

SPARK+AI SUMMIT 2020

Organized by  databricks

Spark + AI Summit 2020 Guide



Table of Contents

1. Overview and Agenda
2. Ways to Consume Content
3. Navigating the Platform
4. Networking and Activities
5. Get Social!
6. Questions and FAQ

SPARK+AI SUMMIT 2020

Welcome to Spark + AI Summit 2020! We have assembled another outstanding slate of trainings, tutorials, keynotes, and sessions for you. The theme of the conference is **Data Teams Unite!** This is an important moment for data teams everywhere. Together, we can solve the world's toughest problems. That's why we transformed this year's Summit into a global event – totally virtual and open to everyone, free of charge.

My friend DJ Patil, former Chief Data Scientist in the Obama White House, is fond of saying that "Data Science is a team sport." Enjoy the conference, and take time to use the networking features that are available on our conference platform. For the next few days, we are all one data team.

Ben Lorica
Spark + AI Summit Program Chair

SPARK+AI
SUMMIT 2020
JUNE 22-26 | VIRTUAL

THE VIRTUAL EVENT FOR DATA TEAMS

Organized by  databricks



Data Teams Unite!

There's never been a more important moment for data teams. The current environment only underscores the urgency to tackle the biggest challenges by giving data teams what they need. These teams are often geographically dispersed – and right now they are working remotely – but open-source data platforms can unite them wherever they are.

This event is a great manifestation of our theme. It's an incredible opportunity to bring data teams together, show how data teams can become more connected, share innovative ideas and ultimately help move the industry forward.

SPARK+AI SUMMIT

#Datateams #SparkAISummit

Summit by the Numbers

DATA PEOPLE

- 11K+ Data Engineers
- 10K+ Data Scientists
- 4K+ Architects
- 7K+ Software Engineers
- 2K+ Academia

COUNTRIES

136

SESSIONS

- 220+ Breakouts
- 18 Keynotes
- 75+ Training Courses

REGISTRANTS

50K+

Week at a Glance

Pacific Time

TRAINING DAY 1

Monday, June 22

6 – 9 Training

10 – 1 Training

2 – 3:30 Spark Meetup

9 – 9:45 APJ Networking Event

[Online Agenda](#)

Access to platform and agenda builder will be June 18

TRAINING DAY 2

Tuesday, June 23

9 – 12 Training

1 – 4 Training

DAY 1

Wednesday, June 24

8:30 – 10:30 Keynotes

10:30 – 11 Delta Lake GA AMA

11 – 12:45 Sessions

11:30 – 1:00 Government Event

11:30 – 1:00 Education Event

1 – 2:30 Keynotes

2:30 – 3 Nate Silver Q&A

2:30 – 4 Sessions

2:30– 4 Media & Entertainment Event

4 – 5:30 Welcome Party

4:30 – 5 Spark 3.0 + Koalas AMA

DAY 2

Thursday, June 25

6 – 8:30 Financial Services Event

7:30 – 8:15 EMEA Networking Event

9 – 10:30 Keynotes

10:30 – 11 AM Delta Lake VIP AMA

11 – 12:45 Sessions

11:30 – 1:00 Retail & Consumer Goods Event

11:30 – 1:00 Health & Life Sciences Event

11:30 – 12:30 Women in UDA Event

1 – 2 Keynotes

2 – 2:30 Matei Zaharia VIP AMA

2:30 – 4 Sessions

4 – 5:30 Attendee Party

4:30 – 5 MLflow GA AMA

DAY 3

Friday, June 26

9 – 10:30 Keynotes

10:30 – 12:50 Sessions

Keynotes



Nate Silver
Founder, FiveThirtyEight.com



Prof. Jennifer Chayes
Associate Provost, Division of Computing, Data Science, and Society (CDSS)



Adam Paszke
Maintainer of PyTorch



Clemens Mewald
Director of Product Management



Amy Heineke
Principal Product Architect



Rohan Kumar
Corporate Vice President, Azure Data



Kim Hazelwood
West Coast Head of Engineering of FAIR



Dr. Phillip Atiba Goff
Co-Founder and President



Prof. Hany Farid
Digital Forensics Pioneer



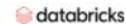
Vish Subramanian
Director of Data and Analytics Engineering



