



STATEMENT BY  
VIJAY KARUNAMURTHY  
FIELD CHIEF TECHNOLOGY OFFICER  
SCALE AI

BEFORE THE  
HOUSE COMMITTEE ON FINANCIAL SERVICES

ENTITLED  
“AI INNOVATION EXPLORED:  
INSIGHTS INTO AI APPLICATIONS IN FINANCIAL SERVICES AND HOUSING”

JULY 23, 2024

Chairman McHenry, Ranking Member Waters, and members of the House Financial Services Committee, thank you for the opportunity to be here today. It is my honor to discuss with you the exciting role that Artificial Intelligence (AI) is playing in the Financial Services and Housing sectors, how industry is already adopting the technology, and the trends we are seeing about its future adoption.

## **Introduction**

My name is Vijay Karunamuthy, and I am the Field Chief Technology Officer at Scale AI (Scale). As a technologist, I have devoted my career to building next-generation technologies that would make people's lives better. Early on, this meant harnessing the power of machine learning to improve media and video applications at companies like YouTube and Apple. Now I am proud to work for Scale where we are helping the world's leading technology companies build, apply, and evaluate AI systems.

Since 2016, Scale's mission has been to build the data foundry for AI. We have been on the forefront of nearly every major AI innovation and worked across the industry to power high-quality AI systems.

Initially, Scale worked with leading automotive companies such as General Motors and Toyota to build high-quality computer vision datasets for autonomous vehicles and other autonomy programs. But then, as AI evolved, we adapted our existing Data Engine to power what is now referred to as Generative AI or Large Language Models (LLMs) by working with companies like OpenAI, Meta, and Anthropic.

More recently, Scale has helped some of the world's most prominent companies, such as Global Atlantic Financial Group, Morgan Stanley, and Howard Hughes, apply AI and harness the power of their proprietary data. We are also proud to work with the U.S. government to bring best-in-class commercial AI technology to the Department of Defense and other agencies.<sup>1</sup>

Scale worked alongside the leading tech companies and stakeholders to conduct early-stage research on techniques to build higher quality and more human-like

---

<sup>1</sup> See, <https://www.businesswire.com/news/home/20240521674964/en/Scale-AI-Raises-1-Billion-Series-F-to-Push-The-Frontier-of-AI-Data>

responses. Today, these techniques are better known as Supervised Fine Tuning (SFT) and Reinforcement Learning from Human Feedback (RLHF).

Once AI systems were able to routinely produce higher quality outputs, Scale recognized that understanding model safety and performance was critical, and we developed methodologies to test and evaluate it so its benefits and limitations could be understood.

## **AI Use in Financial Services & Housing**

AI has been around in various forms for decades, and the financial services and housing industries were two of the earliest sectors to recognize its potential. Since the Turing Test in 1950,<sup>2</sup> AI technology continued to evolve at a steady pace with sectors, such as manufacturing, beginning to adopt it in the late 1970s.

In 1982, the financial services industry began to embrace AI, and today nearly every sector of the economy utilizes AI for some type of service, ranging from customer service to assistance with portfolio management.<sup>3</sup> Due to its early adoption, the financial services industry is relatively further along in the path of applying cutting-edge AI technologies.

The housing industry has also begun to embrace AI for functions such as real estate management and speeding up the permitting process and other administrative tasks.<sup>4</sup>

These combined decades of experience using AI have made the financial services and housing sectors knowledgeable about the impact that AI can have on an industry. And importantly, their early adoption allowed federal and state regulators to incorporate many aspects of AI use in existing regulations for those sectors.<sup>5</sup>

---

<sup>2</sup> See, [https://en.wikipedia.org/wiki/Turing\\_test#:~:text=The%20test%20was%20introduced%20by,at%20the%20University%20of%20Manchester](https://en.wikipedia.org/wiki/Turing_test#:~:text=The%20test%20was%20introduced%20by,at%20the%20University%20of%20Manchester).

<sup>3</sup> See, [https://www.turing.ac.uk/sites/default/files/2019-04/artificial\\_intelligence\\_in\\_finance\\_-\\_turing\\_report\\_0.pdf](https://www.turing.ac.uk/sites/default/files/2019-04/artificial_intelligence_in_finance_-_turing_report_0.pdf)

<sup>4</sup> See <https://bipartisanpolicy.org/explainer/ai-housing-industry-innovations/#:~:text=AI%20presents%20a%20host%20of,credit%2C%20the%20potential%20is%20significant>.

