decide what to disclose to regulators.

Consumer Watchdog has found that AI patent applications have exploded over the past five years. In 2020, the most current data provided by the the USPTO published nearly 80,000 AI-related patent applications, the most in its history.

Since 2017, the top five banks produced half of AI investment deals and 67 percent of research papers. Over 90 percent of AI patents in the banking sector were filed by five investment banks, according to Evident, which monitors AI implementation in the private sector.[6] Among these banks, JPMorgan Chase leads in AI research, Wells Fargo in investments, and Capital One in patents. JPMorgan and Wells Fargo are in the top five in all three categories. [7]

Powerful banks are seeking to incorporate Large Language Models (LLMs) such as ChatGPT despite ChatGPT being dangerous. ChatGPT has a "high risk of economic harm" due to a "tendency to hallucinate," and should not be used for financial advice. That's not just the assessment of outside experts, but of ChatGPT itself, according to parent company OpenAI's own usage policy and system card. There is little we know about generative AI. Even engineers, coders and people who study and work on AI often don't understand how AI works[8]. We don't know where the underlying training data comes from or how it is synthesized, raising serious concerns about privacy and safety. ChatGPT has scraped the Internet for over 300 billion words, and there are widespread

---

[4] "Wall Street Banks Are Using AI to Rewire the World of Finance," William Shaw and Aisha S. Gani, Bloomberg, May 31, 2023.

[5] "Sensing the Fed's Direction with the Help of AI," Morgan Stanley, April 23, 2023.

[6] "Evident Innovation Banks," Evident, July, 2023.

[7] "Evident Innovation Banks," Evident, July, 2023.

[8] "There's Only One Way to Control AI," Politico, Charles Jennings, August 20, 2023.

concerns that it has sucked up personal information along the way. But unlike previous models, OpenAI decided not to disclose the training data of GPT-4. Despite these concerns, JPMorgan Chase said this year that data and AI will be "critical" to the company's future success. It currently has more than 300 use cases for AI in production.[9]

The increased investment in AI is happening as banks are cutting jobs.[10] The banking industry is simply not prepared for AI compliance.

"AI could push risky investments, loans or mortgages"

"Nor is there evidence that banks are increasing investment in responsible AI talent—even as the environment on this matter turns serious and regulation looms," determined Evident.[11]

The major concerns about AI in the financial services industry are algorithmic complexity, a lack of transparency, and biased or false information. For example, the AI could push risky investments, loans or mortgages, and hallucinate bad advice on managing debt. Goldman's AI could be used to create and price out a new type of dangerous derivative.

"But it could be the case that the pricing algorithms missed something. Kind of like the financial crisis, those who valued mortgages back then thought they were diversified to housing price collapses. They were wrong," said Hoberg. "And that was the mistake the humans made. But with derivatives they get more complex. If banks end up taking on too much risk that their stock price starts to crash, and they can barely pay their bills because they took on a big loss, that would create another recession."

Current U.S. Securities and Exchange Commission Chair Gary Gensler warned about the dangers of AI in the financial services industry, arguing in a 2020 paper for MIT that AI breeds uniformity due to the reliance by multiple companies on one dataset.

"The 2008 crisis exposed the over-reliance of the financial sector on the three main credit agencies Standards & Poor's (S&P), Moody's, and Fitch to underwrite Collateral Debt Obligations (CDOs)," wrote Gensler. "These agencies used models with similar methodologies and evaluations of mortgage debt, all of which proved to be faulty."[12]

---

[9] JPMorgan Shareholder Letter, Jamie Dimon, 2023

[10] "Evident Talent Report," Evident, June 2023.

[11] "Evident Talent Report," Evident, June 2023.

[12] "Deep Learning and Financial Stability," Gary Gensler and Lily Bailey, MIT, Nov. 13, 2020.

As a result, 6 million Americans lost their homes. Complicated financial instruments enabled Wall Street to gamble with billions of dollars on the housing market. Then it all went bust.

With unprecedented reservoirs of data now at its disposal, Wall Street AI may even more precisely redline vulnerable communities and cause harm. Due to the complexity and opacity of financial algorithms, the public won't see the effects of a skewed algorithm until it's too late.

An AI that is not finely calibrated can also start to "drift," meaning it can deviate from its intended use and perpetuate bias.

"If an algorithm is not finely tuned, and it starts getting, say, 10 loan denials in a row for African American families of four, the algorithm will automatically sort of shift and be like, okay, Black families of four equals bad credit," said Richard Anthony, a tech advocate for Public Citizen.

Basically, AI can sway in any direction unless its inputs and outputs are constantly monitored.

