Request for Information and Comment on Financial Institutions’ Use of Artificial Intelligence, including Machine Learning

Scale AI, Inc.

Table of Contents

Scale AI Overview

Company Mission & Background

Scale AI's mission is to accelerate the development of Artificial Intelligence for our customers. We work with the federal government as well as companies like Square, Paypal, Brex, Flexport, OpenAI, General Motors, Toyota, Lyft, Samsung, DoorDash, and many others to build the data-centric, end-to-end infrastructure layer for managing the entire machine learning lifecycle. Scale delivers solutions across annotation, data management, automated data extraction, model evaluation, and synthetic data generation as part of our complete enterprise AI product.

Scale's headquarters are in San Francisco, CA and currently employs approximately 350 people.

Financials & Investors

Scale is a privately owned company that has raised $600+ million in funding and valued at over $7.3 billion. Our most recent Series E funding round was co-led by Dragoneer, Greenoaks Capital, and Tiger Global with continued investments from Founders Fund, Index Ventures, Coatue, and Y Combinator.

Leadership Team

Scale is proud of the team we've built & our leadership. We believe that building a team of exceptional engineers and operations specialists results in happy customers. Our team of experts is able to quickly adapt the platform and consistently deliver on customer throughput and quality goals.

- **Alex Wang (CEO) –** Leads Scale. Alex is himself a Machine Learning Engineer. He started Scale after working on many ML projects and realizing that the biggest bottleneck in building ML applications is high quality training data. Before starting Scale, he worked at Quora and Addepar as an Engineering Manager and studied Machine Learning at MIT.

- **Brad Porter (CTO) –** Leads Scale's engineering and product teams. Brad is a seasoned executive joining us from Amazon where he spent 13 years spearheading the company's efforts in Robotics and was named Distinguished Engineer, an honor given to a handful of people within the company for their work.

- **Mark Valentine (Head of Federal)** – Leads Scale's federal business unit. Mark has extensive experience working in the US Federal Government space and previously spent the last 6 years at Microsoft working across National Security projects spanning cloud computing, AI, mixed reality, and cybersecurity. Prior to Microsoft, Mark served for over 25-years in the US Air Force as a fighter pilot, commander and staff officer on the Air Staff and the Joint Staff.
- **Melisa Tokmak (GM of Document AI)** – Leads Scale Document across product and engineering. As an executive at Scale, Melisa has built multiple new lines of business including the government team. Prior to Scale, she led New Product Monetization and managed $4B+ in revenue at Facebook and graduated from Stanford University with a degree in Computer Science where she worked on artificial intelligence.

The Future of Financial Services

Over the past few years, increasing digitization efforts in the financial services industry is enabling new product experiences for end users and improving the efficiency of internal workflows. Financial institutions are leveraging various technologies, particularly Artificial Intelligence (AI), to reimagine existing processes with automation and directly impact business metrics including cutting costs. The industry is continuing to see rapid adoption of AI across a variety of use cases and core to accelerating this growth is data.

Investments in cloud computing, data management systems, and API-based integrations have made it easier for financial institutions to collect diverse data sources, and the natural next step is unlocking value from that data to improve financial products. For example, using AI to derive insights from alternative datasets, enable high-quality automated document processing, and analyze product data from customers can be directly used to build a faster, more personalized credit decisioning application for loans while also speeding up backend workflows.

Financial institutions, including a number of government agencies, are sitting on a treasure trove of data across a variety of sources such as consumer/business banking, loans, investments, payments, taxes, etc. However, the challenge is moving beyond small-scale experiments and instead operationalizing production-level AI systems. Beyond requiring machine learning resources and engineering support to build out these systems, one of the biggest struggles that Scale has heard first-hand is managing the data needed to build applications. AI is garbage-in, garbage-out – better data ultimately leads to better AI models. Even the best algorithms won't perform well without high-quality data.

