

## 4 сентября

Время в программе указано для часового пояса UTC+3 (Europe/Moscow).

11:00-11:15	<a href="#">Открытие конференции SmartData 2024</a>		
	Михаил Лукин Андрей Дмитриев Судо JUG Ru Group		
11:15-12:00	<a href="#">Использование вероятностных структур данных для оптимизации ETL-процессов</a>	<a href="#">Шардированный не значит распределенный: что важно знать, когда PostgreSQL мало</a>	<a href="#">Как жить с data swamp</a>
	Дмитрий Вертлиб Честный знак	Евгений Иванов Олег Бондарь Яндекс Яндекс	Александр Боргардт bigdataduck
12:00-12:30	<a href="#">Интервью с Дмитрием Устюговым</a>		
	Дмитрий Устюгов Михаил Лукин NTU Singapore Судо		
12:30-13:15	<a href="#">Эффективные сервисы ML Inference нейросетей в Яндекс рекламе</a>	<a href="#">Airbyte. 2 года в продакшене</a>	<a href="#">Assessing Data Pipeline Quality &amp; Sanity with Data Angiograms</a>
	Дмитрий Ульянов Яндекс	Александра Попова СберЗдоровье	Sri Vishnu Chanderraju CueZen
13:15-13:45	<a href="#">Airbyte или NiFi: два подхода, два инструмента, одна цель?</a>		
	Антон Букреев Василий Сафронов Николай Акимов Бронислав Житников Cryptology Независимый эксперт TaskData Positive Technologies		
13:45-14:30	<a href="#">Как построить пайплайн RAG с использованием LlamaIndex</a>	<a href="#">Проблемы обработки Excel-файлов в Apache NiFi и как их решать</a>	<a href="#">От хайпа до продакшена: data mesh на Airflow + dbt</a>
	Алсу Нурутдинова Алина Кочева Positive Technologies Positive Technologies	Игорь Шемаров Axenix	Никита Юрасов Леонид Кожин Toloka Toloka
14:30-15:00	<a href="#">Альтернативные варианты восприятия информации в будущем программирования</a>		
	Павел Филонов Никита Соболев Николай Марков Независимый консультант wemake.services Altener		
15:00-15:20	<a href="#">Оптимизация распределения партиций в последовательности задач распределенной обработки данных</a>	<a href="#">Debezium Engine: практическое руководство по использованию</a>	
	Милена Булкина Т-Банк	Анастасия Сашина Т-Банк	
15:20-15:35	<a href="#">Подведение итогов онлайн-части конференции SmartData 2024</a>		
	Михаил Лукин Андрей Дмитриев Судо JUG Ru Group		

	Зал 1	Зал 2	Зал 3
10:15-10:30	<b><a href="#">Открытие офлайн-части конференции SmartData 2024</a></b> Владимир Озеров CedrusData   Олег Кочергин Positive Technologies		
10:45-11:30	<b><a href="#">Пишем свой cluster manager для Apache Spark</a></b> Александр Токарев Яндекс	<b><a href="#">Как мы сократили TTM создания дашбордов</a></b> Анар Багиров Авито	<b><a href="#">Как кролик съел зеленую сливу и не умер: сказ о миграции на Iceberg</a></b> Валентин Пановский BestDoctor
12:00-12:45	<b><a href="#">Как работает Apache Iceberg на примере Trino</a></b> Владимир Озеров CedrusData	<b><a href="#">Как мы Apache Kafka на Redpanda меняли</a></b> Роман Ананьев Авито	<b><a href="#">Data Vault 2.0. В каких случаях внедрять, разбор основных проблем применения методологии при построении DWH на Greenplum</a></b> Денис Лукьянов Ecom.tech
12:45-13:30	<b>Обед</b>		
13:30-14:15	<b><a href="#">Инструменты Data Quality: как, зачем, почему. Опыт Т-Банка</a></b> Дмитрий Руднев Т-Банк	<b><a href="#">dbt в деле: реальные кейсы и лайфхаки</a></b> Антон Александров Детский мир	<b><a href="#">Data sketches — быстро, дешево и (почти) точно!</a></b> Сергей Жемжицкий Arenadata
14:15-15:00	<b>Обед</b>		
15:00-15:45	<b><a href="#">Data mesh: ожидания vs реальность</a></b> Петр Гуринов Лемана ПРО (Леруа Мерлен)	<b><a href="#">CI/CD в большом on-premise Datalake-- проекте</a></b> Александра Чекмарева (Китченко) Честный знак Никита Благодарный Честный знак	<b><a href="#">Как мы делаем облачный Greenplum</a></b> Леонид Борчук Yandex Cloud
16:15-17:00	<b><a href="#">Каждый байт на вес золота. Опыт построения DMP в рекламе Яндекса</a></b> Алексей Стыценко Яндекс	<b><a href="#">Data Lineage: как настроить в зоопарке технологий и зачем это нужно</a></b> Булат Усманов Кулер (ex-СберМаркет)	<b><a href="#">Выбрасываем Java икратно ускоряем Spark/Presto... Или пока нет?</a></b> Павел Солодовников CedrusData
17:30-18:15	<b><a href="#">Как быстро запустить процесс ведения каталога данных в компании на примере DataHub</a></b> Наталья Журавлева Ozon	<b><a href="#">Schema Registry: Ultimate Guide</a></b> Тимофей Брунько Yandex Cloud	<b><a href="#">Snowplow: утерянное руководство</a></b> Олег Кочергин Positive Technologies