Sue Ann Hong
Software Engineer



Lauren Richie
Software Engineer



Ali Ghodsi
Co-founder & CEO
Original Creator of Apache Spark™



Matej Zaharia
Co-founder & Chief Technologist, Databricks
Original Creator of Apache Spark™ & MLflow



Brooke Wenig
Machine Learning Practice Lead



Reynold Xin
Co-founder & Chief Architect



Sarah Bird
Principal Program Manager



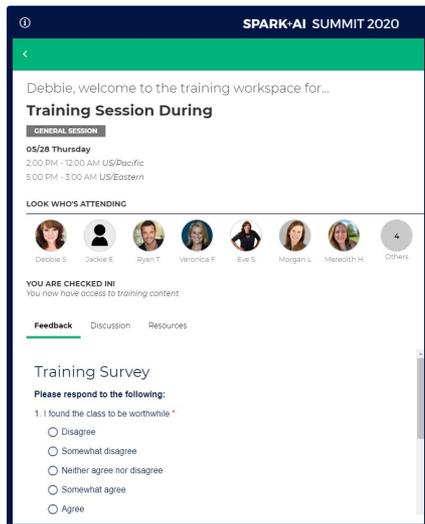
Prepare for the Virtual Environment

1. Our virtual platform is optimized for desktop viewing. Although it will still work with mobile devices, it is recommended to use a desktop computer if you can
2. Make sure to use Google Chrome or Firefox as your browser for the best experience
3. If you signed up for pre-conference training, make sure that you have Zoom
4. Enable your microphone and camera, when you login to take advantage of our fully interactive conference
5. Get Excited!

Content Delivered in a Variety of Ways

Pre-Conference Training

- Instructor-led, live training will offer an engaging and interactive way to learn.
- Utilizing Zoom and our platform to offer an optimal learning experience, with direct access to TAs and other students.

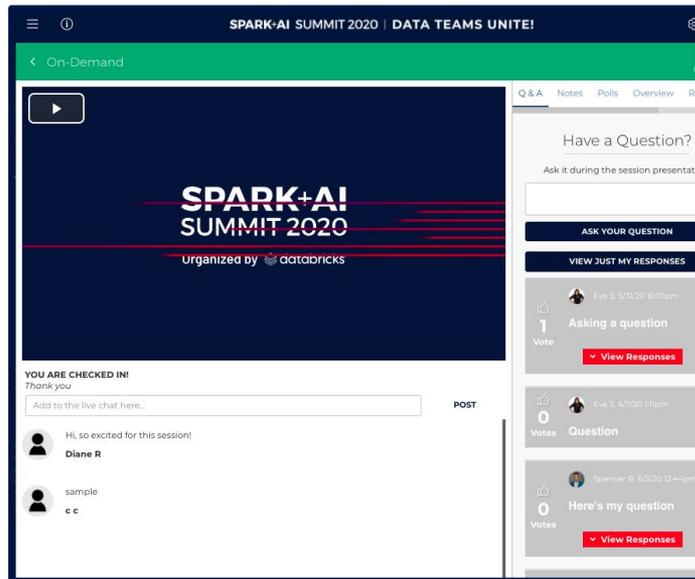


Keynotes

- 5 general sessions featuring industry thought leaders like Ali Ghodsi, Matei Zaharia and Reynold Xin.
- Luminary keynotes from Nate Silver, Hany Farid, Dr. Phillip Goff, Adam Paszke, and many others.

Breakout Sessions

- Over 200 breakout sessions with 9 tracks covering a wide variety of topics in data and AI.
- Session rooms are meant to keep you engaged with the ability to chat with other attendees, ask questions, take notes, provide feedback, and download relevant material



OnDemand Library

Missed a session? Not to worry, sessions will be available on demand immediately after their scheduled time. This does not include pre-conference training

Fireside Chats

There will be a series of fireside chats with industry leaders and experts at the Dev Hub + Expo. Listen in on live discussions at the industry lounges, Databricks booth and more. Be sure to check out the expo agenda for scheduling

SESSION LIBRARY

Recommended

- Breakout**
Wednesday, Jun 24
00 seconds
No Thumbnail Available
- On-Demand**
Wednesday, Jun 24
00 seconds
Distilling Actionable Insights From Your Data With MapR
- On-Demand**
Wednesday, Jun 24
00 seconds
Distilling Actionable Insights From Your Data With MapR