Building the Platform to Enable AI for Financial Services

In order to deploy the best models, it's critical to leverage the right infrastructure that can prepare high-quality data to power more accurate and robust AI applications. Scale works closely with some of the largest financial institutions to accomplish this goal by helping them build an AI-ready Knowledge Graph to centralize data across silos in order to stack on the AI algorithms that leverage the data for different use cases. There are four key aspects for constructing a Knowledge Graph with high-quality data for financial services AI applications (Scale provides the tooling built in-house for each of these areas):

1. Annotation – Unstructured, unlabeled data is common in financial services given the amount of information that can be collected, but that raw data is useless for training AI models without clean, high-quality annotations. Labeling data is typically an expensive, manual process and it's difficult to ensure quality. Scale takes a technology-first approach combining both humans and machine learning to speed up operational efficiency and automatically ensure accurate results, enabling us to deliver guaranteed high-quality SLAs at low costs.

2. Extraction – Data extraction is a common task, especially across the variety of documents that financial institutions process. These documents contain critical information that power workflows like credit decisioning, application processing, etc. However, data extraction involves labor intensive workflows to manually enter information into systems. Existing automation solutions like Optical Character Recognition (OCR) or "intelligent" processing only transcribe text without identifying the most important fields you need and rely on rules-based templates that don't work if a document format changes. Scale can help with its Adaptive AI technology, where we deliver a tailored machine learning solution to each of our customers for higher accuracy data extraction without needing any upfront or continuous maintenance effort.

3. Linking – Identifying the relationships across different sources of data is critical for building an AI-ready Knowledge Graph for downstream applications. These links serve as a foundation for modeling detailed patterns in datasets as well as enable explainability since you can directly visualize the connections and decisions made. Algorithmic approaches can process large volumes of data but oftentimes miss out on nuanced relationships that require human-insight, resulting in lower quality. Scale's approach combines custom built AI algorithms that automatically link data using contextual information with expert review to ensure high accuracy and coverage.

4. Debugging – Once the knowledge graph is built, it's important to continuously monitor your data and understand where your model isn't performing well. Best-in class debugging tools enable teams to quickly identify a model's biggest problems, whether

it's across biased data or inaccurate predictions and investigate those issues individually. Scale provides the platform to interpret large-scale knowledge graphs and stacks on extensive validation metrics and visualization tooling to monitor integrity for use in AI applications.

Applications of AI in Financial Services

The consolidated knowledge graph of data unlocks the potential for a number of impactful AI applications in financial services. We highlight a few areas where Scale has delivered AI models that drove business impact for customers:

Document + Text Understanding AI – Documents and unstructured text are one of the most common data types in financial services. Typically, processing this information involves humans manually reviewing each document or text snippet individually, which is expensive and prone to error. AI models can automate the extraction of key fields of interest from this unstructured data to speed up downstream workflows. Scale's approach is unique because our solution adapts to variations in content and layout without sacrificing on quality or requiring any manual configuration. Document and text understanding models are oftentimes the first building block for other AI applications in financial services (described below).

Risk Management AI – The key in preventing fraud is minimizing false positives (results in poor customer experience) and false negatives (leads to regulation/compliance risks) in order to process large volumes of data at high quality. One of the challenges in fraud detection is the variability in the data that is collected across text, audio, documents, etc. as well as the constant race to prevent new bad actor attempts. AI solutions are valuable for automating data extraction to reduce errors and identify information that might be indicative of fraudulent activity without needing extensive manual review. In addition, the knowledge graph serves as a useful tool for detecting anomalies by clustering related incidents and estimating risk levels.

Insurance AI – Generating accurate quotes, modeling risk, and processing claims efficiently are all core parts of the insurance stack where AI can improve workflows. Alternative data sources, such as satellite imagery, social media, etc, can be ingested via the knowledge graph and used to develop AI models that price risk more accurately while streamlining processes. In addition, rather than relying on individual auditors to review claims submissions, AI can more quickly extract and surface key information from reports to identify fraud risk and respond to the claims more quickly.

Personalized Assistant AI – Current "personalized" assistants rely on large call centers of people triaging questions or simple rules-based chat bots. However, the former approach is operationally challenging, costly, and can significantly reduce customer satisfaction. The latter approach, while more automated, lacks the ability to learn over time and adapt to unstructured contextual information related to finances. Instead, AI models, built on latest advancements in natural language understanding and deep learning, are more sophisticated in analyzing deeper interactions in an organization's knowledge graph to answer more complex financial questions (ex. "What factors should I consider when applying for loan A vs loan B?").

