

Как меняется рынок 2025 в РФ и мире

Пчелин Юрий

Немного об эволюции развития больших данных

Эра хранилищ данных (Data Warehouse)

Период:

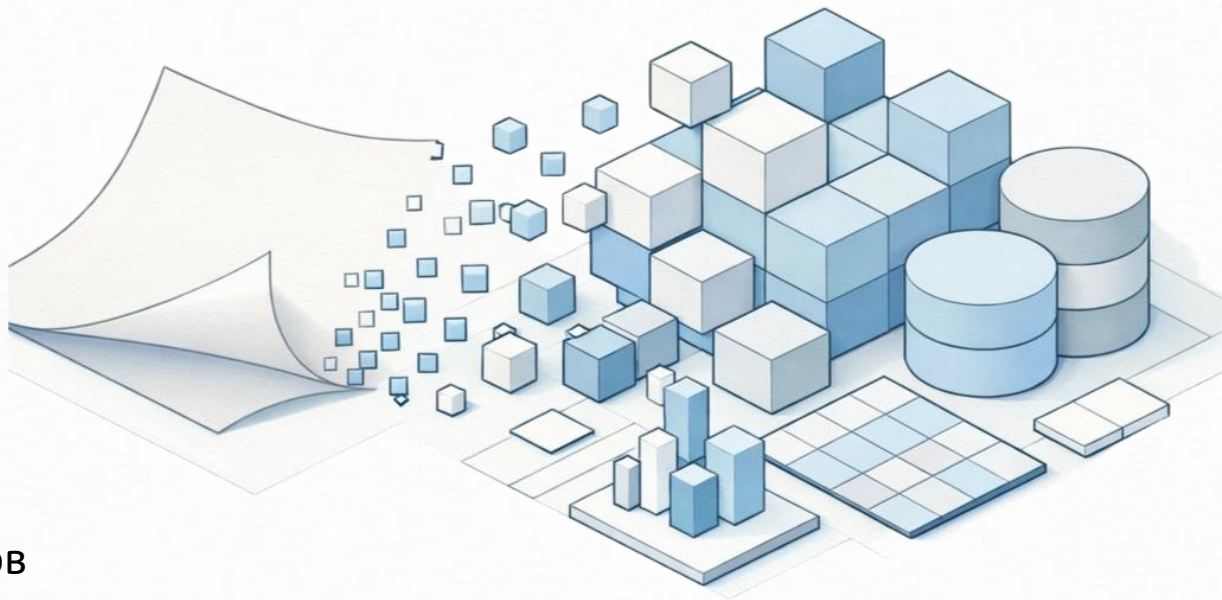
1995–2005

Ключевые игроки:

Oracle, IBM, Microsoft

Решаемые задачи:

- Финансовая консолидация,
- Регуляторная отчётность,
- План-факт анализ,
- Контроль складских остатков



Рождение Big Data

Период:

2005–2015

Ключевые игроки:

Hadoop, Apache Software Foundation

Решаемые задачи:

- Анализ clickstream в e-commerce
- Рекомендательные системы
- Fraud detection в банках
- Телеком-аналитика churn

Старт Machine Learning



Облака и Data-as-a-Service

Период:

2015–2020

Ключевые игроки:

Amazon Web Services, Google Cloud,
Microsoft Azure

Решаемые задачи:

- Облако – место экспериментов, A/B тесты
- Real-time аналитика

Характерные примеры:

- Netflix рекомендации, Uber динамическое ценообразование, Spotify



Автоматизация решений

Период:

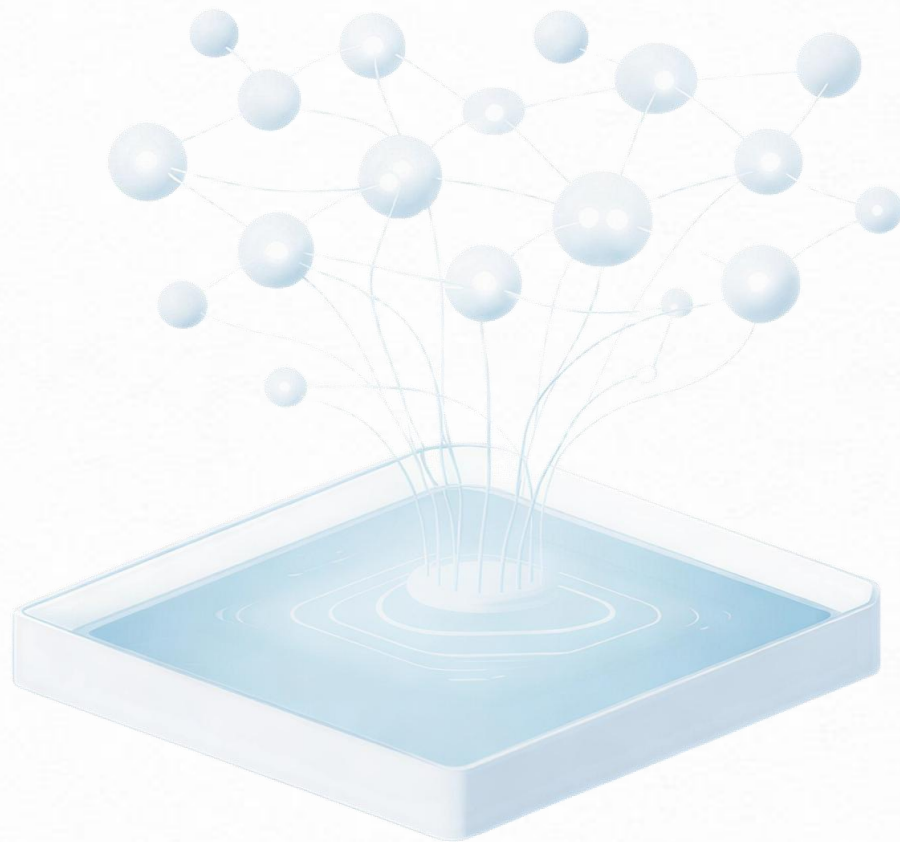
2020—н.в.