18:15-22:00

Нетворкинг

18:30-19:30 Зал 2

[Круглый стол: Data Mesh](#)

18:45-20:45 Зал 1

[Lightning Talks](#)

	Зал 1	Зал 2	Зал 3
11:00-11:45	<p><a href="#">Apache Arrow: быстрее, ниже, сложнее</a></p> <p>Евгений Готов Autotech</p>	<p><a href="#">Как мы тестировали 5 способов загрузки данных в Greenplum и что из этого вышло</a></p> <p>Татьяна Дидова АЭРО</p>	<p><a href="#">NiFi. Пишем код для codeless-системы</a></p> <p>Бронислав Житников Positive Technologies</p>
12:15-13:00	<p><a href="#">Как мы строим систему распределенного трейсинга, в которой можно терять данные</a></p> <p>Игорь Балюк Авито</p>	<p><a href="#">One More Way to Make Backup in Ignite</a></p> <p>Николай Ижиков Apache Software Foundation</p>	<p><a href="#">От ручного труда к автоматической генерации проверок качества данных</a></p> <p>Александр Мадумаров Инновационный центр «Безопасный транспорт» ГКУ ЦОДД</p>
13:00-13:45	Обед		
13:45-14:30	<p><a href="#">CI/CD для большого хранилища данных</a></p> <p>Арсений Бирюков Т-Банк</p>	<p><a href="#">Рецепт платформы потоковой обработки данных на Apache Flink</a></p> <p>Данил Сабиров Яндекс.Такси Технологии</p>	<p><a href="#">Обработка событий в Snowplow: от сбора до аналитики</a></p> <p>Мария Бученкова Детский мир</p> <p>Владимир Седелников Детский мир</p>
14:45-15:15	<p><a href="#">CDC от источника до хранилища: как в банке Синара построили CDC с применением продуктов Arenadata</a></p> <p>Иван Клименко Arenadata</p>		
15:15-16:00	Обед		
16:00-16:45	<p><a href="#">Как навести порядок в двух эксбайтах данных?</a></p> <p>Максим Гудзикович Яндекс</p> <p>Кирилл Осинцев Яндекс</p>	<p><a href="#">Автогенерация синтетических данных с использованием алгоритмов ML</a></p> <p>Данила Самошкин Neoflex</p> <p>Анастасия Коткова Neoflex</p>	<p><a href="#">Оптимизации сериализатора ВКонтакте</a></p> <p>Илья Асадуллин VK</p> <p>Илья Кокорин VK</p>
17:15-18:00	<p><a href="#">Оркестратор пайплайнов для небольшой команды инженеров и аналитиков: как мы используем Dagster</a></p> <p>Алексей Завальский Bestplace</p> <p>Дмитрий Крылов Bestplace</p>	<p><a href="#">The State of Data, RU Edition</a></p> <p>Олег Кочергин Positive Technologies</p>	<p><a href="#">Storage для lake</a></p> <p>Александр Казанский Т-Банк</p>
18:30-19:15	<p><a href="#">Это реальный космос: space-индустрия сегодня</a></p> <p>Яна Харлан Бюро 1440</p>		
19:25-19:40	<p><a href="#">Заккрытие конференции SmartData 2024</a></p> <p>Владимир Озеров CedrusData</p> <p>Олег Кочергин Positive Technologies</p>		