"Because if you don't, it will just go and go and go and now it's racially discriminating because it got a few bad credit applications. And it decided to make that a part of itself," said Anthony.

The specter of unchecked generative AI doling out financial advice is scary, and needs to be regulated. The public needs to know when it is dealing with AI and the level of risk involved. Much like the European Union's soon-to-be approved AI Act, there needs to be a law in the United States that requires warning labels. Warning labels for AI giving financial advice would fall under "high risk."

In California, the California Privacy Protection Agency (CPPA) has proposed new draft automated decisionmaking rules under the country's landmark California Consumer Privacy Act (CCPA) that are a model for the nation.

Once implemented, the regulations will require:

- A pre-use notice alerting consumers that it uses automatic decisionmaking before a business collects personal information, including when the decision "results in access to, or the provision or denial of, financial or lending services."
- The right to know about the logic and to opt-out of the decisionmaking process.
- An opt out right even if a company "*uses computation as* ***whole or part*** *of a system* ***to make or execute a decision*** *or facilitate human decisionmaking."*

And for the first time, consumers will also be able to opt out of their personal data being used to train AI like ChatGPT under draft rules for algorithms and risk assessments.

A model for federal securities and AI regulation developed by Public Citizen would address the specific systemic risks and dangers of the AI in the financial space. Among the actions regulators should take:

- Ensure that AI investment systems are held to the same standard as human investment brokers/advisors.
- Mandate the external review of black box data of AI investment algorithms to ensure the AI serves the best interest of investors.
- Require AI investment system models to be regularly externally audited for groupthink pattern biases to prevent AI models from "herding."

Introduction

Few topics have dominated the news this year more than Artificial Intelligence. But it's been accelerating at a rapid pace, well before the media became heavily interested in it. Wall Street has been using algorithms for years now, but deep learning is taking them to another level.

Although AI is being evangelized as igniting massive progress in society, there are serious concerns about it doing the opposite. It can deepen fissures that already exist such as inequality, groupthink, and racism.

There is a revolving door between large tech companies and Wall Street. Wall Street is populated by Big Tech alumni, and vice versa. The chief operating officer of OpenAI, the company that developed ChatGPT, is a former analyst for JPMorgan.[13] Marco Argenti, Goldman Sach's Chief Information Officer, previously worked at Amazon Web Services and Nokia.

[13] Brad Lightcap, LinkedIn

Anthony, the tech advocate at Public Citizen who previously worked in the banking industry and the crypto currency sector, witnessed the speed of AI integration.

"Given my experience, the financial service sector is going to be the hardest hit in terms of AI integration because it's the furthest along," said Anthony. "Everyone's more worried about job losses. But for me, the bigger concern in the near future is going to be somebody getting denied a mortgage."

Now with OpenAI bringing ChatGPT to market, we have crossed the Rubicon. ChatGPT is like an overly eager intern who will always give you some sort of answer—quickly and confidently—even if it is the wrong one. It's up to society to set limits on the use and abuse of artificial intelligence or *caveat emptor*, aka buyer beware.

OpenAI agrees there should be warning labels on AI.

Artificial Intelligence is a Mystery

"But if deep learning predictions were explainable, they wouldn't be used in the first place. Instead, we would use linear models, table lookups, if-then statements, fixed rules and other, simpler approaches."

— Gary Gensler

ChatGPT-4, the most advanced LLM, often gives bad information, and it should not be used to make decisions in areas such as financial services, education, or healthcare. The National Institute of Standards and Technology, as well as the U.S. Office of Science and Technology Policy have warned about these AI biases and harmful discrimination.

Who else agrees? OpenAI, the company that created ChatGPT.

"GPT-4 has the tendency to 'hallucinate,'" i.e. "produce content that is nonsensical or untruthful in relation to certain sources," according to OpenAI's system card, a company document explaining how its LLM works.

The company said it is harmful when ChatGPT is used to determine credit, jobs, education or public benefits, according to the company's usage policy.

"Activity that has high risk of economic harm, including… Automated determinations of eligibility for credit, employment, educational institutions, or public assistance services," according to OpenAI. [14]

Yet despite these warnings, AI and Large Langue Models are being deployed without regulation in some of the most sensitive areas, such as financial services, human resources, and healthcare.

The protection of personal information is another issue. OpenAI concedes that ChatGPT could lead to "Activity that violates people's privacy, including…Unlawful collection or disclosure of personal identifiable information or educational, financial, or other protected records."