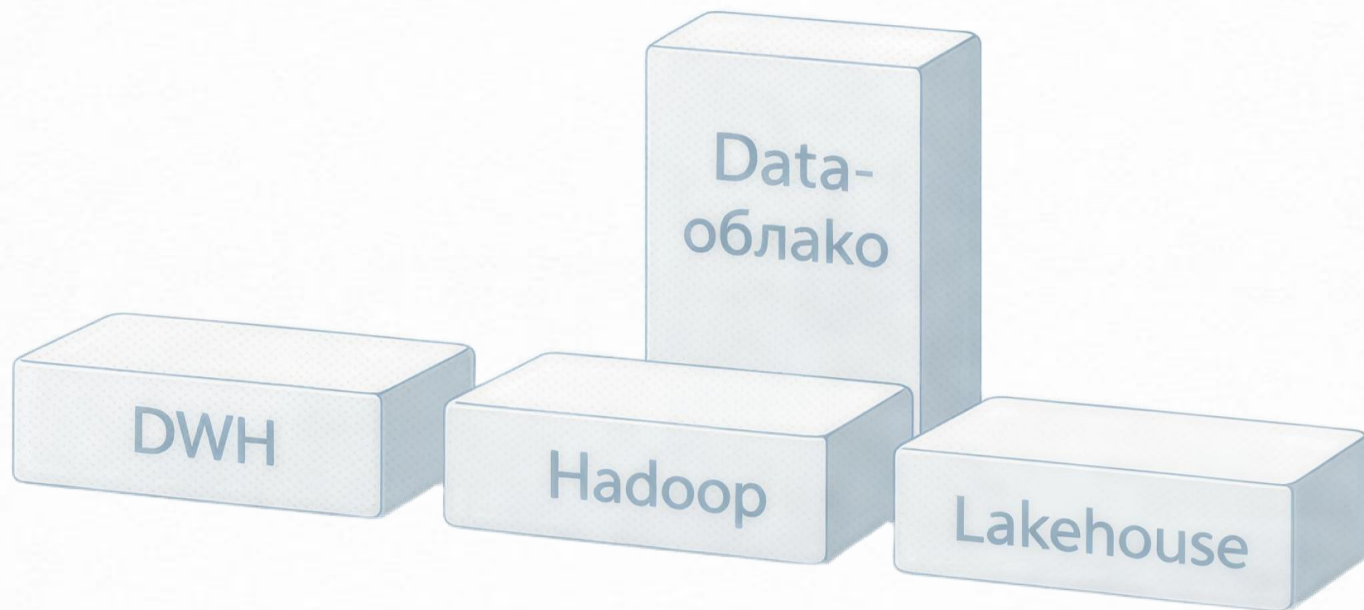
Ключевые игроки:

Lakehouse, Databricks, AI

Решаемые задачи:

- Дешевые как Hadoop,
- Управляемые как DWH,
- ML-нативны





Что изменилось на рынке в 2025 году

В Море:

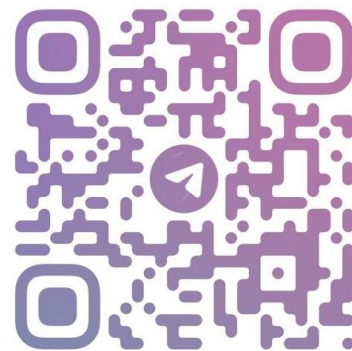
- Открытие исходного кода Unity Catalog от Databricks и поддержка Apache Iceberg всеми крупными игроками (Snowflake, AWS, Google)
- Появление "агентов-аналитиков" для сложных SQL-запросов и визуализации
- Внедрение интерфейсов на естественном языке (NLQ)
- AI Агенты, NotebookML

РФ:

- Массовое импортозамещение стека Hadoop
- Первые российские Lakehouse-платформы
- Фокус на безопасности озер данных | Применение синтетических данных
- Яндекс выложил в открытый доступ один из крупнейших мировых датасетов «Yambda» для рекомендательных систем

Обо мне

- Директор ИТ-Сервисов и отказоустойчивости
- @Дикси
- ex CDTO | CPO | Head of Support
- @ X5, HomeCredit Bank, SberTech, СУЭК
- Участник сообщества GetMentors
- Увлекаюсь бегом, архитектурой и музыкой



@YPCHELIN