(https://spark-summit.org/eu-2015)

DATA SCIENCE AND ENGINEERING AT SCALE

AMSTERDAM · OCTOBER 27 – 29, 2015

SOLD OUT ()

Home (https://spark-summit.org/eu-2015/)   Schedule (https://spark-summit.org/eu-2015/schedule/)

Spark Training (https://spark-summit.org/eu-2015/spark-training/)   Speakers (https://spark-summit.org/eu-2015/speakers/)

Venue (https://spark-summit.org/eu-2015/venue/)   Sponsors (https://spark-summit.org/eu-2015/sponsors/)

Job Board (https://spark-summit.org/eu-2015/jobs/)

Jump to: Day 1 ⌄   Day 2 ⌄   Day 3 ⌄

DAY 1 · TUESDAY, OCTOBER 27 · CONFERENCE

07:00 **Registration**

	ADMINISTRATIEZAAL	BERLAGE ZAAL	GRAANBEURSZAAL	VEILINGZAAL
09:00	**Training: Spark Essentials (https://spark-summit.org /eu-2015/events/training-spark-essentials/)** Adam Breindel (https://spark-summit.org/eu-2015/speakers /adam-breindel/) (Databricks) (3 hours)	**Training:Advanced: Exploring Wikipedia with Spark (https://spark-summit.org/eu-2015 /events/trainingadvanced-exploring-wikipedia-with-spark/)** Sameer Farooqui (https://spark-summit.org /eu-2015/speakers/sameer-farooqui/) (Databricks) (3 hours)	**Training: Data Science with Apache Spark (https://spark-summit.org /eu-2015/events/training-data-science-with-apache-spark/)** Jon Bates (https://spark-summit.org/eu-2015/speakers /jon-bates/) (Databricks) (3 hours)	**Training: Spark Essentials (https://spark-summit.org /eu-2015/events/training-spark-essentials/)** Brian Clapper (https://spark-summit.org/eu-2015/speakers /brian-clapper/) (Databricks) (3 hours)

12:00 **Lunch**

ADMINISTRATIEZAAL	BERLAGE ZAAL	GRAANBEURSZAAL	VEILINGZAAL

| 13:00 | Training Continues: Spark Essentials (https://spark-summit.org/eu-2015/events/training-continues-spark-essentials/)

(4 hours) | Training Continues: Advanced: Exploring Wikipedia with Spark (https://spark-summit.org/eu-2015/events/training-continues-advanced-exploring-wikipedia-with-spark/)

(4 hours) | Training Continues: Data Science with Apache Spark (https://spark-summit.org/eu-2015/events/training-continues-data-science-with-apache-spark/)

(4 hours) | Training Continues: Spark Essentials (https://spark-summit.org/eu-2015/events/training-continues-spark-essentials/)

(4 hours) |

17:00 **End of Day 1**

18:00 **Spark Meetup**

DAY 2 · WEDNESDAY, OCTOBER 28 · CONFERENCE

07:00 **Registration**

09:00 **How Spark Usage is Changing in 2015 (https://spark-summit.org/eu-2015/events/keynote1/)**

Matei Zaharia (https://spark-summit.org/eu-2015/speakers/matei-zaharia/) (Databricks)

EFFECTENBEURSZAAL

2015 was a year of continued growth for Spark, with numerous additions to the core project and very fast growth of use cases across the industry. In this talk, I'll look back at how the… (30 minutes)

09:30 Keynote (https://spark-summit.org/eu-2015/events/keynote2/)

Ion Stoica (https://spark-summit.org/eu-2015/speakers/ion-stoica/) (Databricks)

EFFECTENBEURSZAAL

09:50 Spark in the Enterprise: Lessons Learned From the Past Two Years (https://spark-summit.org/eu-2015/events/keynote3/)

Charles Zedlewski (https://spark-summit.org/eu-2015/speakers/charles-zedlewski/) (Cloudera)

EFFECTENBEURSZAAL

10:00 Apache Spark: The Analytics Operating System (https://spark-summit.org/eu-2015/events/keynote4/)

Anjul Bhambhri (https://spark-summit.org/eu-2015/speakers/anjul-bhambhri/) (IBM)

EFFECTENBEURSZAAL

In this keynote, IBM's Anjul Bhambhri will highlight how Spark is stretching the boundaries of existing Business Intelligence applications into a new generation of analytics infused applications, and IBM's contribution to value creation with Spark. (10 minutes)

