

How the power of data and AI can transform financial services

Participants at the Databricks Financial Services Symposium shared their experience tackling today's financial services challenges

Databricks, inventor and pioneer of the data lakehouse, recently hosted a symposium on the use of data and AI to transform the financial services industry. Participants at the in-person and live-streamed Financial Services Symposium explored how their organizations can unleash the potential of their data. Featured speakers from global brands in banking, insurance, capital markets and fintech shared innovative data and AI use cases with solutions like open-source collaboration and the Databrick's Lakehouse for Financial Services, a single platform that brings together all data and analytics workloads to power transformative innovations in modern financial services institutions.

How FSIs benefit as they move along the Data Maturity Curve

Financial institutions are under more pressure than ever to compete with technology and fintech companies for the lead in data and analytics. In an opening keynote, Junta Nakai, Databricks Global Head of Financial Services and Sustainability Leader at Databricks, spoke of the need for the financial services institutions (FSI) to progress along the data maturity curve, moving from things like clean data and ad hoc queries to prescriptive analytics and automated decision making. "The farther [along the curve] you go, the more collaboration, the more AI, the more cloud," he said. Anup Segu, Senior Software Engineer at YipitData, said during the Executive Panel Discussion that YipitData's prospects have also recently progressed along the curve. "I think this trend is only going to continue where we have our investor clients being much more data-savvy," he said. Participants discussed the competition FSIs are facing from fintech companies, who are, in large part, far along the maturity curve.

"The farther [along the data maturity curve] you go, the more collaboration, the more AI, the more cloud."

Junta Nakai, Global Head, Financial Services and Sustainability, Databricks

Every company is a technology company

Financial services is the fastest-growing vertical at Databricks, Nakai said, due in part to urgency around becoming more data and AI-driven to compete. "Every company, in our opinion, has to become a data and AI

company to survive,” he said, and that’s reflected in job postings. The number of job postings at one bank that mentioned the word “cloud” increased 378% from 2019 to 2021 (job postings that mentioned “machine learning” increased 2,713%).

FSIs struggle with data and AI

Participants discussed the challenges FSIs face on the road to data maturity. In his keynote, Nakai said the main pain points are:

- Data is trapped in silos across multiple business units
- Inability of legacy data architecture to scale in real time
- Data that is not easily governed or available
- Collaboration is hindered by highly complex workflows, disparate technologies and spreadsheet culture

Participants spoke of seeking to remove data engineers from the day-to-day processes, in favor of more automated decision making and analysis. Segu added, “Data teams need to be able to instrument their own insights and produce their own analyses without really having to interact with data engineers and [without] technical barriers for them to produce the insights that our clients care about.” Nicolas Doyen, Head of Product at TickSmith, said during the Executive Panel Discussion that the maturity curve often starts with data being internal and ends with it being fully externalized. “Once you’ve mastered your own data and you’ve engaged with partners about how this data is valuable to them and the different ways they can use it...you get new insights into how your customers want to engage with you as an organization, how they want to engage with the data, and you can even find new ways that you can use that data to solve new questions and new use cases that come up.”

Data Sharing and Data Democratization

Nakai spoke to the need for FSIs to use data and AI to “make better decisions on an automated basis.” Everything on the left of the maturity curve — clean data, reports, ad hoc queries, and data exploration — is backward-looking, and serves to assess what has happened in the past, Nakai said. That’s typically powered by a data warehouse, which consolidates “data in a structured, clean way,” so that question is easily answered. At the same time, “you need a data lake for all your unstructured data, your real-time data, for your use cases that involve a lot more machine learning and advanced analytics” Nakai said. However, the challenges that arise from managing both a data warehouse and a data lake, with vastly different personas engaging in each, are immense.