ChatGPT has scraped over 300 billion words online. It could just be giving the same advice as a random Reddit poster whose information it assimilated, and we would never

[14] OpenAI Privacy Policy.

know it because we do not know how it weighs information. In fact, we know nothing about the training data on which ChatGPT was built.

"In a departure from its previous releases, the company is giving away nothing about how GPT-4 was built—not the data, the amount of computing power, or the training techniques,"[15] wrote MIT Technology Review Senior Editor Will Douglas Heaven.

Florenta Teodoridis, an economist at the USC School of Business who studies AI, told Consumer Watchdog, "It is true that these algorithms, at least up until this point, are a bit of a black box, in the sense that not even the most advanced computer scientists understand what exactly happens inside for a prediction to come out at the other end. So they cannot be reverse engineered, which makes it harder because we don't know what we are missing along the way."

This is a problem, as science relies on robust transparency and reasoning.

"Without knowing how these systems are built, there is no reproducibility," said Dr. Kate Crawford, the former director of research at the AI Now Institute at New York University. "You can't test or develop mitigations, predict harms, or understand when and where they should not be deployed or trusted. The tools are black boxed."[16]

In order to build good policy around LLMs, transparency is needed, according to a 2023 Stanford University report on AI transparency.

"Black-box decisionmaking remains a challenge for policymakers, researchers, company executives, and the public seeking to understand why an AI model is generating a particular output."[17]

---

[15] "GPT-4 is bigger and better than ChatGPT—but OpenAI won't say why," Will Douglas Heaven, *MIT Technology Review,* March 14, 2023.

[16] https://x.com/katecrawford/status/1638524013432516610?s=20

[17] "Improving Transparency in AI Language Models," Rishi Bommasani, Daniel Zhang, Tony Lee, Percy Liang, Stanford University, February 2023.

Teodoridis said the AI being developed by banks are good at predicting past patterns, but not so much future ones. She analogized it to earthquakes.

"We have plenty of data about earthquakes, and yet not enough to predict when the earthquake is going to happen. Nobody knows when the next earthquake is going to happen, and that's because it's not enough to observe a pattern in the past that's gonna tell us with some certainty when the next earthquake is gonna happen, the magnitude of it and so on," said Teodoridis. "Prediction works to the extent that we have good data."

Writing in 2020, Gensler said in an MIT paper about AI and finance, "We conclude that deep learning is likely to increase systemic risks. " In October 2023, Gensler said it was "nearly unavoidable" that a financial crisis due to AI could happen within a decade[18].

"I do think we will in the future have a financial crisis…[and] in the after action reports people will say 'Aha! There was either one data aggregator or one model…we've relied on,'" said Gensler.

[18] "Gary Gensler urges regulators to tame AI risks to financial stability," Stefania Palma and Patrick Jenkins, *Financial Times*, October 14, 2023.

How AI is Transforming Wall Street

"AI and the raw materials that feeds it, data, will be critical to our company's success—the importance of implementing new technologies simply cannot be overstated."

— Jamie Dimon, CEO, JPMorgan Chase

Despite flaws and secrecy surrounding Large Language Models, the most powerful banks in the world are moving fast to deploy them. Customers already interface with bank chatbots on a regular basis, getting basic accounting assistance and other help from the robots. The top 10 largest banks in America use chatbots.[19] But judging from the types of patents Wall Street is applying for, banks are envisioning a future where AI will increasingly run the show. These generative AI chatbots will be the new financial advisors, both internally for employees and externally for customers. Almost half of all open jobs at major banks in 2023 are for AI-related roles.[20]

Wall Street's hunger for data is feeding the AI boom, as spending on financial data hit a record high of $37 billion in 2022.[21] Wall Street banks are leading the way in AI investment. Between 2017 and 2023, the top five banks made half of AI investment deals and produced 67 percent of research papers. By 2027, bank spending on AI will hit $97 billion, the most across industries.[22] Generative AI is expected to significantly boast Wall Street's value, between $200 billion and $340 billion, according to the corporate consulting firm McKinsey.[23]

Among banks in North America, JPMorgan Chase leads in the AI space. JPMorgan leads in AI research, spending $12 billion on technology in 2022, a near 10 percent increase from 2020.[24] The bank is also bankrolling a rival to ChatGPT named Cohere.

"We already have more than 300 AI use cases in production today for risk, prospecting, marketing, customer experience and fraud prevention, and AI runs throughout our

---

[19] Chatbots in Consumer Finance, Consumer Financial Protection Bureau, June 6, 2023.

[20] "Evident Talent Report," Evident, June 2023.

[21] "AI's Reverberations Across Finance," International Monetary Fund, December 2023.