10:10 How Spark is Making an Impact at Goldman Sachs (https://spark-summit.org/eu-2015/events/keynote5/)

Vincent Saulys (https://spark-summit.org/eu-2015/speakers/vincent-saulys/) (Goldman Sachs)

EFFECTENBEURSZAAL

In this keynote, Vincent Saulys will discuss how Spark was brought into Goldman, the impact its making today (lessons learned along the way), and where they expect to apply Spark in the future (20 minutes)

10:30 Break

EFFECTENBEURSZAAL   GRAANBEURSZAAL

	DEVELOPER TRACK	APPLICATIONS TRACK
		DATA SCIENCE TRACK
11:00	**Spark DataFrames: Simple and Fast Analysis of Structured Data (https://spark-summit.org/eu-2015/events/spark-dataframes-simple-and-fast-analysis-of-structured-data/)** Michael Armbrust (https://spark-summit.org/eu-2015/speakers /michael-armbrust/) (Databricks) (30 minutes)	**3 Trillion App recommendations, with less than 100 Lines of Spark Code in less than 25 Minutes (https://spark-summit.org/eu-2015/events/3-trillion-app-recommendations-with-less-than-100-lines-of-spark-code-in-less-than-25-minutes/)** Ayman Farahat (https://spark-summit.org/eu-2015/speakers /ayman-farahat/) (Yahoo) (30 minutes)
11:35	**Magellan: Geospatial Analytics on Spark (https://spark-summit.org/eu-2015/events/magellan-geospatial-analytics-on-spark/)** Ram Sriharsha (https://spark-summit.org/eu-2015/speakers /ram-sriharsha/) (Hortonworks) (30 minutes)	**Hundreds of queries in the time of one (https://spark-summit.org/eu-2015/events/hundreds-of-queries-in-the-time-of-one/)** Gianmario Spacagna (https://spark-summit.org/eu-2015 /speakers/gianmario-spacagna/) (Barclays) (30 minutes)
12:10	**Elasticsearch and Apache Lucene for Apache Spark and MLib (https://spark-summit.org/eu-2015/events /elasticsearch-and-apache-lucene-for-apache-spark-and-mlib/)** Britta Weber (https://spark-summit.org/eu-2015/speakers/britta-weber/) (elastic) (30 minutes)	**Combining the Strengths of MLlib, scikit-learn, and R (https://spark-summit.org/eu-2015/events/combining-the-strengths-of-mllib-scikit-learn-and-r/)** Joseph Bradley (https://spark-summit.org/eu-2015/speakers /joseph-bradley/) (Databricks) (30 minutes)

12:45 **Lunch**

<table>
<tr><td></td><td align="center">EFFECTENBEURSZAAL</td><td align="center">GRAANBEURSZAAL</td></tr>
<tr><td></td><td>DEVELOPER TRACK</td><td>APPLICATIONS TRACK</td></tr>
</table>

13:50 **Spark in Production: Lessons from 100+ production users (https://spark-summit.org/eu-2015/events/spark-in-production-lessons-from-100-production-users/)**

Aaron Davidson (https://spark-summit.org/eu-2015/speakers/aaron-davidson/) (Databricks)

(30 minutes)

Big Telco Real-Time Network Analytics (https://spark-summit.org/eu-2015/events/big-telco-real-time-network-analytics/)

Yousun Jeong (https://spark-summit.org/eu-2015/speakers/yousun-jeong/) (SK Telecom)

(30 minutes)

14:25 **Building a REST Job Server for interactive Spark as a service (https://spark-summit.org/eu-2015/events/building-a-rest-job-server-for-interactive-spark-as-a-service/)**

Romain Rigaux (https://spark-summit.org/eu-2015/speakers/romain-rigaux/) (Cloudera)
Erick Tryzelaar (https://spark-summit.org/eu-2015/speakers/erick-tryzelaar/) (Cloudera)

(30 minutes)

Making better news with Spark (https://spark-summit.org/eu-2015/events/making-better-news-with-spark/)

Philip Wills (https://spark-summit.org/eu-2015/speakers/philip-wills/) (The Guardian)

(30 minutes)

15:00

A Scalable Implementation of Deep Learning on Spark (https://spark-summit.org/eu-2015/events/a-scalable-implementation-of-deep-learning-on-spark/)