Enter Databrick’s Lakehouse for Financial Services. “Lakehouse is putting the best parts of a data warehouse and a data lake together on one platform,” Nakai said. “If you have a Lakehouse, you can drive everything from AI to BI from that single place.” It moves FSIs from having incomplete support for use cases to having all machine learning, SQL, BI, and streaming use cases; from incompatible security and governance models to one approach for all

“If you have a Lakehouse, you can drive everything from AI to BI from that single place.”

Junta Nakai, Global Head, Financial Services and Sustainability, Databricks

assets on all clouds; and from disjointed and duplicative data in silos to an open and reliable data platform to efficiently handle all data types, he said. Antoine Amend, Senior Technical Director, Financial Services, Databricks, said, "It's not just about AI, it's about culture change, it's about operating AI, it's about democratizing [data]."

Moving data out of silos

"The data lake can be really swampy," Ben Tallman, Chief Technology Officer at M Science, said on the panel, adding that the challenge inherent in data transformation is to "define the deep, clear areas of your data lake" and "keep people out of the swampy murky spots at the edge." Rizwaan (Riz) Sahib, CITO at Brookfield Renewable, spoke about the data transformation aimed at optimizing the firm's hydro assets, which removed the silos that had often prevented data insights. "The way our previous process was set up was, it was very manual," he said. "You had a lot of tribal knowledge, you had a lot of experienced people that everything was [only stored] in their heads, and we used Excel quite a bit." At that time, it took a day to run a schedule. He described the transformation process: "The first thing we did was we sat down with our teams, took everything that was in their heads, and started to define our requirements." By the end of the transformation, they had created an optimizer that, "within an hour, will take all the inputs, spit out a schedule, and submit it directly into our trade entry system. So now, our team is spending a lot less time building the schedule and a lot more time analyzing the results, and actually seeing how well we're performing, and starting to look at future ways to optimize our business." During the panel, Francis Wenzel, CEO & co-founder at TickSmith, also stressed the importance of taking data out of silos and making it work together. "We have tools like Databricks and others that make the data usable, put the data in one place, and make it accessible," he said. Yipit Data's Segu added, "For us, the lakehouse architecture is a philosophy of how to centralize your data assets and then unify the analytics around it...[to] answer key investment questions for the investor clients that we work with."

Sharing business logic

In a keynote, Stephen Goldbaum, Executive Director at Morgan Stanley, discussed the company's open-source project, named Morphir, designed to share business logic. Morphir arose out of frustration with rewriting business logic every time the technology changed. "We went about seeing if we could save the business logic in something that's not technology, that's independent from the technology, so that as the technology changes,

"We went about seeing if we could save the business logic in something that's not technology, that's independent from the technology, so that as the technology changes, the business logic is safe."

Stephen Goldbaum, Executive Director at Morgan Stanley

the business logic is safe," he said. "The goal of Morphir is not to dictate a single way of writing business logic. The goal is that when you do write your business logic, we want to be able to extract information out of it. By opening this up to the open-source community ...we were able to start collaborating with leaders in other aspects of technology."

Hyper-personalization

“Historically, financial services firms were product-centric, not customer-centric,” said Nakai. However, with the wide adoption of hyper-personalization in areas like streaming and shopping, customer expectations have changed. Now, FSIs are wise to take on hyper-personalization to be “more relevant to customers,” Nakai said.

It’s about the customer experience

Nakai discussed how the measure of a good bank used to be the number of branches, but today it is also the customer experience on the app. “A good example of that [is] Starbucks,” he said. “Why are people attracted to a company like Starbucks? Is it because they have the best possible coffee or is it because they have the best possible customer experience, the best possible accommodation, the best possible app? What makes Starbucks so successful today is not just about the [number of locations], it’s about their app that runs personalization engines on your transaction history, the weather, your geolocation, and propels me to go buy a pumpkin latte in October. That’s an example of how companies are starting to harness the foundations of capital and scale and augmenting that with the data.”

“What makes Starbucks so successful today is not just about the [number of locations], it’s about their app that runs personalization engines on your transaction history, the weather, your geolocation, and propels me to go buy a pumpkin latte in October.”

