

Система обучения Arenadata

Федор Кирдяшов
Директор по обучению



О себе



Спикер

Федор Кирдяшов

Директор по обучению ООО «Аренадата Софтвер»,
кандидат технических наук

Эксперт в обучении и развитии персонала

Более 20 лет занимаюсь подготовкой кадров в
отрасли IT

Опыт построения корпоративных академий,
внутреннего и внешнего обучения

Автор методик по построению систем обучения



Группа Arenadata –
ведущий российский
разработчик на рынке
систем управления и
обработки данных

Проект «Кадры для экономики данных»

Цель проекта:

построение системы подготовки IT-специалистов в части СУБД (системы управления базами данных) и сопутствующих сервисов для работы с большими данными

Наш успех – доверие клиентов

150+ клиентов размещают свои хранилища, витрины, озера данных на продуктах Arenadata

Банки и страховые: **40+** клиентов



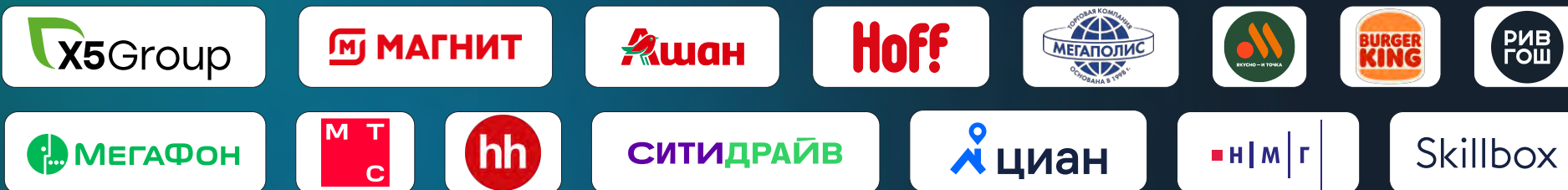
Промышленность, ТЭК, Транспорт: **40+** клиентов



Госсектор: **30+** клиентов



Retail, FMCG, Фарма, Telecom, Digital, Media: **30+** клиентов



Роль CLO

Связь обучения с бизнесом

- Понимание потребностей бизнеса
- Создание и реализация стратегии по обучению
- Сотрудничество с бизнес-лидерами
- Продвижение непрерывного обучения
- Работа с внешними поставщиками обучения

Система обучения и её атрибуты



Вопрос 1

Кого учить?



Категории специалистов



Администраторы и специалисты:

- Администраторы баз данных
- Специалисты сопровождения СУБД



Аналитики:

- Аналитики данных
- Бизнес-аналитики
- Системные аналитики



Архитекторы и исследователи:

- Архитекторы данных
- Архитекторы СУБД
- Исследователи данных



Инженеры:

- Инженеры данных
- Инженеры внедрения СУБД
- DevOps-инженеры



Разработчики:

- Разработчики баз данных
- Разработчики КХД
- ETL-разработчики



Руководители:

- Руководители проектов
- Руководители IT подразделений

Вопрос 2

Чему учить?



Источники данных

Держатели данных в индустрии:

- Сотрудники HR подразделений
- Руководители команд
- Эксперты

Открытые источники:

Онлайн ресурсы для рекрутинга:

- Headhunter
- Superjob
- и другие

Продукты и технологии:

- Дата-стратегия
- Функционал и возможности ПО
- Бизнес-процессы
- Запросы от клиентов

Глоссарий

Обучение

Процесс получения знаний, умений, навыков для формирования конкретной компетенции

Демонстрация

Процесс передачи информации без подтверждения степени усвоения (показ функционала продукта)

Компетенция

Совокупность знаний, умений, навыков, необходимых для выполнения трудовой функции

Знания

Усвоенная теория, с подтверждением степени усвоения

Умения

Способность применять на практике знания

Навыки

Наработанные на практике умения, позволяющие выполнять сложные действия в ограниченное время