Alexander Ulanov (https://spark-summit.org/eu-2015/speakers/alexander-ulanov/) (Hewlett-Packard Labs)

(30 minutes)

Using Spark in an IoT Analytics Platform Enable breakthroughs in Parkinson Disease Research (https://spark-summit.org/eu-2015/events/using-spark-in-an-iot-analytics-platform-enable-breakthroughs-in-parkinson-disease-research/)

Ido Karavany (https://spark-summit.org/eu-2015/speakers/ido-karavany/) (Intel)

(30 minutes)

15:30 Break

EFFECTENBEURSZAAL

APPLICATIONS TRACK

DEVELOPER TRACK

GRAANBEURSZAAL

APPLICATIONS TRACK

16:00

Jaws - Data Warehouse with Spark SQL (https://spark-summit.org/eu-2015/events/jaws-data-warehouse-with-spark-sql/)

Ema Orhian (https://spark-summit.org/eu-2015/speakers/ema-orhian/) (Atigeo)

(30 minutes)

Dictionary Based Annotation at Scale with Spark, SolrTextTagger and OpenNLP (https://spark-summit.org/eu-2015/events/dictionary-based-annotation-at-scale-with-spark-solrtexttagger-and-opennlp/)

Sujit Pal (https://spark-summit.org/eu-2015/speakers/sujit-pal/) (Elsevier)

(30 minutes)

16:35 Productionizing Spark and the Spark REST Job Server (https://spark-summit.org/eu-2015/events /productionizing-spark-and-the-spark-rest-job-server/)

Evan Chan (https://spark-summit.org/eu-2015/speakers/evan-chan/) (TupleJump)

(30 minutes)

How Spark enables the Internet of Things : Efficient integration of multiple Spark components for Smart City use cases (https://spark-summit.org/eu-2015/events /how-spark-enables-the-internet-of-things-efficient-integration-of-multiple-spark-components-for-smart-city-use-cases/)

Paula Ta-Shma (https://spark-summit.org/eu-2015/speakers /paula-ta-shma/) (IBM)

(30 minutes)

17:10 The Pushdown of Everything (https://spark-summit.org /eu-2015/events/the-pushdown-of-everything/)

Stephan Kessler (https://spark-summit.org/eu-2015/speakers /stephan-kessler/) (SAP)
Santiago Mola (https://spark-summit.org/eu-2015/speakers /santiago-mola/) (Stratio)

(30 minutes)

Real-Time Anomaly Detection with Spark ML and Akka (https://spark-summit.org/eu-2015/events/real-time-anomaly-detection-with-spark-ml-and-akka/)

Natalino Busa (https://spark-summit.org/eu-2015/speakers /natalino-busa/) (ING)

(30 minutes)

17:40 Attendee Networking Reception

18:00 Spark R Meetup

DAY 3 · THURSDAY, OCTOBER 29 · CONFERENCE

07:00 **Registration**

09:00 **Spark—The Ultimate Scala Collections (https://spark-summit.org/eu-2015/events/spark-the-ultimate-scala-collections/)**

Martin Odersky (https://spark-summit.org/eu-2015/speakers/martin-odersky/) (Typesafe)

EFFECTENBEURSZAAL

Mike Olson from Cloudera has called Spark the killer App for Scala. Not only is it written in Scala but it is also deeply integrated in its toolchain, making intensive use of the Scala REPL... (30 minutes)

09:30 **Keynote (https://spark-summit.org/eu-2015/events/keynote7/)**

Franz Faerber (https://spark-summit.org/eu-2015/speakers/franz-faerber/) (SAP HANA Vora)

EFFECTENBEURSZAAL

09:45 **A Look Ahead at Spark's Development (https://spark-summit.org/eu-2015/events/keynote8/)**

Reynold Xin (https://spark-summit.org/eu-2015/speakers/reynold-xin/) (Databricks)

EFFECTENBEURSZAAL

10:00 **Spark and Cassandra: An Amazing Apache Love Story (https://spark-summit.org/eu-2015/events/keynote9/)**

Patrick McFadin (https://spark-summit.org/eu-2015/speakers/patrick-mcfadin/) (Datastax)

EFFECTENBEURSZAAL

It's been a little over a year since the release of the Spark Connector for Apache Cassandra and there has been incredible interest and adoption. In that time, I've been traveling the world, talking to... (15 minutes)

