

## Программа курса «ARMG: Архитектура Данных»

**О курсе:** Курс для ИТ-архитекторов и специалистов по проектированию и разработке архитектуры данных (Big Data), планированию инфраструктуры озер данных (Data Lakes) и проектов Big Data в компаниях.

### Аудитория:

- ИТ-архитекторы,
- системные аналитики и разработчики,
- руководители Big Data-проектов,
- команды в области аналитики,
- специалисты направления Data Governance,
- ИТ-менеджеры,
- руководители проектов по цифровизации.

**Уровень подготовки:** предварительный опыт не требуется.

**Продолжительность курса:** 24 академических часа, 6 дней по 4 ак. часов дистанционно

## Содержание программы

### 1. Архитектура данных на уровне предприятия

- **Data Driven подход.** Успехи. Проблемы. Эволюция. Внедрение Data Lake: что может пойти не так?
- Какие проблемы мы не можем решить на уровне DataLake/DWH.
- Задачи **Data Governance**.
- **DAMA DMBOK.** Стандартизация подхода к управлению данными на уровне предприятия. Data Management Principles. Data Management Framework.
  - DMBOK: Data Architecture.
  - DMBOK: Enterprise Data Model (EDM).
  - DMBOK: Data Flow Design
  - DMBOK: Data Architecture implementing. Risk Assessment. Cultural Change
- **Application Centric vs Data Centric.** Когда, внедряя Big Data, мы приближаемся к Data Centric, а когда отдаляемся?

### 2. Основы проектирования моделей данных

- **Data Modeling.** Введение
- Уровни проектирования моделей данных
- Концептуальное проектирование
- Классическая ER-Model и ее расширения
- Модель данных приложения (логический и физический уровни проектирования)
- Эффективное использование Data Modeling Tool
- **Live Demo**
- Реляционная модель данных. Нормализация и денормализация
- NoSQL Introduction
- Работа с референсными моделями данных (на примере банковской области)

### 3. Проектирование моделей данных для аналитических систем

- Концептуальный дизайн архитектуры
- Обобщенная схема архитектуры Big Data-решений. Требования к компонентам и моделям данных.
- Управление временем в аналитических системах.

- Паттерны проектирования моделей данных
- **Live Demo**
- Подходы к проектированию ядра
- CIF, Anchor modeling, Data Vault. Преимущества и ограничения. Кастомизация.
- Итеративное развитие модели данных ядра
- Моделирование аналитических витрин. Подход Р. Кимбалла и его развитие.
- **BEAM** (Business Event Activity Modeling) - подход к проектированию моделей для аналитических задач
- **Примеры и кейсы**