Профиль должности

Трудовая функция 1	Компетенция 1	Навык 1.1	Умение 1.1	Знание 1.1.1
				Знание 1.1.2
				Знание 1.1.n
	Компетенция 2	Навык 1.2	Умение 1.2	Знание 1.2.1
				Знание 1.2.2
	Компетенция 2	Навык 2.1	Умение 2.1	Знание 2.1.1
				Знание 2.1.2
Трудовая функция 2	Компетенция 1	Навык 2.1	Умение 2.1	Знание 2.1.n
	Компетенция 2	Навык 2.2	Умение 2.2	Знание 2.2.1
				Знание 2.2.2
	Компетенция 3	Навык 3.1	Умение 3.1	Знание 3.1.1
				Знание 3.1.2
Трудовая функция 3	Компетенция 1	Навык 3.1	Умение 3.1	Знание 3.1.n
	Компетенция 2	Навык 3.2	Умение 3.2	Знание 3.2.1
				Знание 3.2.2
	Компетенция 3	Навык 3.3	Умение 3.3	Знание 3.3.1
				Знание 3.3.2

Профиль должности (простой вариант)

Трудовая функция 1	Навык 1	Знание 1
Компетенция 1		Знание 2
		Знание 3
	Навык 2	Знание 4
		Знание 5
	Навык 3	Знание 6
		Знание 7
		Знание 8
	Навык n	Знание 9
		Знание n

Профиль администратора СУБД

Компетенции	Навыки	Знания	
Развертывание систем управления базами данных	Создание проекта развертывания и выполнение развертывания	Особенности подключения к ADB (Greenplum/Greengage) средствами psql Модель распределения ресурсов	Основные параметры конфигурирования СУБД (ADB/Greenplum/Greengage)
Настройка резервного копирования	Восстановление данных с использованием журналов транзакций	Варианты и средства резервного копирования	Принципы отказоустойчивой работы ADB (Greenplum/Greengage)
Настройка мониторинга работы и оптимизация	Контроль использования предоставленных разрешений, аудит Оптимизация работы СУБД	Цели и методы мониторинга запросов Механизмы обеспечения параллельной работы транзакций	Основы оптимизации запросов Принципы работы статистики оптимизатора

Профиль разработчика БД

Компетенции	Навыки	Знания	
Проектирование баз данных	Выбор типа баз данных для решения задачи	Сущности и атрибуты в ADB (Greenplum/Greengage)	Нормализация таблиц и нормальная форма
	Проектирование схем данных	Типы данных	Индексы
	Настройка ресурсов групп	Особенностей ролевой модели и работы с объектными правами	Принципы проектирования и настройка схем данных
Работа с данными в таблицах и с транзакциями		Механизмы для управления и распределения ресурсов между пользователями или запросами	
	Объединение таблиц	Основы массовой загрузки и выгрузки данных	Основы блокировки и их мониторинг, оптимизация
	Сортировка, группировка и фильтрация данных.	Особенности сортировки, группировки и фильтрации данных	Базовый синтаксис изменения данных
Работа с процедурными объектами в базах данных	Изменение данных в таблицах и работа с транзакциями	Знание основных операций с партициями (Rename, Split, Exchange Partiton)	Транзакции в базе данных
	Применение пользовательских процедур, функций и триггеров	Языки для написания функций	
		Курсоры и табличные функции	

Вопрос 3

Как правильно
выбрать курсы?



Технологии, в рамках которых реализуется подготовка специалистов

Greengage

Greenplum

PostgreSQL

Hadoop

Clickhouse

Kafka

NiFi

Data Governance &
Data Quality

В линейке присутствуют как курсы по отдельным продуктам, так и индивидуальные программы для крупных заказчиков, а также совместные треки обучения с партнёрами Группы Arenadata



Учебные курсы

Курсы по Greenplum/Greengage:

- Эксплуатация Arenadata DB
- Arenadata DB для разработчиков
- Arenadata DB для аналитиков
- Введение в Arenadata DB

Курсы по PostgreSQL:

- Основы Arenadata Prosperity
- Arenadata Prosperity для аналитиков данных

Курсы по ClickHouse:

- Arenadata QuickMarts для разработчиков
- Эксплуатация Arenadata QuickMarts
- Arenadata QuickMarts для пользователей

Курсы по Hadoop:

- Эксплуатация Arenadata Hyperwave
- Управление информационной безопасностью кластера Arenadata Hyperwave
- Основы работы с Apache Spark в экосистеме Arenadata Hyperwave
- Основы Trino

Совместные курсы с партнерами:

- Курс «Стюард данных Arenadata Catalog» в партнерстве с IM Center (ООО «ИМ ЦЕНТР»)
- Совместный курс с ООО «Нетология» «Аналитик данных для атомной отрасли»
- Совместный курс с ООО «Скала-Р» - «Arenadata DB и Машина больших данных Скала^р МБД.Г

Курсы по Data Governance & Data Quality:

