

Теоретический курс для Архитекторов Модели Данных

Длительность: 24 ак. часа

Необходимая предварительная подготовка

- Знание основ Hadoop, DWH, СУБД (желателен)

Программа курса

1. Основные концепции и понятия теории хранилищ данных

- Классическое реляционное хранилище
 - ✓ Слои данных (Stage, ODS, DDS, витрины)
 - ✓ Методики моделирования: подходы Р. Кимбалла и Б. Инмона, Data Vault, Anchor modeling
- Озеро данных
 - ✓ Schema-on-write vs Schema-on-read
 - ✓ Виртуализация данных (Data Virtualization)
- Big Data и NoSQL технологии
- ETL\ELT процессы
- Лямбда- и Каппа-архитектуры

2. Планирование хранилища данных

- Определение используемого подхода и техники моделирования
- Проектирование модели данных
- Определение стека используемых технологий
- Планирование жизненного цикла и развития хранилища
- Разработка проектной документации

3. Имплементация хранилища данных

- Организация контуров разработки, тестирования и UAT
- Разработка управляющих потоков
- Разработка процессов загрузки и обработки данных
- Создание справочников и общих измерений
- Миграция исторических данных

4. Эксплуатация хранилища данных

- Разработка эксплуатационной документации
- Регламентные процессы обслуживания
- Обеспечение отказо- и катастрофоустойчивости
- Организация технической поддержки

5. Смежные вопросы

- Управление данными (Data Governance)
- Качество данных (Data Quality)
- Информационная безопасность