[22] "AI's Reverberations Across Finance," International Monetary Fund, December 2023.

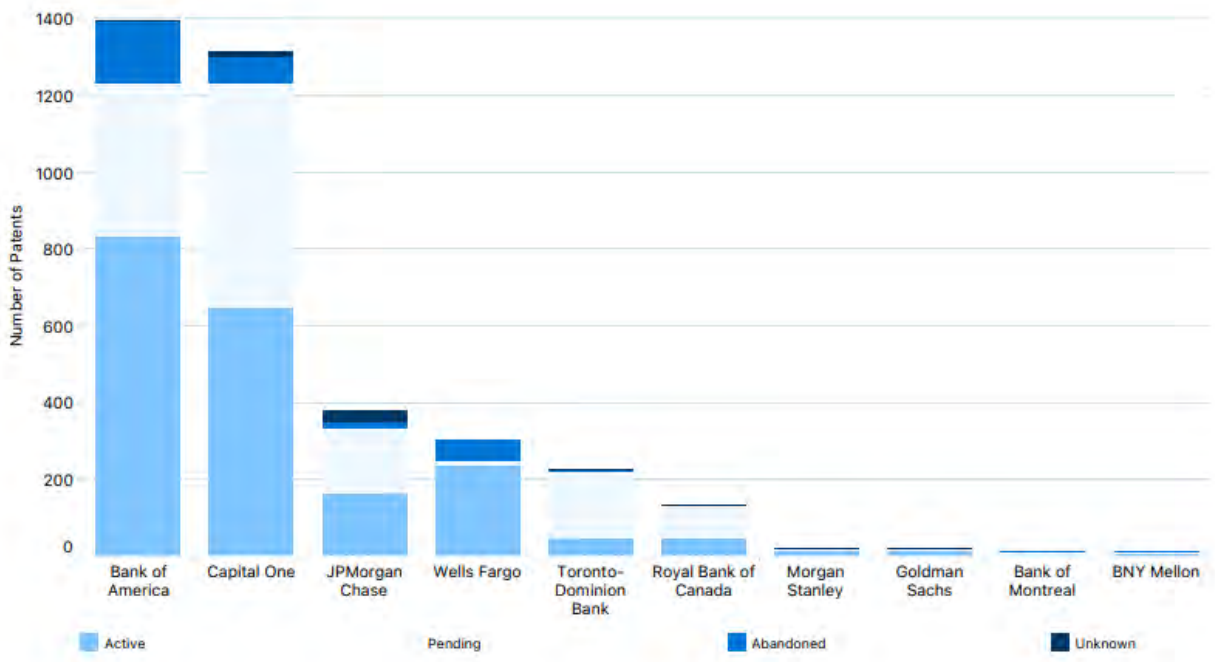
[23] "Capturing the Full Value of Generative AI in Banking," McKinsey, December 5, 2023.

[24] This $12 Billion Tech Investment Could Disrupt Banking, JPMorgan Chase

payments processing and money movement systems across the globe," said JPMorgan CEO Jamie Dimon in a letter to shareholders this year.

Most AI patents filed by banks are related to trading, followed by payments and user experience. Over 90 percent of AI patents were filed by five investment banks.[25] Bank of America leads in patents, with 1,396 filed between 2010 and 2021.[26]

Number of AI Patents Filed by Bank, 2010-2021



For example, JPMorgan filed a trademark this year called IndexGPT, consisting of "cloud computing software using artificial intelligence" for "analyzing and selecting securities tailored to customer needs," according to the patent application.[27]

The patent office also published an AI application this year from JPMorgan for the purposes of linking investors with companies. The AI dissects a company's size and how much money it has and compares it to an investor's "funding style." A rating is then produced informing you if the investment is a good fit.[28]

---

[25] "Evident Innovation Banks," Evident, July, 2023.

[26] "Evident Innovation Banks," Evident, July, 2023.

[27] United States Patent and Trademark Office, May 11, 2023.

[28] United States Patent and Trademark Office, (20230110458).pn., April 13, 2023.

Morgan Stanley is also rolling out an OpenAI chatbot to aid its army of over 15,000 financial advisors. The bank is doing this despite previously saying chatbots spit out "wrong" answers.[29] The chatbot will have access to company data to give insight into "companies, sectors, asset classes, capital markets, and regions around the world," the company announced this year.[30]

Wells Fargo is using generative AI to decide what its clients should or should not to disclose to regulators.[31] And this summer, the USPTO published a couple notable patents held by Goldman Sachs. One is an interface that will synthesize virtually all the data a trader would need to better predict stock prices[32].