All Videos

FILTER TAGS (0) Alphabetical Live Keyword Search CLEAR FILTER

- A Production Quality Sketching Library for the Analysis of Big Data**
Wednesday, Jun 24
No Thumbnail Available
- A Thorough Comparison of Delta Lake, Iceberg and Hudi**
Wednesday, Jun 24
No Thumbnail Available
- Accelerating Apache Spark Shuffle for Data Analytics on the Cloud with Remote Persistent Memory Pools**
Thursday, Jun 25
No Thumbnail Available
- Accelerating Data Processing in Spark SQL with Pandas UDFs**
Wednesday, Jun 24
No Thumbnail Available

Accenture Fireside Chat: Scaling Mission AI Faster with AIP IQ + Databricks

Attendee Journey

Networking

Enhanced Personalization

Using ML SmartFeed technology we create personalized recommendations across the platform.

REMOVE ME AS SPEAKER

STOP SHARING SCREEN

Presenting Now
Joe

Notes Overview Recommended Feedback

MUTE ALL USERS

Default - Microphone (4- HC)

HD Pro Webcam C920 (046)

Jake Tyler Eve

Networking

Networking questions that are part of your onboarding process will help our platform recommend like-minded people to connect with, as well as suggest sessions to attend.

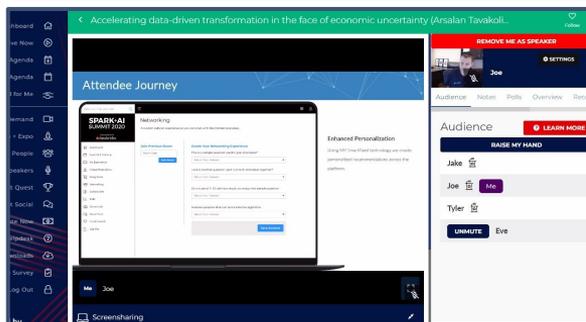
Check out the “Suggested for Me” tab to make the most of this.

Valuable networking can be found across the entire conference from birds of a feather rooms, live chat logs in each of the sessions, special events, and parties.

Special Event

We have a variety of special events that includes industry forums and fireside chats, a Women in Unified Data Analytics panel discussion, networking events for the EMEA and APJ regions, an Apache Spark™ Meetup and more.

Filter out the agenda to see the special event schedule.



Ask Me Anything (AMA) Sessions

Check out the agenda for engaging conversation with industry leaders and experts. Show up early as they are first come, first serve.

Wednesday, June 24:

- 10:30 AM: Delta Lake - General Admission
- 2:00 PM: Nate Silver - General Admission
- 4:30 PM: Spark 3.0 + Koalas - General Admission

Thursday, June 25:

- 10:30 AM: Delta Lake - VIP Only
- 2:00 PM: Matei Zaharia - VIP Only
- 4:30 PM: MLflow - General Admission

Navigating the Platform

SPARK+AI SUMMIT

Dashboard

1. Your personalized dashboard that houses everything you need to know about the conference.
2. The left navigation panel will help you explore all aspects of the conference.
3. Make sure to look out for notifications and your inbox, so you don't miss any updates.

The screenshot displays a personalized dashboard for the SPARK+AI SUMMIT 2020. The header includes the event name and the slogan "DATA TEAMS UNITE!". A search bar is located at the top left. The left navigation menu contains options such as Dashboard, Live Now, My Agenda, Summit Agenda, Suggested for Me, OnDemand, Dev Hub + Expo, Data People, Speakers, Summit Quest, Get Social, My Downloads, Conference Survey, and Log Out. The main content area is titled "Welcome, Ali" and shows a schedule for Thursday, May 28, 2020, Saturday, June 6, 2020, and Monday, June 22, 2020. A video player is featured on the right, showing a woman speaking at a microphone. Below the video are several red buttons with white text and arrows: "LIVE NOW", "NOTIFICATIONS", "MY INBOX", "SUGGESTED FOR ME", and "MY MEET-UPS". At the bottom, it says "Thank you to our sponsors" and "Organized by databricks".

Build Your Profile!

The screenshot shows a profile editing form with a navigation sidebar on the left. The sidebar includes links for OnDemand, My Profile, Speakers, My Queue, My Social, My Agenda, My Survey, and Log Out. The main form area contains a profile picture placeholder with 'ROTATE LEFT' and 'ROTATE RIGHT' buttons, and a 'DELETE PROFILE IMAGE' button. Below the picture are input fields for Title (pre-filled with 'VP, Strategic Accounts'), Company (pre-filled with 'Type A Events'), and Department (pre-filled with 'Conferences + Events'). There is also a field for 'Twitter Handle'. A 'Bio' section has a text area with the placeholder 'Bio text would go here.' Below the bio are two columns of toggle switches for privacy and notification settings. The form ends with an 'UPDATE PROFILE' button.