10:15 **Keynote (https://spark-summit.org/eu-2015/events/keynote10/)**

Arun Murthy (https://spark-summit.org/eu-2015/speakers/arun-murthy/) (Hortonworks)
EFFECTENBEURSZAAL

10:30 **Break**

EFFECTENBEURSZAAL	GRAANBEURSZAAL	VEILINGZAAL
DEVELOPER TRACK	DATA SCIENCE TRACK	RESEARCH TRACK

11:00   **Securing your Spark Applications (https://spark-summit.org/eu-2015/events/securing-your-spark-applications/)**

Developer Track:

Securing your Spark Applications (https://spark-summit.org/eu-2015/events/securing-your-spark-applications/)

Kostas Sakellis (https://spark-summit.org/eu-2015/speakers/kostas-sakellis/) (Cloudera)
Marcelo Vanzin (https://spark-summit.org/eu-2015/speakers/marcelo-vanzin/) (Cloudera)

(30 minutes)

Data Science Track:

Data Science at Scale: Using Apache Spark for Data Science at Bitly (https://spark-summit.org/eu-2015/events/data-science-at-scale-using-apache-spark-for-data-science-at-bitly/)

Sarah Guido (https://spark-summit.org/eu-2015/speakers/sarah-guido/) (Bitly)

(30 minutes)

Research Track:

Data-centric Metaprogramming (https://spark-summit.org/eu-2015/events/data-centric-metaprogramming/)

Vlad Ureche (https://spark-summit.org/eu-2015/speakers/vlad-ureche/) (EPFL)

(30 minutes)

11:35

Exploiting GPUs for columnar DataFrame operations (https://spark-summit.org/eu-2015/events /exploiting-gpus-for-columnar-dataframe-operations/)

Kiran Lonikar (https://spark-summit.org /eu-2015/speakers/kiran-lonikar/) (Informatica)

(30 minutes)

Enabling exploratory data science with Spark and R (https://spark-summit.org/eu-2015/events /enabling-exploratory-data-science-with-spark-and-r/)

Hossein Falaki (https://spark-summit.org /eu-2015/speakers/hossein-falaki/) (Databricks)

(30 minutes)

Distributed Data Processing Using Spark in Radio Astronomy (https://spark-summit.org/eu-2015 /events/distributed-data-processing-using-spark-in-radio-astronomy/)

Panos Labropoulos (https://spark-summit.org/eu-2015/speakers/panos-labropoulos/) (Bright Computing, Inc. / Univ. of Athens)
Sarod Yatawatta (https://spark-summit.org/eu-2015/speakers/sarod-yatawatta/) (ASTRON)

(30 minutes)

12:10

Time series stream processing with Spark and Cassandra (https://spark-summit.org/eu-2015/events/time-series-stream-processing-with-spark-and-cassandra/)

Christopher Batey (https://spark-summit.org/eu-2015/speakers /christopher-batey/) (Cloudance Ltd)

(30 minutes)

Using Natural Language Processing on Non-Textual Data with MLLib (https://spark-summit.org/eu-2015 /events/using-natural-language-processing-on-non-textual-data-with-mllib/)

Casey Stella (https://spark-summit.org /eu-2015/speakers/casey-stella/) (Hortonworks)

(30 minutes)

Spark Deployment and Performance Evaluation on a Petascale HPC Setup (https://spark-summit.org/eu-2015 /events/spark-deployment-and-performance-evaluation-on-a-petascale-hpc-setup/)

Jordi Torres (https://spark-summit.org /eu-2015/speakers/jordi-torres/) (Barcelona Supercomputing Center)

(30 minutes)

12:45 Lunch

	EFFECTENBEURSZAAL	GRAANBEURSZAAL
	DEVELOPER TRACK	RESEARCH TRACK
		DATA SCIENCE TRACK

13:50

SparkUI visualization: a lens into your application (https://spark-summit.org/eu-2015/events/sparkui-visualization-a-lens-into-your-application/)

Andrew Or (https://spark-summit.org/eu-2015/speakers/andrew-or/) (Databricks)

(30 minutes)

Sparkling Random Ferns: From an academic paper to spark-packages.org (https://spark-summit.org/eu-2015/events/sparkling-random-ferns-from-an-academic-paper-to-spark-packagesorg/)