- Arenadata Catalog. Специалист
- Arenadata Catalog. Администратор

Индивидуальные курсы для крупных заказчиков:

- Курс для ПАО «Сбербанк» – «Greenplum для администраторов»
- Курс для ПАО «Мегафон» – «Arenadata QuickMarts (Clickhouse) для аналитиков»

Курсы по Apache Kafka:

- Эксплуатация Arenadata Streaming» (Kafka)
- Основы работы с Apache Kafka на базе платформы Arenadata Streaming

Курсы по Apache NiFi:

- Эксплуатация Arenadata Streaming (NiFi)

Общие курсы:

- Архитектура платформы данных

Карта курса Arenadata DB (Greenplum/Greengage) для разработчиков

№п/п	Формируемые знания и навыки	
	Навыки	Знания
1	Использование основных подходов к построению хранилища данных	Понимание процессов ETL и ELT
		Знание основных принципов построения хранилищ
		Знание особенностей настройки PXF JDBC
		Понимание процесса обновления справочников и таблиц
2	Использование ролевой модели	Понимание особенностей ролевой модели
		Знание деталей работы с объектными правами
3	Настройка ресурсных групп	Понимание работы ресурсных групп
4	Установка расширения Diskquota	Знание особенностей расширения Diskquota
5	Создание простых таблиц (с различными настройками сжатия и распределения)	Знание особенностей работы в ADB ограничений, View и Materialized View
		Понимание принципов распределения данных

Карта курса Arenadata DB (Greenplum/Greengage) для разработчиков

№п/п	Формируемые знания и навыки	
	Навыки	Знания
6	Создание партиций	Понимание механизмов партиционирования
		Знание основных операций с партициями (Rename, Split, Exchange Partititon)
		Понимание ограничений партиционирования
7	Использование типов данных (включая JSON и XML)	Знание основных типов данных
		Понимание процессов работы с различными типами данных
		Знание особенностей библиотек JSON и XML
8	Оптимизация SQL запросов	Понимание процесса выполнения запроса
		Знание о плане запросов и команде Explain
		Понимание основ оптимизации запросов
9	Мониторинг запросов с помощью ADB Control	Понимание функциональности ADB Control
		Знание особенностей мониторинга запросов
		Знание ролевой модели ADB Control
10	Сбор статистики оптимизатора	Понимание работы статистики оптимизатора
11	Применение различных типов индексов в ADB	Знание возможностей различных индексов

Карта курса Arenadata DB (Greenplum/Greengage) для разработчиков

№п/п	Формируемые знания и навыки	
	Навыки	Знания
12	Управление блокировками	Знание уровней блокировок и их мониторинга
		Понимание особенностей Deadlock
13	Управление транзакциями	Понимание механизма транзакций
		Знание уровней изоляции
		Понимание особенностей MVCC в ADB
		Понимание особенностей команды Vacuum
14	Загрузка и выгрузка данных с помощью внешних таблиц	Понимание механизма внешних таблиц
		Знание особенностей работы утилит GPFDIST и GPLOAD
		Понимание архитектуры PXF (включая Pushdown и User Impersonation)
		Знание настроек PXF для Hadoop и JDBC
		Понимание использования коннекторов ADB Enterprise Edition
15	Загрузка и выгрузка данных с помощью команды COPY	Знание синтаксиса команды COPY
		Понимание работы утилит ADB/Greenplum/Greengage

Карта курса Arenadata DB (Greenplum/Greengage) для разработчиков

№п/п	Формируемые знания и навыки	
	Навыки	Знания
16	Создание пользовательских функций	Понимание концепции хранимого кода
		Понимание уровней волатильности функций и уровней доверенности программных языков
		Знание особенностей работы функций в ADB
17	Использование оконных функций, агрегатов, расширений MADlib и PostGIS	Знание синтаксиса оконных функций и пользовательских агрегатов
		Понимание работы расширений MADlib и PostGIS
18	Использование основных приемов миграции в ADB	Знание возможностей расширения
		Знание особенностей миграции в ADB из Oracle Database