Lending AI – Similar to insurance, accurately underwriting loans can directly impact both top and bottom line for a business. Rather than relying on a team of people who manually collect information to make a loan decision, AI can enable end-to-end automation to accelerate human workflows. By parsing the raw documents (ex. 1004 appraisals, paystubs, title reports, etc.) customers send when applying for a loan, extracting data with machine learning models, and structuring that data via a knowledge graph, the underwriting team can make more informed decisions at a much faster pace. In addition, AI models that parse alternative datasets beyond typical documents submitted in a loan application (bank transactions, investment reports, etc.) can be used to improve credit decisions.

Key Areas for Assessing AI-Readiness

Explainability

Question 1: How do financial institutions identify and manage risks relating to AI explainability? What barriers or challenges for explainability exist for developing, adopting, and managing AI?

AI explainability is crucial for financial services because many of the decisions made directly impact the inflow and outflow of money for an institution. Models built on decision trees or regressions are beneficial because outputs are directly impacted by a defined set of parameters. However, these models are simplistic and lack the robustness to handle more complicated data types or tasks. Deep learning models are becoming a standard for AI applications because they are able to more effectively learn nuanced features in datasets and perform at higher accuracy across a variety of financial services workflows. However, one of the key drawbacks of deep learning is that it's more difficult to directly explain the rationale behind predictions, a challenge when deploying AI models for financial services use cases where understanding the decision making process is important (ex. loan applications).

Explainability is crucial throughout the end-to-end AI lifecycle. During the development phase, it's important for financial institutions to get a strong understanding of the input data, including features like data distribution, variance, and formats. When both deploying and maintaining those models, having systems in place to monitor predictions ensures visibility into how the AI is making decisions.

Question 2: How do financial institutions use post-hoc methods to assist in evaluating conceptual soundness? How common are these methods? Are there limitations of these methods (whether to explain an AI approach's overall operation or to explain a specific prediction or categorization)? If so, please provide details on such limitations.

Post-hoc explainability methods are applied after a black-box AI model to more effectively interpret the predictions that the model is making for a given input. Some of the most common methods for post-hoc interpretation include visualization of results, model simplification, feature ranking, model feature visualization, etc. These approaches are all typically used for better understanding the process in which a model makes an output by analyzing changes in predictions across a large input dataset. However, the key limitation of these approaches is that it's difficult to get an exact understanding of exactly how the model arrived at a specific output. Since many AI models are now deep learning-based, there is no "rules-based" decision tree or pre-defined parameters that make interpretability straightforward. Despite this, post-hoc explainability methods still provide insights into model predictions that can be useful for determining how a model performs across inputs. In addition, even simpler approaches such as having a platform to debug model outputs and quickly query through results can also be a useful tool for interpreting model performance on a dataset.

Question 3: For which uses of AI is lack of explainability more of a challenge? Please describe those challenges in detail. How do financial institutions account for and manage the varied challenges and risks posed by different uses?

In general, explainability becomes a challenge when using AI for complex data types such as images, text, documents, etc. where there's no defined set of rules to cover the possible set of output predictions. This problem is top-of-mind especially for high priority use cases that can impact business metrics, such as loan decisioning, risk mitigation, etc. For example, when trying to parse and extract key information from long, unstructured documents, the AI model needs to not only parse the layout of the document (i.e. the location of fields) but also understand the content of the text on a page. These documents might have millions of subtle variations, making it difficult to encode with a simple set of "if-then" rules. However, even though AI models are powerful in understanding these complex relationships, explainability is more challenging because it's hard to pinpoint exactly what the model is doing under-the-hood. Whenever an AI system ingests data, it iteratively constructs sets of numeric values (called vector embeddings) that models various features of the data. These vectors encode information about the inputs in a familiar format that the AI can understand, even if it looks meaningless to humans.

To mitigate issues with explainability, it's important for financial institutions to invest resources into tools that can continuously monitor data inputs and debug model outputs. Identifying the types of inputs to a given model can give insights into the output predictions, but variations in real-world data can directly impact the model in production. This occurs especially when the inputs are previously unseen from training the AI. Moreover, the ability to backtest and quickly identify specific instances where the model is underperforming enables rapid debugging to better understand a model's decision making process. Platforms to efficiently query large datasets and identify issues with model outputs, such as Scale's Nucleus product, can help financial institutions mitigate the risks associated with AI explainability through a simple interface to better understand predictions.

