

ФАБРИКА ДАННЫХ

Digital Q.DataFactory

СДЕЛАЙТЕ ДАННЫЕ ПОЛЕЗНЫМИ С ФАБРИКОЙ ДАННЫХ

DIGITAL Q.DATAFACTORY – решение для работы с данными, автоматизирующее весь процесс обработки – от загрузки до визуализации

СОСТАВ РЕШЕНИЯ



Компоненты для разработки
аналитических приложений:
загрузка, контроль качества и
визуализации



Набор инфраструктурных
компонентов
обеспечивающие
выполнение процессов
фабрики данных



Единый портал данных для
доступа пользователей к
данным

ПРИМЕНЕНИЕ ПЛАТФОРМЫ



Создание и управление хранилищами и озерами данных



Создание и разработка дашбордов и ситуационных центров



Автоматизация подготовки любых видов отчетности, в том числе оперативной



Внедрение Data Governance процессов в организации

ВОЗМОЖНОСТИ ПЛАТФОРМЫ



РАБОТА С СОВРЕМЕННЫМИ АРХИТЕКТУРАМИ ДАННЫХ DATA WAREHOUSE и LAKEHOUSE

Платформа позволяет создавать аналитические приложения для различных современных архитектурных парадигм хранения и обработки данных.



IN-MEMORY ОБРАБОТКА ДАННЫХ

Встроенные в платформу обработка данных в оперативной памяти позволяет создавать конвейеры обработки данных в близком к реальному времени режиме



LOW-CODE ВОЗМОЖНОСТИ ДЛЯ ОБРАБОТКИ ДАННЫХ

Low-code возможности фреймворка построения конвейеров обработки данных позволяют ускорить разработку и снизить уровень требований к квалификации дата-инженеров для разработки на платформе



ПОДХОДИТ ДЛЯ РАЗЛИЧНЫХ БИЗНЕС ЗАДАЧ

Все компоненты платформы взаимодействуют и позволяют настроить конвейер обработки данных для разных бизнес задач: оперативная онлайн отчетность, построение дашбордов, построение хранилища данных



DATA MESH ПОДХОД ИЗ КОРОБКИ

Платформа позволяет создавать аналитические приложения, как отдельные приложения и обеспечивает единые принципы взаимодействия и обмена данными между приложениями.



ПОЛЬЗОВАТЕЛЬСКИЙ ПОРТАЛ

Централизованное единое место доступа к данным и отчетам для конечных пользователей. С настройкой разделения прав доступа к данным.

ФУНКЦИОНАЛЬНАЯ АРХИТЕКТУРА ФАБРИКИ ДАННЫХ

ТЕХНОЛОГИЧЕСКИЕ ПЛАТФОРМЫ

РАБОТА ПОЛЬЗОВАТЕЛЕЙ

ЕДИНЫЙ ПОРТАЛ ДАННЫХ

УПРАВЛЕНИЕ ДОСТУПОМ

ПЛАТФОРМА ДЛЯ РАЗРАБОТКИ АНАЛИТИЧЕСКИХ ПРИЛОЖЕНИЙ

DIGITAL Q.DATAFLOWS:

DATA CATALOG

DATA STREAMER

DATA QUALITY

DIGITAL Q.SENSOR

DIGITAL Q.REPORT ENGINE

ИНФРАСТРУКТУРНЫЕ ПЛАТФОРМЫ

ИНТЕГРАЦИЯ ДАННЫХ

DIGITAL Q.NIFI

ОБРАБОТКА И ХРАНЕНИЯ ДАННЫХ

СУБД DIGITAL Q.DATABASE:

РЕЛЯЦИОННАЯ

IN-MEMORY

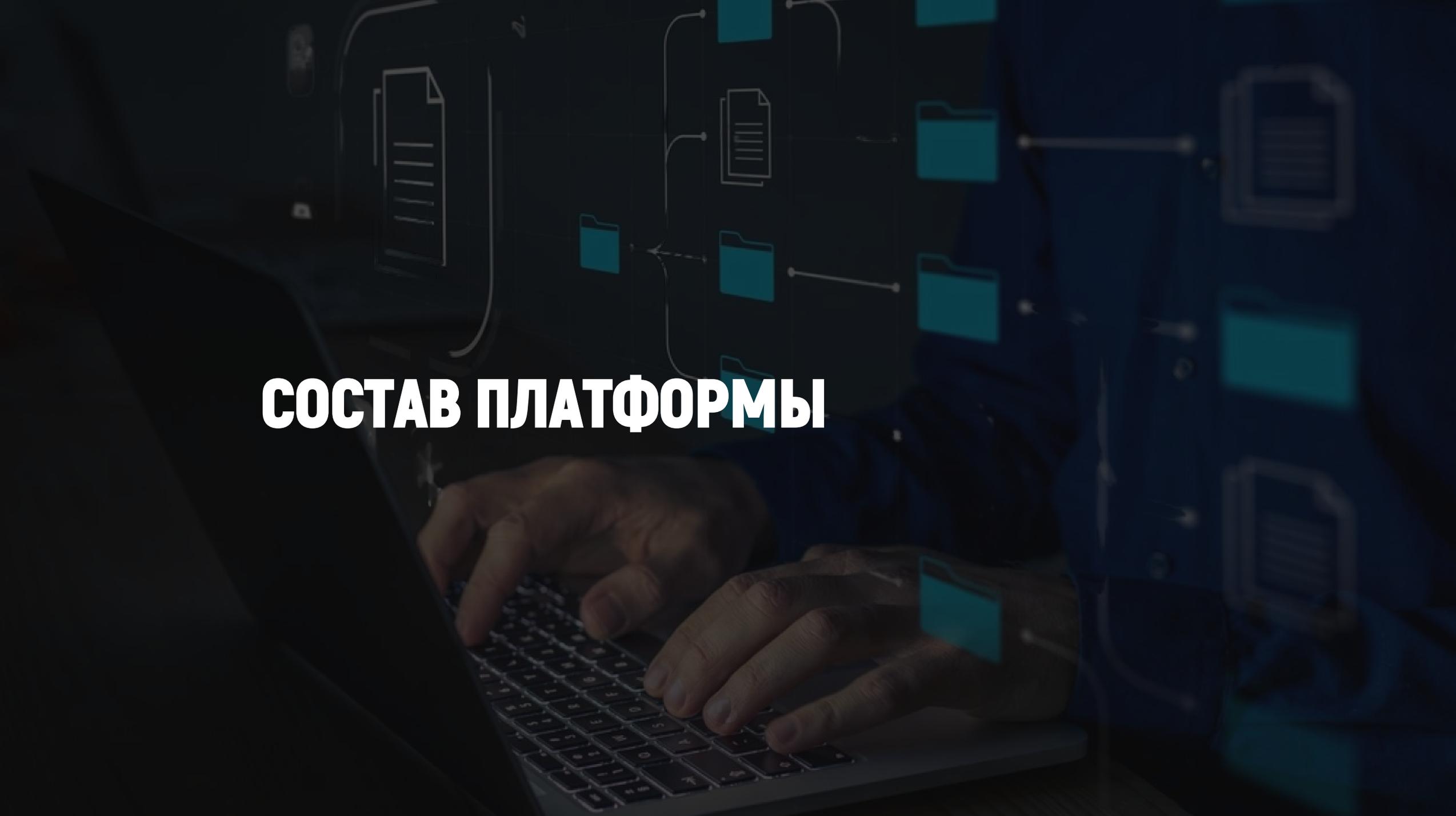
АНАЛИТИЧЕСКАЯ

РАСПРЕДЕЛЕННАЯ
ОБРАБОТКА ДАННЫХ:

АРАСНЕ HADOOP

S3 OBJECT STORE

АРАСНЕ SPARK

The image features a dark, moody background with a person's hands typing on a laptop keyboard. Overlaid on the scene is a digital file system interface, consisting of various icons for folders and documents connected by lines, suggesting a complex network or data structure. The overall aesthetic is technical and modern.

СОСТАВ ПЛАТФОРМЫ

ЕДИНЫЙ ПОРТАЛ ДАННЫХ

- ✓ Специализированное рабочее место для работы с данными: отчётностью и дашбордами организации.
- ✓ Управление доступом к данным пользователей организации.
- ✓ Мониторинг и аналитика по использованию данных в организации.



КАТАЛОГ ДАННЫХ

НАЗНАЧЕНИЕ



Автоматизирует управление метаданными из любой информационной системы в едином месте, предоставляя пользователям полный контекст для аналитики данных и отчетности.



Собирает информацию из всех источников данных организации: учетные системы, файлы, хранилище, отчетность.



Визуализирует путь данных на детальном уровне – от источников до конечных витрин и отчетов.



Собирает информацию о потоках обработки данных.

DATA STREAMER

Позволяет создавать процессы извлечения и преобразования данных в едином интерфейсе с использованием low code возможностей.

Генерирует код для различных движков обработки данных:

- Apache NIFI — для автоматизации потока данных между системами;
- Apache Spark — для распределенной обработки данных в оперативной памяти;
- SQL для обработки данных на стороне СУБД PostgreSQL, Greenplum или Clickhouse.

Обработка данных как в режиме, близком к реальному времени (near-real-time), так и в режиме пакетной обработки (batch).

Позволяет обрабатывать информацию из различных источников данных:

- реляционные базы данных: Oracle, MS SQL, PostgreSQL, MySQL;
- аналитические СУБД: Greenplum, Clickhouse;
- интеграция с web-сервисами через Rest API, SOAP;
- документарные базы данных Elasticsearch;
- поддержка Hadoop;
- очередь сообщений в Kafka.

DATA QUALITY

- ✓ Позволяет проверять качество данных с помощью настроенных шаблонов проверок.
- ✓ Встраивает процессы проверки в потоки данных для контроля качества информации в рамках общих процессов организации.
- ✓ Формирует аналитику и визуализирует результаты проверок и протоколов в метриках качества данных.
- ✓ Настраивает последовательность и периодичность контроля на любых источниках данных.