Junta Nakai, Global Head, Financial Services and Sustainability, Databricks

Personalization drives insights

Newer banks are focused on the customer “experience that attracts the hearts and minds of a younger generation,” Nakai said in his keynote. “Any transaction generates new data points that can be exploited to benefit the end consumer.” Understanding the customer drives personalization, personalization creates more engagement, engagements drive more data, so it is a virtuous cycle, Nakai explained. “How many times a day do we look at Facebook, TikTok, Snapchat, Twitter?” he said. “How many times a day do we look at our mobile banking application? We want the future of mobile banking to be as sticky as a Facebook, as a TikTok. And why? Because we will leverage that information, those behaviors of our customers [in ways] that will benefit them.”

Behavioral segmentation drives hyper-personalization

Behavioral analytics is key to driving hyper-personalization in retail banking, Amend said in a presentation. The ability to track customer preferences against real-time data helps banks provide personalized marketing and offers, he said. “[What] drives the future of customer segmentation [is] not demographics, not who you are, but how you bank,” he said. “Understanding that all those groups of customers that are all exhibiting the same pattern will be grouped in the same category. Customer segmentation is no longer [the] nirvana on the top of the maturity curve, it’s that foundation that will drive all the downstream use cases, from credit decisioning to fraud detection to pricing and targeting, and cross-sell and up-sell opportunities. It can all be driven by AI at its core.”

“Customer segmentation is no longer [the] nirvana on the top of the maturity curve, it’s that foundation that will drive all the downstream use cases, from credit decisioning to fraud detection to pricing and targeting, and cross-sell and up-sell opportunities. It can all be driven by AI at its core.”

Antoine Amend, Senior Technical Director, Financial Services at Databricks

Modern compliance and risk management

Financial services organizations operate in a highly regulated environment. “Doing machine learning in a highly regulated environment is hard,” Nakai said. “So, how can we create the right infrastructure — the people, the data, lineage, the governance, all these controls — so that we could do machine learning? We’ve brought in powerful ecosystem partners, like Deloitte and Avanade and Accenture, that can help you implement Lakehouse in a highly regulated industry, with all the controls in place.” Participants discussed navigating compliance and risk management amid a data transformation.

Delivering regulations as code

“[Regulations] usually come to us in these big, hundreds and hundreds of pages legal documents,” Goldbaum said in his keynote, where he discussed delivering regulations as code. “That’s inefficient because every firm needs to interpret that and write that, so there’s a very big efficiency gain if we could just deliver it as code. Of course, there’s a catch to this, which is that every firm has different technologies. So, we want to be able to deliver it as code, while still allowing them to execute it in their technologies, without having to rewrite everything...When we went to the open-source community, we immediately got feedback that there are even better tools, like this Bosque programming language, these verification tools that catch all kinds of bugs that we would never, ever be able to catch before.”

Using AI responsibly

“An emerging topic especially in the chief risk office world is compliance, and even in the strategy world is around, ‘are we using AI technologies responsibly?’” said Anand Premasundar, Chief Data Scientist, Applied Intelligence Financial Services, Accenture, in his keynote. “If so, am I able to explain what I am doing with these deep data science models that are being built?” He described using the lakehouse to answer these questions: “We’re working with another bank where they’re using the lakehouse architecture to drive reviewing what their risk management processes are, both from a data perspective as well as from a model build perspective. And then eventually designing [an] ethical design framework [for] the responsible AI world.”

Sourcing data in a compliant fashion

“Compliance has a role to play, not just when you source data, making sure that the data is sourced in a compliant fashion, but it’s something you have to think about on a continuous basis,” Christopher Darringer,

Software Engineer, Point 72, said in his keynote. “So as we pull this data in-house we have to monitor it, manage it, and make sure that it continues to meet compliance requirements.”