Risks from Broader or More Intensive Data Processing and Usage

Question 4: How do financial institutions using AI manage risks related to data quality and data processing? How, if at all, have control processes or automated data quality routines changed to address the data quality needs of AI? How does risk management for alternative data compare to that of traditional data? Are there any barriers or challenges that data quality and data processing pose for developing, adopting, and managing AI?

As described above, data quality is critical for building the best AI models. The performance of AI solutions can suffer without high quality training data because these models are dependent on the quality of the input data. Financial institutions typically have standard pre-processing workflows to get raw data into a format ingestible by machine learning algorithms. However, oftentimes the raw data itself is unlabeled or unstructured, meaning it can't be used to build AI models or automate workflows, respectively. For example, 1004 forms, paystubs, bank statements, etc. are useful sources of information for loan applications, but to actually improve processing efficiency, it's necessary to extract the information from these unstructured documents.

One of the biggest challenges for ensuring data quality for AI applications is building the infrastructure to manage and get that high-quality data. Typically, financial institutions have in-house teams that curate and label data for other teams to build AI applications. However, from Scale's experience working with some of the largest organizations in the space, a common challenge is building and maintaining the platform for labeling data but also ensuring quality control checks. The reason why this is challenging is because it typically requires rigorous human review for each item in the dataset, which can quickly become expensive and operationally complex to manage.

Question 5: Are there specific uses of AI for which alternative data are particularly effective?

There are a number of use cases where alternative data is effective, including:

- Loan/credit underwriting – On personal lines, leveraging datasets from social media and unstructured documents can provide additional information about an individual's ability to repay a loan. Importantly, AI solutions can automate those workflows so alternative data is analyzed without requiring extensive human review. Similarly, social media, documents, and even satellite imagery can all be useful in underwriting business loans.
- Chatbots/personalized recommendations – raw text or audio from conversations or interactions with customers can be tapped to build more robust personalized assistements and AI models that can recommend the best products for users.
- Fraud detection – social media analysis and information from business documents can be useful across a variety of financial business units, such as loans, tax repayments, and onboarding.
- Investment decisions – this is a common use case where alternative data from websites, satellite imagery, and raw text can help provide differentiated insights into certain investment opportunities.
- Insurance underwriting – images, video data, 3D scans, and raw text from claims documents are all sources or useful information for insurance applications when constructing a policy as well as arbitrating a filed claim.

Across all of these areas, Scale has helped financial institutions leverage their alternative datasets to drive business impact. We found that these unconventional data types, managed within an AI-ready knowledge graph, can enable a variety of downstream AI applications that perform at higher accuracy and improve customer experience.

Overfitting

Question 6: How do financial institutions manage AI risks relating to overfitting? What barriers or challenges, if any, does overfitting pose for developing, adopting, and managing AI? How do financial institutions develop their AI so that it will adapt to new and potentially different populations (outside of the test and training data)?

Overfitting is a key challenge that all financial institutions should keep in mind when developing AI because it can impact the quality of solutions on real-world data. When transitioning from experimentation to a production deployment, it's important that the training data for the model is representative of what might actually be seen. By doing so, the model can perform at similar expected levels of quality on real-world data as it would on the training/test set during development. To build AI in this way, it all comes down to getting high-quality data that is labeled with detailed information. As mentioned above, this is typically a challenge that financial institutions face because ensuring quality and efficiency in acquiring enough labeled data for AI models is a hard task. Moreover, it's important that any production AI system is able to identify specific segments of data where the model is underperforming and quickly retrain and deploy to improve performance on those examples.

Thus, to address the challenge with overfitting, financial institutions need to invest in systems across two areas: first, getting high-quality, representative data for training AI models to prevent overfitting and second, building the infrastructure to continuously monitor model performance.