OnDemand
My Profile
Speakers
My Queue
My Social
My Agenda
My Survey
Log Out

Rotating image controls: ROTATE LEFT, ROTATE RIGHT, DELETE PROFILE IMAGE

Title: VP, Strategic Accounts
Company: Type A Events
Department: Conferences + Events
Twitter Handle: [Empty]
Bio: Bio text would go here.

Make my profile public: ON
Display my email address on my profile page: ON
Allow notifications: ON
Notify me for activity on my posts: ON
Notify me for individual and group chat messages: ON
Notify me for agenda related activity: ON
Allow users to video chat with me: OFF
Allow users to send me messages: ON
Allow users to send me Meetup requests: OFF
Provide my information to sponsors: ON

Update Profile: UPDATE PROFILE

Upon logging in, you will be able to build your profile, manage all of your settings, answer networking questions, and plug information that you would like public. You can go back to change your settings at any time.

The screenshot shows the user profile page for Debbie Seabloom. The top navigation bar includes a search icon, a user icon, and the event name 'SPARK+AI SUMMIT 2020'. The profile header shows the name 'Debbie Seabloom' and title 'VP, Strategic Accounts', with 'about' and 'follow' buttons. A navigation sidebar on the left lists: Dashboard, Live Now, My Agenda, Summit Agenda, Suggested for Me, OnDemand, Dev Hub + Expo, Data People, Speakers, Summit Quest, Get Social, My Downloads, Conference Survey, and Log Out. The main content area has an 'EDIT YOUR PROFILE' button, followed by sections for 'COMPANY' (Type A Events), 'EMAIL' (debbie@typeaevents.com), and 'BIOGRAPHY' (Bio text would go here.). An 'ADD NOTE' button is at the bottom.

Search, SPARK+AI SUMMIT 2020, Debbie Seabloom, VP, Strategic Accounts, about, follow, EDIT YOUR PROFILE, COMPANY: Type A Events, EMAIL: debbie@typeaevents.com, BIOGRAPHY: Bio text would go here., ADD NOTE

Navigation: Dashboard, Live Now, My Agenda, Summit Agenda, Suggested for Me, OnDemand, Dev Hub + Expo, Data People, Speakers, Summit Quest, Get Social, My Downloads, Conference Survey, Log Out

Anatomy of a Breakout Session

The screenshot shows the main interface of the SPARK+AI SUMMIT 2020. At the top, there is a search bar and the event title. Below that, a green header bar displays the session title 'Preventing Abuse Using Unsupervised Learning'. The main content area features the event logo 'SPARK+AI SUMMIT 2020' and the organizer 'Organized by databricks'. A callout box at the bottom left, highlighted with an orange border, contains the text 'YOU ARE CHECKED IN! You may now post discussion comments.' and a text input field 'Add to the live chat here...' with a 'POST' button. A yellow arrow points from this callout to the text 'Check in for points and participate in the chat log' located below the event title. The left sidebar contains navigation options like 'Dashboard', 'Live Now', 'My Agenda', and 'Suggested for Me'. The bottom right has a navigation menu with 'Notes', 'Overview', 'Recommended', and 'Feedback'.

This screenshot shows a close-up of the 'Notes' panel. At the top, there are tabs for 'Notes', 'Overview', 'Recommended', and 'Feedback', with 'Notes' being the active tab. Below the tabs is a text input field with the placeholder 'Type note here...'. To the right of this field is an 'ADD NOTE' button. Below the input field, the text 'Notes Taken' is displayed in green. Underneath, a light blue box contains the message 'You have not taken any notes'.

Use the right panel to take notes, add questions for live Q&A, give your feedback, and download resources specific to the session.

Dev Hub + Expo

Visit our Dev Hub and Expo where you can find the Advisory Lounge, Industry Lounges, Sponsor Booths and Theater.