<sup>5</sup> See, <https://www.thomsonreuters.com/en-us/posts/corporates/ai-compliance-financial-services/#:~:text=Firms%20must%20also%20perform%20a,AI%20Enabled%20Regtech>

## Emerging Trends in AI Adoption

So where do these early adoption sectors, like financial services and housing, go from here?

From the perspective of making AI adoptable at scale, perhaps the greatest step change occurred in November 2022 when OpenAI released ChatGPT.<sup>6</sup> The broad release of LLMs have democratized AI by allowing nearly any person with an internet connection able to utilize AI on a daily basis. Beyond the broad commercial appeal, LLMs also have shown tremendous enterprise benefits.

Due to this promise, the financial service industry is projected to spend \$97 billion on AI by 2027,<sup>7</sup> and nearly every AI company is working to utilize their model to make it more powerful for financial service specific use cases.

Based on our work with the leading tech companies developing AI models and enterprises deploying AI, Scale has unique insights into the emerging trends across the financial and housing sectors. I want to share three of those key trends in AI deployment today:

***Foundational AI elements are critical:*** At a high level, AI is only as good as the data that it is trained on, and high quality data produces high quality outputs from the AI system. The leading AI labs have recognized they need a runaway data advantage to stay ahead of their competition, and for that reason are, collectively, spending billions of dollars a year on data alone.

As the leading companies in the world deploy AI systems, they have to make sure that they have data strategies and budgets in place to turn raw data into AI-ready data, data that is curated, transformed, and annotated in order to be leveraged by an AI system.

***Fine-tuned, proprietary data is a game changer:*** AI training has two main steps. During pre-training, the first step, the open internet is scrubbed and the information is used to teach the model to understand the nuances of language

---

<sup>6</sup> See, <https://en.wikipedia.org/wiki/ChatGPT#:~:text=ChatGPT%20is%20a%20chatbot%20and,launched%20on%20November%2030%2C%202022.>

<sup>7</sup> See, <https://www.statista.com/statistics/1446037/financial-sector-estimated-ai-spending-forecast/#:~:text=The%20financial%20sector's%20spending%20on,billion%20U.S.%20dollars%20in%202027.>

and generate coherent responses.<sup>8</sup> It is a necessary first step in training AI and results in models that can respond to prompts at the level of a high schooler.<sup>9</sup>

Second, pre-trained models can undergo post training by applying techniques such as SFT and RLHF. Applying these enables models to go from responding at a high school level to the level of masters or even PhD students with expertise in particular topics.

However, companies need not only masters & PhD-level capabilities for AI systems, but also industry specific, expert-level capabilities that come from years of institutional knowledge within the company itself.

That is where proprietary data comes in. By training models with a company's own data, this gives the AI system the institutional knowledge that enables enterprises to derive value from generative AI for their particular applications.

As an example, it has been reported that JP Morgan currently has roughly 150 Petabytes of proprietary data,<sup>10</sup> or 150 times the entirety of Chat GPT-4's initial training data.<sup>11</sup> Harnessing the power of the data it currently houses for AI would create incredible insights for that bank to better serve its customers because this data is specifically relevant to its business.

While AI must be deployed responsibly, this capability could enable a newer employee to access institutional knowledge and data from the past 20 years and, as appropriate for their responsibilities, more easily solve problems of today.

If companies want to truly harness the power of AI for their own use, they must not only use best-in-class commercial models, but also leverage the power of their own data to build the highest quality AI systems.

***Deploying AI is top of mind and must be done safely and responsibly:*** Industries across the board are looking for the next AI innovation to give them a competitive

---

<sup>8</sup> See, <https://www.chatgptguide.ai/2024/03/01/what-is-pre-training-llms-explained/>

<sup>9</sup> See, <https://www.chatgptguide.ai/2024/03/01/what-is-pre-training-llms-explained/>

<sup>10</sup> See, <https://www.projectpro.io/article/how-jpmorgan-uses-hadoop-to-leverage-big-data-analytics/142>

<sup>11</sup> See, <https://dataloop.ai/blog/keeping-up-with-the-ai-wave-auto-gpt-sets-new-standards/>

advantage, but even in this market race, we are seeing companies express a commitment to deploying AI responsibly. That is essential.

Regular test and evaluation is necessary to produce and deploy the highest quality AI systems and ensure that they are safe to deploy.