"Successfully predicting how the price of a given asset will evolve over time can be highly profitable," according to the bank's patent application.

Another Goldman Sachs patent is for AI to predict a hedging portfolio. The AI "identifies a recommended hedging portfolio that track a target asset and provides one or metrics indicating a predicted performance of the hedging portfolio relative to the target asset," according to the patent application. [33]

The Top 10 largest banks in the U.S. already use basic chatbot technology:

1. JPMorgan Chase
2. Bank of America
3. Citibank
4. Wells Fargo
5. US Bank
6. Truist
7. PNC
8. Goldman Sachs
9. Capital One
10. TD

Deutsche Bank uses artificial intelligence to scan the portfolios of clients, while ING Group is screening for potential defaulters.[34]

There is even AI to translate "Fedspeak," so banks can tell if statements by regulators like the Federal Reserve are "dovish" or "hawkish."[35] The AI is calibrated to detect the earliest pangs of a policy shift such as whether there will be a rise or fall in interest rates.

---

[29] ChatGPT will keep 'hallucinating' wrong answers for years to come and won't take off until it's on your cellphone, Morgan Stanley says, Zahra Tayeb, *Markets Insider,* Feb. 23, 2023.

[30] Morgan Stanley Wealth Management Announces Key Milestone in Innovation Journey with OpenAI, Morgan Stanley, March 14, 2023.

[31] "Wall Street Banks are Using AI to Rewire the World of Finance," William Shaw and Aisha S. Gani, *Wealth Management*, June 2, 2023.

[32] https://ppubs.uspto.gov/pubwebapp/ (20230214923).pn.

[33] United States Patent and Trademark Office, (20230267544).pn.

[34] "Wall Street Banks Are Using AI to Rewire the World of Finance," William Shaw and Aisha S. Gani, *Bloomberg*, May 31, 2023.

[35] "Sensing the Fed's Direction with the Help of AI," Morgan Stanley, April 23, 2023.

Why Wall Street's Use of AI is Risky

"The more these companies use automation, and they start to use it on bigger markets, then there is some systemic risk," said Gerard Hoberg, a University of Southern California professor of finance who studies artificial intelligence.

AI could push risky investments, loans or mortgages, and give bad advice on managing debt. Or it could create and price out a new type of dangerous derivative.

"But it could be the case that the pricing algorithms missed something. Kind of like the financial crisis, those who valued mortgages back then thought they were diversified to housing price collapses. They were wrong," said Hoberg. "And that was the mistake the humans made. But with derivatives they get more complex. If banks end up taking on too

much risk such their stock price starts to crash, and they can barely pay their bills because they took on a big loss, that would create another recession."

Algorithms can also be complex, mysterious, and biased or incorrect.

Bias

For example, the same amount of people in a geographically affluent area seek mortgages compared to people in a lower-income area. Generative AI could start denying mortgage applicants from the low-income area based on geolocation, even if an applicant from the low-income area has a better credit score than someone from the affluent area. That's because the AI will start unfairly associating a lower income address with bad credit, according to the group. AI can redline people based on location, and they won't even know why.

In fact, this has already happened. An investigation by *The Markup* found that algorithmic home mortgage lenders gave out 40%-80% more loans to White people than people of color in scenarios where both groups had *similar financial characteristics*[36]. In addition, high-earning Black applicants with less debt were denied loans more than high-earning White applicants with more debt.

"If that data is incomplete, then the algorithm is going to make decisions only based on that incomplete data. And it's going to perpetuate whatever is missing into the prediction," said Teodoridis.

Herding

Herding is when financial group-think develops because individuals or financial companies rely on the same dataset or model. Herding contributed to the Savings and Loan Crisis of the 1980s, the dot-com bubble, and the 2007 Quant Meltdown. Because of herding and the interconnectedness of financial markets, the damage inflicted by a flawed algorithm can spread like a contagion. And due to the complexity, speed and opacity of financial algorithms, the public won't know the problem until it's too late.

As Gensler noted, uniformity in finance can happen when many actors in the financial sector rely on a centralized dataset or model. In 2007, the "Quant Meltdown," many hedge funds recorded hundreds of millions of dollars in losses over a period of a few weeks. Financial analysts have since attributed the problem to herding and the

[36] "The Secret Bias Hidden in Mortgage-Approval Algorithms," *The Markup*, Emmanuel Martinez and Lauren Kircher, August 25, 2021.

overreliance of a dataset.[37] The issues run deep. A reliance on just a few data models in the hands of only a handful of very powerful banks could trigger a lending or equity crisis.