Sustainable finance

Companies in every sector are facing pressure to make net-zero commitments, but the pace of change is not meeting these ambitions, said Betty Jiang, Head of US ESG Research at Credit Suisse, in a keynote. “In a world where everyone is pledging decarbonization, if you are not telling the market how you are getting there, how much it is going to cost, and the data and business transformation for you to get there, that’s just empty words.” Regulations coming down the pike could force companies to back up their ambitions with data. “[There is a] big data problem that we have in ESG analysis,” she said. “There are hundreds of metrics being used to measure companies’ performance on these ESG considerations. But how they’re being measured, or reported by

“In a world where everyone is pledging decarbonization, if you are not telling the market how you are getting there, how much it is going to cost, and the data and business transformation for you to get there, that’s just empty words.”

Betty Jiang, Head of US ESG Research at Credit Suisse

companies, is different company by company. How it’s weighed by different ESG rating agencies varies. How it gets applied and interpreted by the investors also varies from fund to fund and asset manager to asset manager. Using the same ESG data, an oil company could be a bad ESG investment for one investor, and it could be a great opportunity for another investor.”

Data and AI underpin the net-zero transformation

“You have the regulators that are setting policies and decarbonization goals...we’re going to see more standardization of ESG data,” Jiang said. “The international reporting frameworks are being developed and we are going to see more enhanced comparability and transparency of data.” An outlier in a conference of data brains, as an ESG researcher, Jiang focused more on the role she sees data playing in the net-zero transformation. “My presentation will be different because I’m coming from a non-data background. I’m normally a consumer of data. Things magically happen, visualization happens, and I’m trying to [determine] what it means for investment decisions. I’m responsible for US ESG research. Despite my background, hopefully, I will provide some unique perspective on why data is critical for ESG analysis. And in fact, for transforming the entire economy.”

It’s challenging to calculate a carbon footprint

Data standardization is key to tracking ESG progress, Jiang said, but it’s complicated. “You’re not only looking at the emissions that you can control from your direct operations, but the entire value chain, from upstream, your products, how that’s getting produced and sourced and transported, to the use of product and services,” she said. “It’s incredibly difficult for a company to ferret that out.”

How the power of data and AI can transform financial services

Scaling up to meet net-zero goals requires data collaboration

More data and AI is required to meet the demands of these goals, Jiang said, noting that utility companies are providing data analytics to their customers so that they can help them to decarbonize. “Every day, I’m learning about new innovators and disruptors [that are] creating smart grids, smart homes, smart everything, to help to advance this sustainable economy,” Jiang said. “Underpinning all of that is scaling up collaboration...We need companies to work together — Industry customers, suppliers — all need to work together.”

“Underpinning all of that is scaling up collaboration...We need companies to work together — Industry customers, suppliers — all need to work together.”

Betty Jiang, Head of US ESG Research at Credit Suisse

Conclusion

Throughout the event, participants gave strong examples of how data and AI can empower FSI leaders, and how Lakehouse for Financial Services helps organizations build a data asset strategy to enable multiple use cases, including but not limited to: banking and payments; risk management; regulatory compliance; fraud detection, personalized offers, and customer 360. All of a company’s data — transactions, credit risk, macro, demographics, log data, market data, trades, derivatives, exposure, and ESG — enters the lakehouse “across any cloud, any system, any tool, in real-time,” Nakai said. “So now, you can hydrate your Lakehouse, and then execute important use cases from this Lakehouse as quickly as possible.” During the panel, Segu said, “I think Databricks has done a great job of giving you flexibility, allowing you to bring in your own cloud, bring in your own infrastructure, be able to pull in whatever data sources in whatever format that it might exist, to interact with it in a single place.”

About Databricks

Databricks is the data and AI company. Founded by the creators of Apache Spark™, Delta Lake and MLflow, Databricks offers an open and unified platform that simplifies data and AI so data engineers, scientists and analysts can collaborate and innovate faster. Learn more on [Twitter](#), [LinkedIn](#) and [Facebook](#).

Watch the Financial Services Symposium **on-demand**, and learn more about Databricks at dbricks.co/fiserv.