Piotr Dendek (https://spark-summit.org/eu-2015/speakers/piotr-dendek/) (University of Warsaw, Interdisciplinary Centre for Mathematical and Computational Modelling)
Mateusz Fedoryszak (https://spark-summit.org/eu-2015/speakers/mateusz-fedoryszak/) (University of Warsaw, Interdisciplinary Centre for Mathematical and Computational Modelling)

(30 minutes)

14:25

Streaming Analytics with Spark, Kafka, Cassandra, and Akka (https://spark-summit.org/eu-2015/events/streaming-analytics-with-spark-kafka-cassandra-and-akka/)

Helena Edelson (https://spark-summit.org/eu-2015/speakers/helena-edelson/) (TupleJump)

(30 minutes)

Best Practices for Running PySpark (https://spark-summit.org/eu-2015/events/best-practices-for-running-pyspark/)

Juliet Hougland (https://spark-summit.org/eu-2015/speakers/juliet-hougland/) (Cloudera)

(30 minutes)

15:00   **Spark Streaming: Pushing the Throughput Limits, the Reactive Way (https://spark-summit.org/eu-2015/events /spark-streaming-pushing-the-throughput-limits-the-reactive-way/)**

Gerard Maas (https://spark-summit.org/eu-2015/speakers/gerard-maas/) (Virdata (Technicolor S.A))

François Garillot (https://spark-summit.org/eu-2015/speakers /franois-garillot/) (Typesafe)

(30 minutes)

Sparking Science up with Research Recommendations (https://spark-summit.org/eu-2015/events/sparking-science-up-with-research-recommendations/)

Maya Hristakeva (https://spark-summit.org/eu-2015/speakers /maya-hristakeva/) (Mendeley Ltd)

(30 minutes)

15:30   **Break**

EFFECTENBEURSZAAL

APPLICATIONS TRACK

GRAANBEURSZAAL

DATA SCIENCE TRACK

16:00   **Leverage Mesos for running Spark Streaming production jobs (https://spark-summit.org/eu-2015/events/leverage-mesos-for-running-spark-streaming-production-jobs/)**

Iulian Dragos (https://spark-summit.org/eu-2015/speakers/iulian-dragos/) (Typesafe)

Luc Bourlier (https://spark-summit.org/eu-2015/speakers /luc-bourlier/) (Typesafe)

(30 minutes)

Data Science lifecycle with Zeppelin and Spark (https://spark-summit.org/eu-2015/events/data-science-lifecycle-with-zeppelin-and-spark/)

Moon soo Lee (https://spark-summit.org/eu-2015/speakers /moon-soo-lee/) (NFLabs)

(30 minutes)

16:35

Production Spark and Tachyon Use Cases (https://spark-summit.org/eu-2015/events/production-spark-and-tachyon-use-cases/)

Gene Pang (https://spark-summit.org/eu-2015/speakers/gene-pang/) (Tachyon Nexus, Inc)

(30 minutes)

17:10

Adding Complex Data to Spark Stack (https://spark-summit.org/eu-2015/events/adding-complex-data-to-spark-stack/)

Tug Grall (https://spark-summit.org/eu-2015/speakers/tug-grall/) (MapR)

(30 minutes)

Share and analyse genomic data at scale with Spark, Adam, Tachyon and the Spark Notebook (https://spark-summit.org/eu-2015/events/share-and-analyse-genomic-data-at-scale-with-spark-adam-tachyon-and-the-spark-notebook/)

Andy Petrella (https://spark-summit.org/eu-2015/speakers/andy-petrella/) (Data Fellas)
Xavier Tordoir (https://spark-summit.org/eu-2015/speakers/xavier-tordoir/) (Data Fellas)

(30 minutes)

17:40 **End of Day 3**

Home (https://spark-summit.org/eu-2015/)
Schedule (https://spark-summit.org/eu-2015/schedule/)
Spark Training (https://spark-summit.org/eu-2015/spark-training/)
Speakers (https://spark-summit.org/eu-2015/speakers/)
Venue (https://spark-summit.org/eu-2015/venue/)
Sponsors (https://spark-summit.org/eu-2015/sponsors/)
Job Board (https://spark-summit.org/eu-2015/jobs/)

Event Policy (https://spark-summit.org/event-policy/)
About Spark Summit (https://spark-summit.org/about/)

Spark Summit Home (/)
Past Events
Subscribe to the Newsletter

Organized by (https://databricks.com/)

If you have questions, or would like information on sponsoring a Spark Summit, please contact organizers@spark-summit.org (mailto:organizers@spark-summit.org).