The pains associated with infrastructure development is an example of why financial institutions use Scale. We integrate AI models directly with our technology-enhanced human-in-the-loop platform to get high-quality training data to tailor solutions for our customers while continuously improving based on patterns where the model is underperforming. Financial institutions simply plug into Scale's platform without worrying about any setup.

Cybersecurity Risk

Question 7: Have financial institutions identified particular cybersecurity risks or experienced such incidents with respect to AI? If so, what practices are financial institutions using to manage cybersecurity risks related to AI? Please describe any barriers or challenges to the use of AI associated with cybersecurity risks. Are there specific information security or cybersecurity controls that can be applied to AI?

Cybersecurity threats from vulnerabilities in AI models are an important factor when deploying models for financial services applications. These data engineering exploits are most notable for images, for example when an AI model misclassifies objects based on subtle, unnoticeable changes in an input image. However, these attacks can still occur for other data types as well, so building systems that can detect and handle edge cases is essential. To mitigate these risks, it's important to train AI models by curating a representative training dataset and debugging any issues with results before moving into production. The challenge, as Scale has seen while working with financial institutions, is building out the necessary tooling to enable end-to-end data and model debugging workflows. However, when the investment in those systems are made, financial institutions gain a number of improvements, including being better prepared for cybersecurity threats. In addition, when an AI model is in production, organizations should always monitor the incoming inputs and resulting predictions to see if there are any anomalies. The same infrastructure used during the training phase can also be applied for detecting these unusual variations and flagging any previously unseen data types as a potential attack on the AI model.

Dynamic Updating

Question 8: How do financial institutions manage AI risks relating to dynamic updating? Describe any barriers or challenges that may impede the use of AI that involve dynamic updating. How do financial institutions gain an understanding of whether AI approaches producing different outputs over time based on the same inputs are operating as intended?

Ensuring AI models dynamically update is critical for adapting to real-world variations in the data distribution, which may not have been seen while training the AI. However, one of the biggest challenges for implementing dynamic updating is the machine learning operations (ML ops) infrastructure investment needed to ensure there's an end-to-end feedback loop for identifying problems, retraining, and redeploying the model. In addition, while most financial institutions are shifting to the cloud, the ones that want on-prem solutions find it even more difficult because it requires building all of the necessary tooling in house rather than leveraging the scalability and flexibility of secure cloud environments.

There are two key issues to address with dynamic updating: first, changes in the inputs that might be unexpected for the model and severely impact output results, and second, after updating the model, the output results for input data might look different. To gain more visibility into AI model performance over time, it's important for financial institutions to have systems that continuously monitor for these potential issues and flag anomalies. This would involve having a production ML ops infrastructure to identify when a model is underperforming, the tooling to visualize those instances, and an integrated platform to quickly get labeled data and retrain the model on certain examples.

AI Use by Community Institutions

Question 9: Do community institutions face particular challenges in developing, adopting, and using AI? If so, please provide detail about such challenges. What practices are employed to address those impediments or challenges?

Community institutions do face challenges in adopting AI because they generally have less investment and resources to commit to new systems. The organizations typically don't have large machine learning or engineering teams and that domain experience is important for building out new AI applications. In addition, there is less incentive to adopt AI given these institutions are much smaller in size and don't employ as many people compared to larger financial institutions. Resources are limited and most of the investment goes into other initiatives rather than building in-house AI teams. As AI becomes more common across financial services, one way in which these community institutions can build up knowledge is through hosted seminars that train folks within those organizations about AI basics. In addition, these community institutions can still benefit from AI by working with third-party machine learning companies, like Scale, who provide the end-to-end infrastructure for getting high quality data, managing data in a centralized platform, and deploying AI models tailored for specific business needs. In this case, community institutions don't need significant AI expertise and instead can work closely with a vendor like Scale to scope out specific projects and develop AI solutions to address those needs.

Oversight of Third Parties

Question 10: Please describe any particular challenges or impediments financial institutions face in using AI developed or provided by third parties and a description of how financial institutions manage the associated risks. Please provide detail on any challenges or impediments. How do those challenges or impediments vary by financial institution size and complexity?