Herding was also at play in the lead up to the 2008 financial crash. An unregulated derivative market is what contributed to the financial crisis of 2008 and the ensuing Great Recession. The expansion of Credit Default Swaps (CDS) led more investors to speculate on bundled mortgages such as Collateralized Debt Obligations (CDO).[38] When the housing market crashed, 6 million Americans lost their homes.

"The 2008 crisis exposed the over-reliance of the financial sector on the three main credit agencies Standards & Poor's (S&P), Moody's, and Fitch to underwrite collateral debt obligations (CDOs)," wrote Gensler in a 2020 paper for MIT. "These agencies used models with similar methodologies and evaluations of mortgage debt, all of which proved to be faulty."[39]

The housing market wasn't the only sector to be affected. School districts, state and local governments, as well as hospitals, bought these derivatives. An entire county government filed for bankruptcy, in part as a result of being sharked into interest rate swaps[40].

"If an algorithm is not finely tuned, and it starts getting, say, 10 loan denials in a row for African American families of four, the algorithm will automatically sort of shift and be like, okay, Black families of four equals bad credit."

RICHARD ANTHONY
PUBLIC CITIZEN

Herding was also responsible for a "flash crash" that occurred in 2010 when an algorithm used by a mutual fund was triggered into a stock buying frenzy. Hundreds of billions of dollars were wiped off the share prices of Proctor & Gamble and General Electric. The market recovered quickly. But the risk is stocks being sold for pennies when they are worth a dollar, or vice versa.

---

[37] "What Happened To The Quants in August 2007?" MIT, Amir E. Khandani and Andrew W. Lo, September 20, 2007.

[38] "Moving to Opportunity? The Geography of the Foreclosure Crisis and the Importance of Location," Chistos Makridis and Michael Ohlrogge, *Journal of Economic Geography*, March 21, 2021.

[39] "Deep Learning and Financial Stability," Gary Gensler and Lily Bailey, MIT, Nov. 13, 2020.

[40] "In Alabama, a County That Fell off the Financial Cliff," Mary Williams Walsh, *New York Times*, Feb. 18, 2012.

"These robots are trained typically on past data. It's quite likely that all the robots are seeing the same past data," said Hoberg, the USC professor. "So they're likely to behave in similar ways. And so an unexpected scenario comes, and they're all going to behave in a similar way. And it might just be the way that we don't want them to behave."

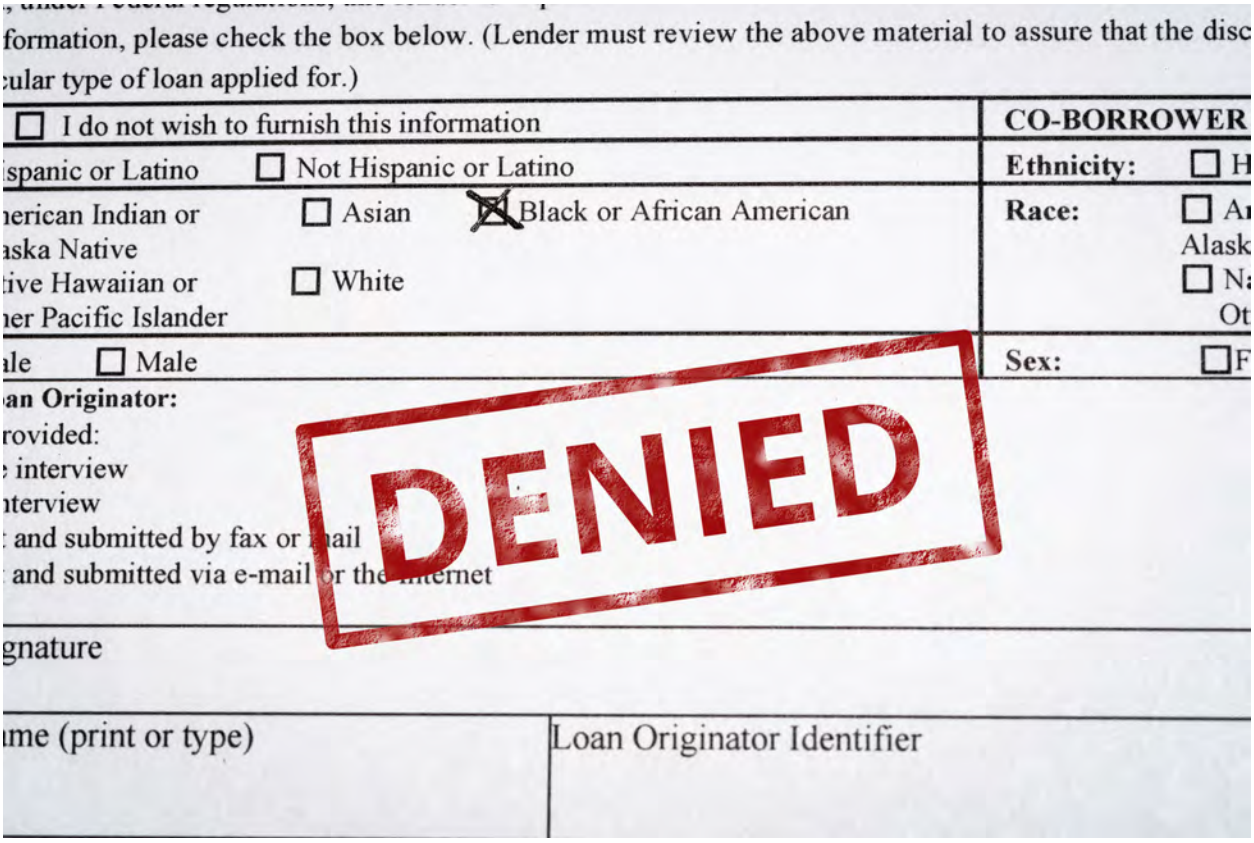
Knight Capital went bankrupt in less than an hour after a faulty algorithm in their trading software bought billions in stocks.