The screenshot displays the SPARK+AI Summit 2020 website interface. The header includes the event title "SPARK+AI SUMMIT 2020 | DATA TEAMS UNITE!" and navigation icons. The main content is divided into two columns. The left column features a "Featured" section with a schedule for Friday, June 12, 2020, and Monday, June 22, 2020. It lists events such as "Practical Problem-Solving in Retail: Real-Time Data Analytics with Apache Spark" and "Databricks Booth Office Hours", each with a "SPONSORED BY:" field and details like time and location. Below this is a "Spark + AI by Industry Lounges and Networking" section with categories like "Financial Services", "Government and Education", "Healthcare and Life Sciences Experience", "Media and Entertainment", and "Retail and Consumer Goods". The right column is titled "Developer Hub" and includes sections for "Advisory Lounge", "Theater", "Job Board", and "Exhibitors". The "Exhibitors" section features a search bar and a list of sponsors, including Databricks (labeled as the Organizing Sponsor), Immuta, Microsoft Azure, Avanade, AWS, and Informatica.

Databricks Booth

Be sure to visit the Databricks booth during Spark + AI Summit. Our team will answer any questions you have during office hours.

You'll also be able to watch demos on:

- Delta Lake
- MLflow
- Apache Spark
- Koalas
- Cybersecurity
- Intro to Databricks
- and more!

You won't want to miss the Transforming Security Analytics for the Cloud Age Fireside Chat on June 25th at 11 am PT!

SPARK+AI SUMMIT 2020

Databricks Booth

If a scheduled broadcast has begun and you do not see the presenter please refresh the page

Exhibitor Schedule

- Databricks Booth Office Hours**
Wednesday June 24, 2020
11:00 US/Pacific - 3:45 PM US/Pacific
[ADD](#)
- Databricks Booth Office Hours**
Thursday June 25, 2020
11:00 US/Pacific - 3:45 PM US/Pacific
[ADD](#)
- Databricks Booth Office Hours**
Friday June 26, 2020
10:00 US/Pacific - 1:00 PM US/Pacific
[ADD](#)

Virtual Plinko! You could win Databricks swag shipped right to your home! Stop by each day for another chance to play and win.

TAP TO BEGIN

VIEW PRIZES

Job Postings Recordings Take Poll About

Fireside Chats Demo's

No Image Available

@@TITLE@@
@TIME@

SPARK+AI SUMMIT 2020
Organized by @databricks

Networking & Activities

SPARK+AI SUMMIT

Suggested for Me

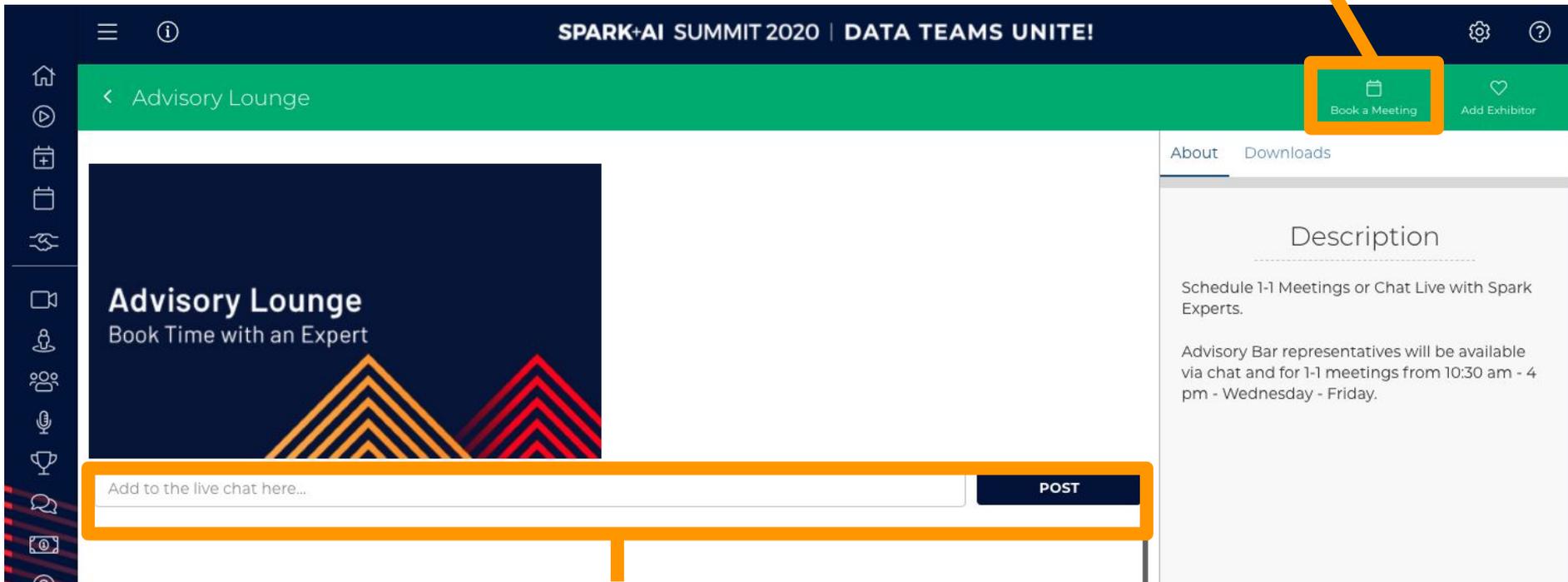
Questions upon onboarding and your activity within the platform will help populate recommended data people to connect with, as well as sessions of interest.