One challenge associated with using third-party AI solutions is that many deliver generic, off-the-shelf models that aren't tailored to a financial institutions' exact use case. Even for a standard task, such as document processing for onboarding workflows, each financial institution deals with specific types of documents from customers and might care about only extracting a certain set of fields. The off-the-shelf solutions are built to work across many institutions as a plug and play solution, but don't fine-tune their models based on exact customer requirements. Ultimately, this means the model may not achieve the highest quality results and is unable to dynamically update based on variations in data. At Scale, we build a knowledge graph based on the data silos for financial institutions and deliver tailored AI solutions that are customized for all of the customers we work with. We found that this approach significantly outperforms traditional third party solutions, while ensuring a more robust system for the long-term.

Another challenge is ensuring that the third parties are using secure systems to process and handle sensitive data. Working with trusted vendors with experience in financial services and government work is important for ensuring data security and regulatory compliance. Financial institutions should be conscious of doing risk assessments prior to starting an engagement and be sure that the vendor they are working with has a strong reputation and credibility in the space. For example, at Scale, security and compliance is top priority since we work with a number of financial institutions and have since taken the necessary steps on our end, including extensive security audits and diligencing, to ensure our systems are protected. Moreover, we are a government contractor and FedRAMP certified to handle sensitive data.

In general, financial institutions should conduct necessary due diligence when selecting a third party vendor, prioritizing the experience and reputation of the company as well as the potential to streamline infrastructure into a single easy-to-use platform. Throughout any engagement, these financial institutions should work closely with a vendor to address any specific regulatory concerns and define key success metrics for a project.

Scale's Past Performance

Federal Prime Contracts

Customer	**Summary**	**Associated Data Types**	**Period of Performance**	**Value**
U.S. Army Research Lab <u>Contract #:</u> W911QX-20-C-0051 <u>POC:</u> Michael Lander michael.d.lander2.civ@mail.mil	Research and develop a series of scientific approaches that will support a foundational methodology designed to facilitate U.S. Department of Defense efforts to experiment, develop and iterate on high-quality annotated datasets for artificial intelligence and machine learning.	Synthetic aperture radar Full-motion video Satellite imagery Acoustic	30 Sep. 20 – 6 Jun. 24	$90.1M
U.S. Air Force Research Lab	Accelerating intelligence, surveillance, and reconnaissance (ISR)	Full-motion video	15 Mar. 21 – 14 Jun. 22	$850K

Contract #: FA864921P0899 POC: Jenah Hernek jenah.hernek@us.af.mil	processing, exploitation and dissemination (PED) for the Advanced Battle Management System (ABMS) SmartONE program.	Satellite imagery		

Federal Subcontracts

Customer	**Summary**	**Associated Data Types**	**Period of Performance**	**Value**
ECS Federal POC: Dan Keller Daniel.Keller@ecstech.com	Data labeling and annotation support to specific U.S. Department of Defense program efforts.	Synthetic aperture radar Full-motion video	1 Apr. 20 – 30 Sep. 21	$5.6M
Palantir POC: Shannon Clark sclark@palantir.com	Named-entity recognition taskbank for the U.S. Army Distributed Common Ground System (DCGS) Capability Drop 2 (CD2).	Document	10 Nov. 20 – 10 Dec. 20	$250K
Hivemapper POC: Neil Sobin neil@hivemap	Provide a USSOCOM component with processing and multi-sensor fusion from non-traditional ISR platforms.	Full-motion video (aerial and ground)	25 May 21 – 25 June 21	$100K

per.com				

Select Commercial Contracts

Customer	**Summary**	**Associated Data Types**	**Period of Performance**	**Value**
Flexport	Data extraction for processing complex, unstructured logistics documents with a machine learning model fine-tuned to Flexports' document use cases.	Documents	20 Jun. 20 - Present	*Unable to disclose*
Brex	Instant in-app invoice processing for automating business onboarding and payment workflows for Brex's new premium products.	Documents	10 Feb. 21 - Present	*Unable to disclose*
Large financial services institution #1	Development of AI knowledge graph with added machine learning solutions for transaction classification and fraud prevention	Documents, Text	25 Feb. 21 - Present	*Unable to disclose*

Conclusion

Scale AI is pleased to submit our response for the *Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, including Machine Learning*. We're excited to continue additional conversations to further discuss how we can be a technical partner to support new AI opportunities and strategic initiatives.