In 2012, a 17-year-old company that managed billions in daily trades went out business in under an hour after its algorithm glitched. An error in the trading software sent the algorithm into a $7 billion stock buying frenzy.[41] More so, the software passed its safety tests, but when it was deployed it activated its testing software instead of its trading program.

An AI that is not finely calibrated can also start to "drift," according to Anthony, the tech advocate at Public Citizen.

[41] "Case Study 4: The $440 Million Software Error at Knight Capital," Henricodolfing, June 5, 2019.

"If an algorithm is not finely tuned, and it starts getting, say, 10 loan denials in a row for African American families of four, the algorithm will automatically sort of shift and be like, okay, Black families of four equals bad credit," said Anthony.



Basically, the AI can be swayed in any direction unless its inputs and outputs are constantly monitored.

"Because if you don't, it will just go and go and go and now it's racially discriminating because it got a few bad credit applications. And it decided to make that a part of itself," said Anthony.

With unprecedented reservoirs of data now at its disposal, and due to the complexity and opacity of financial algorithms, AI may even more precisely redline vulnerable communities and the public won't see the effects of a skewed algorithm until it's too late.

What Can Be Done

As we enter a world where anyone can be defamed in the blink of an eye and the line between reality and make believe is blurred more than ever, it is imperative that people know when they are interfacing with AI.

California has a model for AI regulation and the agency tasked with enforcing the law, the California Privacy Protection Agency (CPPA), is working on finalizing draft regulations surrounding automated decision making technology (ADMT) and risk assessments. The regulations offer guidance for the California Consumer Privacy Act (CCPA), the state law which gives Californians more control and knowledge about how their personal information is used by private companies. Once final, the regulations will give consumers the right to be informed when a decision involving their personal information is made by ADMT, how the decision was reached, and the opportunity to opt out of its use. Under the draft regulations, residents will be able to opt out of ADMT if the decision legally or significantly affects them, if they are profiled as a worker, applicant, or student, or in a public place.

Consumers will be protected with regard to the access or denial of:

- Financial or lending services
- Housing
- Insurance
- Education enrollment or opportunity
- Criminal justice
- Employment or independent contracting opportunities or compensation
- Healthcare services
- Essential goods or services

California's ADMT rules are also meant to complement draft rules governing risk assessments.[42] Those rules require businesses to submit detailed risk assessments for a wide range of personal information, including:

- Selling, sharing, and processing of personal information.
- Using automated decision making technology to make important decisions about housing, lending or healthcare, for example.
- Processing information of people under the age of 16.
- Using technology to monitor employees, freelancers, or job applicants.
- Using technology to monitor people's behavior or movements in public.
- Processing personal information of people to train AI.

[42] CPPA Draft Regulations

The documentation required is also extensive. It includes:

- Categories of personal information processed.
- The context and reason for processing.
- What a consumer could expect regarding the purpose of the processing.
- Retention period and technology used for processing.
- Benefits and negative impacts of processing.
- Any safeguards a business plans to put in place.

Importantly, companies will also have to submit risk assessments when they process personal information to train a large language model. The draft regs define "training" as "teaching artificial intelligence or automated decisionmaking technology to generate a desired output. Training includes determining or improving the parameters of the artificial intelligence or automated decisionmaking technology to achieve the desired output." Under the draft regs, Californians will be able to opt out of the use of their personal information that generative AI such as GPT run on.