Категории курсов

Для пользователей

Введение в ADB

ADC:
Специалист

Архитектура
платформы
данных

Для аналитиков и специалистов

ADB для
аналитиков

Эксплуатация
ADH

ADQM для
аналитиков

Основы Trino

*ADP для
аналитиков*

Основы ADP

Для администраторов и инженеров

Эксплуатация
ADB

Эксплуатация
ADH

*Эксплуатация
ADS (NiFi)*

Управление ИБ
кластера ADH

Эксплуатация
ADQM

ADC:
Администратор

Для разработчиков и архитекторов

ADB для
разработчиков

*ADQM для
разработчиков*

Основы работы с
Apache Spark в
экосистеме ADH

Основы работы с
Apache Kafka на
базе платформы
ADS

Система грейдов Arenadata

Статус	Пользователь	Специалист	Профессионал	Мастер	Эксперт
Уровень грейда	1-3	4-5	6	7-9	10-11
Уровень курсов	От 1 курса для пользователя	От 1 курса для специалиста	От 1 курса для администратора или разработчика	От 2 курсов для администратора и разработчика	От 4 курсов для администратора и разработчика

Основные принципы

1

Профиль роли

Компетенция 1

Навык 1

Умение 1

Знание 1

Умение 2

Знание 2

Навык 2

Умение 3

Знание 3

Оценка знаний и навыков

2

Профиль
специалиста

Компетенция 1

Навык 1

Умение 1

Знание 1

Умение 2

Знание 2

Грейд специалиста

Навык 2

Умение 3

Знание 3

Подбор курса

3

Карта курса

Навык 2

Умение 2

Знание 2

Умение 3

Знание 3

Проект «Кадры для экономики данных» и партнёры


Участвуем в эксперименте Минцифры России по предоставлению разработчикам ПО возможности добровольного подтверждения компетенций, в соответствии с постановлением Правительства РФ от 11 февраля 2025 г. № 130

Являемся партнёром Ассоциации предприятий компьютерных и информационных технологий (АПКИТ) по построению Системы добровольной сертификации ИТ-специалистов

Входим в профессиональное сообщество «Образовательная Лига вендоров» (совместно с ПАО «Газпром нефть»)

Партнёры со стороны университетов и центров ДПО

Подготовка кадров реализуется как самостоятельно, так и совместно с образовательными партнёрами на различных уровнях



01

Уровень высшего образования:
обучение встраивается в дисциплины

02

Уровень ДПО*:
курсы для специалистов

* Дополнительное профессиональное образование

Образовательные партнеры

01 Университеты

- Алтайский государственный университет
- Владимирский государственный университет
- Высшая школа экономики
- Воронежский институт высоких технологий
- Донской государственный технический университет.
- Магнитогорский государственный технический университет им. Г.И. Носова
- Московский государственный технический университет им. Н.Э. Баумана
- Московский технический университет связи и информатики
- Новосибирский государственный технический университет
- Российский государственный аграрный университет — МСХА им. К. А. Тимирязева
- Московский государственный педагогический университет
- Российский государственный университет туризма и сервиса
- Российский технологический университет МИРЭА
- Российский экономический университет имени Г. В. Плеханова
- Санкт-Петербургский государственный университет
- Университет ИТМО
- Университет науки и технологий МИСИС

02 Организации ДПО

Учебные центры:

- АНО ДПО «СофтЛайн Эдюкейшн»
- Учебный центр «Сетевая Академия»
- Учебный центр «СИБИНФОЦЕНТР»

Онлайн-площадки:

- ООО «ИМ ЦЕНТР»
- ООО «Нетология»
- ООО «Отус онлайн-образование»

Программы подготовки и мероприятия

01 Программы подготовки

Направления подготовки:

- Фундаментальная информатика и информационные технологии
- Программная инженерия
- Прикладная математика и информатика
- Информационные системы и технологии
- Математика и компьютерные науки
- Прикладная математика
- Информатика и вычислительная техника
- Бизнес-информатика

Дисциплины:

- Базы данных
- Большие данные
- Визуализация данных
- Интеллектуальный анализ больших данных
- Информационно-аналитические системы
- Обработка больших данных
- Проектирование баз данных
- Проектирование информационно-аналитических систем
- Проектирование систем обработки больших данных
- Разработка программно-информационных систем
- Технологии и инструментарий анализа больших данных
- Управление и автоматизация БД
- Хранение и обработка больших объёмов данных

02 Мероприятия

Мероприятия для преподавателей:

- Обучение преподавателей университетов

Мероприятия для студентов:

- Хакатоны и конкурсы
- Стипендии для студентов
- Стажировки
- Лекции и практика для студентов

Обучение вместе с Arenadata

Федор Кирдяшов

Директор по обучению

education@arenadata.io

<https://arenadata.tech/>

