

Зал 1

Зал 2

Зал 3

State of Data 2025 от Программного комитета SmartData

Олег Кочергин Positive Technologies

Сергей Бойцов

10:00-10:45

10:45 - 11:15 **Перерыв**

Перспективы развития Apache Iceberg

Владимир Озеров CedrusData

Виталий Аксенов

11:15-12:00



Использование кодеков в ClickHouse: преимущества и недостатки

Анастасия Афанасьева GlowByte

Марк Паненко Ozon Банк

11:15-12:00

Как непростые времена вынудили нас построить лучший BI

Роман Назаренко Т-Банк

Екатерина Щербакова Т-Банк

11:15-12:00



12:00 - 12:30 **Перерыв**

Автоматизация конфигурирования ETL-процессов на основе Apache Spark 3, используя RAG и LLM MTC

Илья Кочагин MTC Web Services (MWS)

Сергей Бойцов

12:30-13:15



Как мы улучшили процессы по работе с данными в Airflow: практические кейсы

Дмитрий Морозов Инновационный центр «Безопасный транспорт»

Дмитрий Зуев Positive Technologies

12:30-13:15



Мониторинг DWH: от метаданных до DataOps

Данил Захаров Skyeng

Илья Аникин Авито

12:30-13:15



13:15-13:45 **Обед**

Spark — ВСЁ!

Евгений Гловов Navio

Олег Кочергин Positive Technologies

13:45-14:30



DataRentgen: как записать yet another lineage, не привлекая внимания санитаров

Максим Мартынов MTC Web Services (MWS)

Евгений Ненахов MTC Web Services (MWS)

13:45-14:30



ИИ под замком: как мы развернули безопасный LLM-сервис для 3000 разработчиков

Илья Дарковский Kaspersky

Андрей Кузнецов Positive Technologies

13:45-14:30



14:30-15:00 **Обед**

GP2S3 по-взрослому

Владимир Ермаков Т-Банк

Андрей Кошкин Т-Банк

15:00-15:45



Как в хранилище Яндекс Маркета начали писать документацию к объектам

Павел Колодкин Яндекс Маркет

Михаил Лукин Судо

15:00-15:45



Алгоритмы векторного поиска в современных базах данных

Александр Зевайкин YDB

Сергей Бойцов

15:00-15:45

15:45 - 16:15 **Перерыв**

Hadoop Is Not Dead — Just Secure!

Антон Александров Детский мир

Евгений Ненахов MTC Web Services (MWS)

16:15-17:00



Как X5 Tech дает аналитику по данным без привлечения аналитиков, специалистов и иных посредников

Владимир Ермаченков X5 Tech

Илья Аникин Авито

16:15-17:00

Алгоритмы векторного поиска в YDB

Александр Зевайкин YDB

Сергей Бойцов

16:15-17:00

17:00 - 17:30 **Перерыв**

Как мы строили lakehouse на Ozone

Виталий Моисеев Островок!

Михаил Марюфич Т-Банк

17:30-18:15

Хорошие данные не случаются случайно

Юрий Горынцева Arenadata Catalog

Виталий Аксенов

17:30-18:15

От бакета в S3 к Data Lakehouse: эволюция платформы данных в гонке за автономией

Никита Бандурко Navio

Георгий Попов Navio

17:30-18:15

18:15 - 22:00

Нетворкинг

Круглый стол «Hadoop мертв, да здравствует Hadoop?!»

Михаил Марюфич Т-Банк

Алексей Белозерский VK Cloud

Виталий Моисеев Островок!

Дмитрий Зуев Positive Technologies

18:45 - 19:45, Зал 2

Lightning Talks

Артем Дубинин VK / VK Tech

Дмитрий Швеенков VK

Михаил Лукин Судо

Бронислав Житников Positive Technologies

19:00 - 21:00, Зал 3



Доклады

Spark Connect: новый подход для работы с Apache Spark
 Александр Токарев Яндекс
 Сергей Бойцов

11:00-11:45

Доклады

Semantic RAG: аналитический подход к моделированию знаний для LLM
 Ольга Татаринова ЕрощВ
 Михаил Лукин Судо

11:00-11:45



Доклады

Как мы искали инструменты для DQ и к чему в итоге пришли
 Павел Павлюков Газпромбанк.Тех
 Александр Свяжин Газпромбанк.Тех

11:00-11:45



11:45 - 12:15 Перерыв

Как в Авито обеспечивают self-service разработки и деплоя витрин
 Айк Оганесян Авито
 Николай Огоров Авито

12:15-13:00



Debezium и PostgreSQL после happy-path: какие проблемы ждут в проде и как их решать
 Никита Рьянов
 Евгений Ненахов MTC Web Services (MWS)

12:15-13:00



Data Catalog: искажение метаданных или продуктовый подход
 Анна Мавлютова Т-Банк
 Бронислав Житников Positive Technologies

12:15-13:00



13:00-13:30 Обед

StarRocks — реальность современной платформы данных
 Станислав Лысков
 Олег Кочергин Positive Technologies

13:30-14:15



Как организовать масштабируемый ресерч-кластер для более чем 600 дата-сайентистов с помощью JupyterHub в Kubernetes
 Даниил Понизов Wildberries & Russ
 Владислав Печень Wildberries & Russ

13:30-14:15



AI-ассистенты в управлении данными
 Олег Сагитов Т-Банк
 Марк Паненко Ozon Банк

13:30-14:15



14:15-14:45 Обед

DataOps под микроскопом: CRD и Kubernetes-операторы для жизненного цикла «ETL-пробирок»
 Сергей Бойко Т-Банк
 Михаил Иванов Positive Technologies

14:45-15:30



Дата-контракты: ожидания от данных без иллюзий
 Валерия Торова Яндекс
 Илья Шишов Яндекс

14:45-15:30



Запускаем YugabyteDB в production
 Василий Осадчий 01.tech
 Виталий Аксенов

14:45-15:30



15:30 - 16:00 Перерыв

Сторонние движки исполнения для Apache Spark: опыт использования
 Никита Благодарный Честный знак
 Дмитрий Зуев Positive Technologies

16:00-16:45



DWN на StarRocks: год в production
 Артем Маркин Передовые Платежные Решения
 Евгений Ненахов MTC Web Services (MWS)

16:00-16:45



Data Quality as a Service — инструмент самообслуживания в большой компании
 Андрей Азеев MTC Web Services
 Богдан Петров MTC Web Services

16:00-16:45



16:45 - 17:15 Перерыв

Apache Spark SQL. Расширяй и управляй
 Дмитрий Вертлиб Честный знак
 Михаил Марюфич Т-Банк

17:15-18:00



Критерии хорошей платформы данных от Яндекса Доставки
 Владислав Гоцуляк Яндекс Доставка
 Марк Паненко Ozon Банк

17:15-18:00



Что такое metastore и с чем его едят
 Михаил Иванов Positive Technologies
 Виталий Моисеев Островок!

17:15-18:00



18:00 - 18:30 Перерыв

Искусство и кибернетика
 Дмитрий Булатов
 Виталий Аксенов

18:30-19:30

Закрытие конференции SmartData 2025
 Михаил Лукин Судо
 Бронислав Житников Positive Technologies

19:35-19:50