There are no federal laws on the books right now governing personal data and how financial institutions may use artificial intelligence. The closest we have are federal laws that apply to certain sectors, such as the Gramm-Leach-Bliley Act (GLBA) for banking data.

Jolina Cuaresma, an attorney and Berkeley Law lecturer who has represented clients before the SEC and the Consumer Financial Protection Bureau, said GLBA has limited preemption of state laws such as the CCPA.

"GLBA generally governs the sharing of information between affiliates," said Cuaresma. "The only teeth GLBA has, in my opinion, is it forced companies to provide privacy notices to consumers. It's 'what am I going to do with your data in this one context, with affiliates.' And that's it."

For example, if a bank is using AI to determine a loan using a Californian's personal information, that person under the CCPA should be able to know more about the decision, personal information used, and opt out of it, according to Cuaresma.

GLBA is silent on automated decisions or deleting personal information. California can provide greater protections as long as it doesn't conflict with GLBA's narrow law surrounding the sharing of financial information.

FEDERAL FRAMEWORK

Beyond California's broad rights, specific protections are needed in the financial space to protect individuals and guard against systemic risk. The SEC is in the middle of a rulemaking to eliminate conflicts of interest when it comes to such technology.
The following are some federal regulations proposed by Public Citizen surrounding AI and securities that are critical to reining in the risks of AI.

Fiduciary Duty:

- Regulators should ensure that AI investment systems are held to the same standard as human investment brokers/advisors.
- Fiduciary duty should explicitly cover AI-driven financial advice.
- Mandate the external review of black box data of AI investment algorithms to ensure the AI serves the best interest of investors and empower the SEC to prevent/punish the development/use of AI financial algorithms that don't serve the fiduciary interest
- Mandate investment brokerages establish and publicly share their policies for identifying and eliminating conflicts of interest that may arise with their use of AI, such as an AI investment advisor upselling the value of a stock to a consumer's portfolio because a stock purchase would benefit the brokerage that provides the AI system

Systemic Concerns:

- Mandate investment brokerages/advisors create a working group that shares intel and best practices on preventing/stopping AI driven dis/misinformation in the stock market.
- Require AI investment system models to be regularly externally audited for groupthink pattern biases to prevent AI models from "herding" investment analysts/institutional investors to making decisions that would critically damage our financial system
- Establish high standards for data inputs/usage and cybersecurity in AI investment systems and require regular independent audits of these systems.

President Biden issued an executive order on AI to much fanfare this year, however its enforceable requirements are limited. The most concrete provision of the executive order is that companies developing AI that could pose a risk to national security, national public health or economic security are required to conduct risk assessments and share that information with the government. The EO gives the Commerce Secretary the power to require companies to submit information on training, development, and production of models, as well as the measures taken to protect them. It also directs federal agencies to develop various guidelines and conduct studies, including the creation of standards for watermarking when a photo, audio, video or text is created by AI.

However, while the White House can direct federal agencies' use of AI, its powers are far more limited in the private sector, or state and local government, without action by Congress. For example, companies such as ClearviewAI, which builds a reservoir of face recognition data, contract with state and local police forces who don't have to comply with the Biden Administration's order. And the executive order is not binding and can be undone with the stroke of a pen when the next Administration comes to power.

The European Union is on the verge of approving the AI Act, which requires warning labels for AI in different risk categories:

> ***Minimal:*** Video games.
>
> ***Limited:*** Chatbots in the retail or the food service industry (Amazon, Starbucks).
>
> ***High:*** Chatbots giving advice on investments, debt management, credit, surgery, evidence in a criminal case.
>
> ***Unacceptable:*** Anything dangerous such as social scoring, toys encouraging dangerous behavior in children.

The United States needs a similar law, one that would inform consumers when they are interfacing with AI in financial services, and be told of the proper risk that comes with it.

OpenAI agrees there should be warning labels on AI:

"Consumer-facing uses of our models in medical, financial, and legal industries; in news generation or news summarization; and where else warranted, must provide a disclaimer to users informing them that AI is being used and of its potential limitations," according to OpenAI's usage policy.

Additionally, penalties must be attached for not complying with the law. The AI Act's penalties include $37 million for noncompliance on AI practices, $27 million for noncompliance with certain high-risk AI systems, and $17 million when misleading or wrong information is provided in response to a request.

America is falling behind on AI regulation. It cannot be slow to act like it was regarding social media. Both sound regulation and innovation are possible.