The screenshot displays the SPARK+AI SUMMIT 2020 interface with the header "DATA TEAMS UNITE!". The main content area is titled "LIKE MINDED INDIVIDUALS" and includes a red button labeled "REBUILD YOUR NETWORK". Below this, a message states: "Thanks Debbie! Based on your responses, we think you might enjoy these channels and people." The section "LIKE MINDED INDIVIDUALS" lists four attendees with their names, roles, and "CHAT" buttons: Ashley Andrada (Qubole), Bart K (MeetingPlay), Caitlin Guyton, and Cathy Nguyen Parker (infoworks). Below this is a red button for "BIRDS-OF-A-FEATHER ROOMS". The "SESSIONS OF INTEREST" section lists three sessions: "Databricks Administration", "Unified Data Analytics for Managers", and "Building Time Series Forecasting Models using Neural Network and Statistical Models". Each session entry includes the date (Monday, June 22) and time (6:00 AM US/Pacific - 9:00 AM US/Pacific or 8:00 AM US/Pacific - 11:00 AM US/Pacific). A "ks" logo is visible in the bottom left corner of the interface.

Advisory Lounge

Attendees can chat with Spark and AI Experts or schedule a meeting for a time that's convenient.

Book meetings in advance by clicking here!



Just have a couple of questions? Ask an expert through live chat

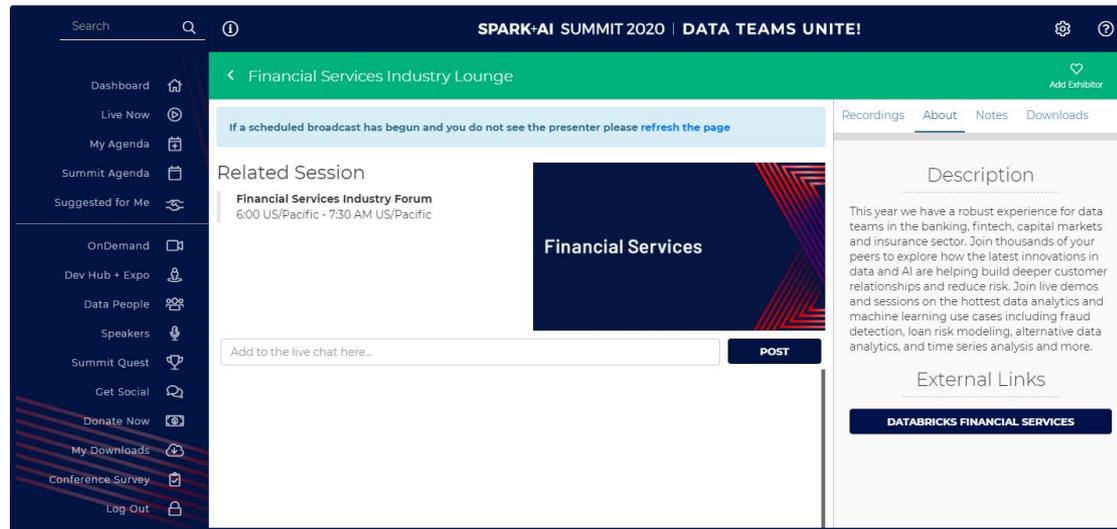
Industry Lounges



Visit the Dev Hub + Expo to check out our industry specific experiences

Explore forums, demos, and curated tracks across these five industries:

1. Media & Entertainment
2. Health & Life Sciences
3. Financial Services
4. Retail & Consumer Goods
5. Government & Education



Summit Quest

Complete challenges to earn points. And the more points you earn, the more swag you can get from the SWAG Store.*

Challenges encourage attendees to engage more deeply in the Summit by visiting sponsor booths, watching sessions, and more.