Safe AI deployment of AI systems takes many forms and can include test and evaluation, developing safe operating environments, leveraging traditional cyber security standards, setting up virtual private clouds and networks to prevent data leaks, and continuous monitoring. From my viewpoint, Test and Evaluation, in particular, is the best way to ensure that AI is safe to deploy because it entails robust testing of the AI system to better understand the potential vulnerabilities and comprehensive evaluations to produce a holistic picture of the AI system. For this reason, the U.S. government has long recognized the importance of developing the standards, frameworks, and methodologies to underpin test and evaluation, as well as public displays to show its benefit.

To further demonstrate this point, in 2023, Scale was selected to build the test and evaluation platform at the DEFCON AI exercise in which 10 of the leading LLMs were tested and evaluated by the world's leading hackers. During this exercise, more than 2,220 participants red-teamed the leading AI models to expose vulnerabilities, demonstrating the critical component that test and evaluation serves in enabling safe, responsible, and trustworthy AI.<sup>12</sup>

Scale has also launched our Safety, Evaluation and Alignment Lab (SEAL), which is continuing to move the discussion of AI safety forward by releasing open source benchmark tests on topics like biosecurity and chemical weapons<sup>13</sup>, and developing proprietary evaluations of the leading AI models, which we release publicly, to ensure that the public is made aware of the strengths and weaknesses of the models themselves prior to using them.<sup>14</sup>

Beyond our work in the commercial sector with companies like OpenAI, we are working with the U.S. Department of Defense's Chief Digital and AI Office to help them develop a framework and begin early testing of the Department's Generative AI systems.<sup>15</sup>

---

<sup>12</sup> See, <https://aivillage.org/defcon%2031/generative-recap/>

<sup>13</sup> See, <https://scale.com/blog/measuring-mitigating-risk-wmdp>

<sup>14</sup> See, <https://scale.com/leaderboard>

<sup>15</sup> See, <https://scale.com/blog/scale-partners-with-cdao-to-test-and-evaluate-llms>

Test and evaluation is also critical for companies looking to deploy AI systems. While the leading developers regularly test and evaluate their AI models, this is not enough. As companies look to leverage fine-tuned models, they need to consistently perform their own tests to ensure that the AI remains safe to deploy. Once they deem it safe to deploy, companies must also continue to monitor the performance of the AI system and periodically retest it to ensure that it has not drifted.

For this reason, Scale has also developed best-in-class products, like our Scale Generative AI Platform (SGP), that allows customers to generate unprecedented insights into the performance and safety of their custom-developed GenAI.

Companies should also regularly red-team their AI systems to identify unanticipated vulnerabilities and whether the systems are producing unintended responses.

### **Recommendations to Ensure Innovation Remains Our North Star**

AI is only able to deliver on its nearly limitless potential if it is deployed safely and responsibly within a regulatory framework that enables innovation. Scale firmly believes that the north star is a framework that is flexible enough to enable innovation while establishing sufficient safety guardrails.

The financial services and housing sectors have both utilized AI for decades, which has allowed regulators to better understand the use cases and incorporate them within existing regulations. For that reason, new regulations may not be necessary, but a thorough and comprehensive gap analysis should be conducted to confirm whether or not existing regulations apply to AI.

Our recommendation is that this gap analysis should focus on ensuring that regulations govern the use of the technology and if gaps exist, we must fill them with risk-based, sector-specific regulations. Risk-based regulations are critical to ensure that the anticipated risk of the AI aligns with the rigor of the testing. This is especially important for the financial services and housing sectors. Both sectors have true high-risk use cases which, if the AI is not thoroughly tested and evaluated, have the potential to produce economic harm to Americans. Therefore, there must rightly be a higher bar to test against.

Industry can not do this alone. One key area is the adoption of AI safety metrics. However, industry needs the government's guidance to establish metrics to test against to determine whether or not the AI is safe to deploy. In practice, this may be regulators simply stating that AI is held to the same standards as humans, but the government agency with jurisdiction over the sector must establish these metrics to ensure they are industry-specific and consistent with existing regulations.

Lastly, government and industry should work together to ensure that the workforce of the future is ready for the technology. AI has nearly limitless potential, but if we do not train the workforce, it will never be reached. This is incumbent on engagement from all layers of the innovation ecosystem—the government needs to ensure that users understand how to properly use and regulate AI, academia must begin to better incorporate AI skill sets into their teachings, and industry must commit to workforce training and development.

## **Conclusion**

AI is the most promising technological innovation of our time, but it must be deployed in a safe, responsible, and thoughtful manner. This is our mission and if done properly, it will set up the United States economy to reap the benefits.

Thank you again for the opportunity to be here today, and I look forward to your questions.