VIP attendees earn 8000 points upon entry as a part of their registration.

* While supplies last, some items limited for shipments outside of the US.

SPARK+AI SUMMIT 2020 | DATA TEAMS UNITE!

SUMMIT QUEST

899 My Points **10** Position

It's game on with Data Quest! Collect points for engaging in the conference and win real swag!

How Do I Earn Points?
You can earn points simply by participating in the conference. The game starts Monday, June 22nd. See below for ways to score big.

What Can I Win?
Redeemed points for swag! The Swag Store will open at up 7:00AM PT on Thursday, June 25th. You will be able to select from a number of Spark + AI Summit Swag Packages based on how many points you've earned. Swag will be available only while supplies last.
Swag will be shipped directly to you! During checkout you will be able to enter a shipping address of your choice. Note, not all items are available for international shipping. Please allow 6-8 weeks following the conference for your shipment to arrive.

HUNT FOR POINTS **SWAG STORE**

My Points

- 179 WATCH SESSIONS LIVE OR ON-DEMAND
- 0 FILL OUT A SURVEY FOLLOWING A SESSION
- 0 TAKE A QUIZ AT THE DATABRICKS BOOTH
- 0 FILL OUT THE CONFERENCE SURVEY
- 720 VISIT SPONSORS AND VERTICAL LOUNGES

The leaderboard updates every 15 minutes

Meredith H.	8029
Eve S.	2101
Morgan L.	2026
Lisa V.	1837

SPARK+AI SUMMIT 2020 | DATA TEAMS UNITE!

Cart Redeem

- Hat (available internationally)**
points: 3,500
100 items in stock
- Water Bottle (only available in US)**
points: 4,000
1,998 items in stock
- Backpack (only available in US)**
points: 4,500
2,000 items in stock
- Third Tier (available internationally)**
points: 5,000
1,500 items in stock
- Second Tier (only available in the US)**
points: 7,000
600 items in stock
- Top Tier Package (only available in the US)**
points: 9,000
400 items in stock

Don't Forget the Fun!

Evening
Events &
Body
Breaks

SPARK+AI SUMMIT 2020

Organized by databricks

Example Domain

This domain is for use in illustrative examples in documents. You may use this domain in literature without prior coordination or asking for permission.

ASK ME ANYTHING DEVELOPER HUB SUGGESTED FOR ME

SPARK+AI SUMMIT 2020 | DATA TEAMS UNITE!

MLflow Yoga: a holistic workout to connect your mind, body, and production data system

Organized by databricks

CHECK INTO SESSION AND EARN SUMMIT QUEST POINTS
Enables session discussion

Join us for a 20 virgasa flow yoga practice that will leave you feeling energized and rejuvenated for Summit. We'll start with a quick warm-up, flow through a breathe-for-movement sequence, include a 3-minute abs circuit, and end with stretching. This all-level class will include modifications if you're brand new to yoga and for our experienced yogis, the ability to level up. See you there.

Get Social!

SPARK+AI SUMMIT

Daily Social Challenges

Attendees are encouraged to participate in social media challenges, ranging from favorite coding language to representing school spirit.

June 18 - 23	June 24	June 25	June 26
<p>Take a minute to introduce your data team to the rest of the community. Share what you're working on, or tell us what you're looking forward to at Summit. Get creative and make sure to use #SparkAISummit and #DataTeams!</p>	<p>Share your favorite library or programming language with us, hold up a sign or wear your gear. Get creative and make sure to use #SparkAISummit and #DataTeams!</p>	<p>Show us where you're from. Wear some gear to rep your hometown or alma mater and take a selfie. Get creative and make sure to use #SparkAISummit and #DataTeams!</p>	<p>Let us know what your favorite learning is from this year's Spark and AI Summit - what are you excited to share with your team? Get creative and make sure to use hashtags #SparkAISummit and #DataTeams!</p>



#Datateams #SparkAISummit

Questions?

Contact us

registration@spark-summit.org

organizers@spark-summit.org

Documentation

[FAQ](#)

[Event Policy](#)