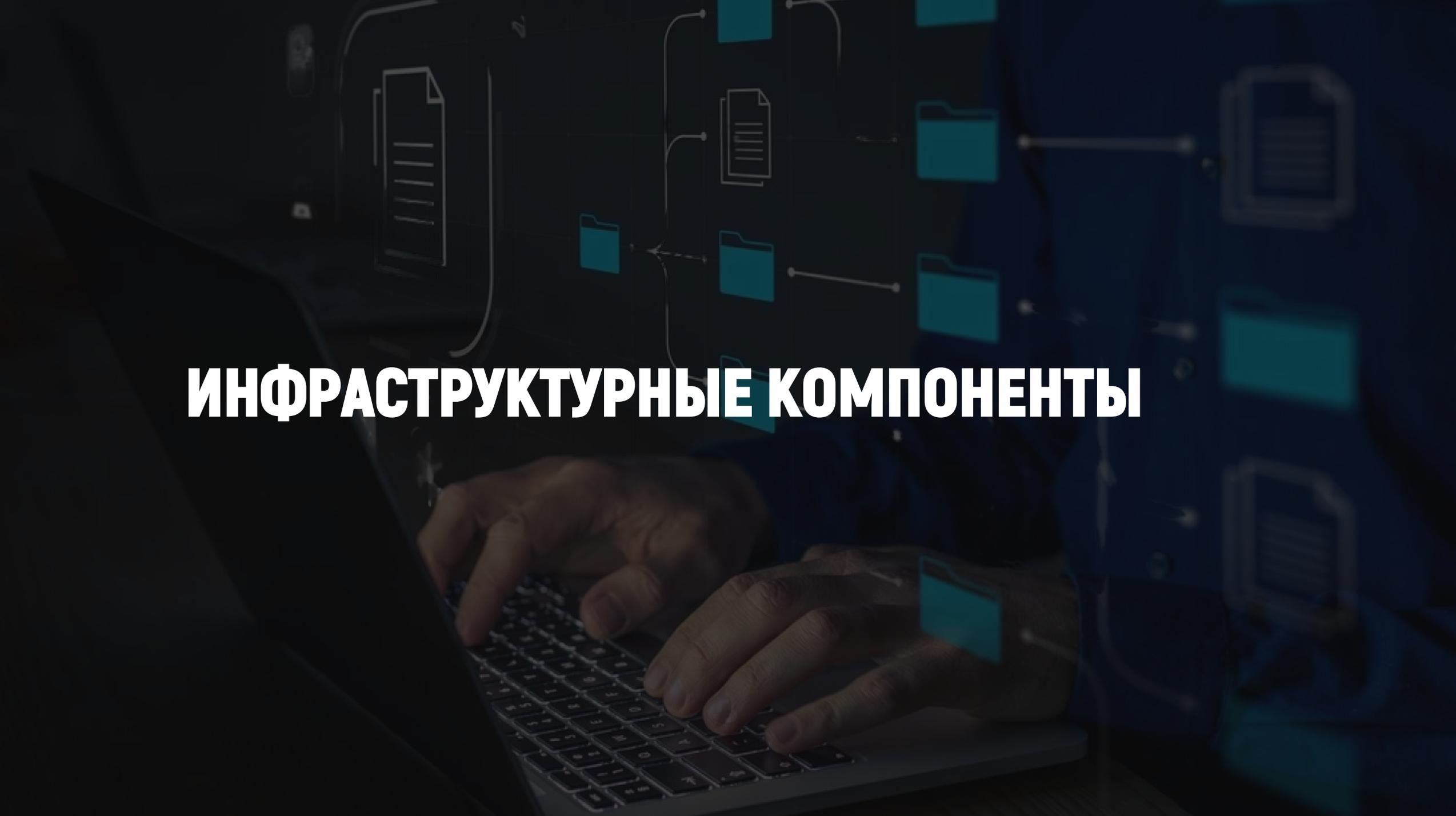
DIGITAL Q.SENSOR

- ✓ Low-code и no-code инструменты проектирования аналитических дашбордов и ситуационных центров
- ✓ Варианты визуализации на любой случай
- ✓ Возможность организации аналитических порталов
- ✓ Возможность встраивания отдельных диаграмм и графиков в интерфейсы личных кабинетов
- ✓ Подписка и получение аналитических данных в мессенджере на телефоне



REPORT ENGINE

- ✓ Генерация отчетов в форматах docx, xlsx, pdf, txt с поддержкой российских офисных пакетов Р7 и «Мой офис».
- ✓ Возможность создания шаблонов отчетов, добавления в шаблоны графических изображений, в том числе факсимильных печатей, логотипов, QR-кодов.
- ✓ Управление местом хранения готовых файлов отчетов.
- ✓ Формирование отчетов и пакетов отчетов в разрезе ролевой модели управления правами.
- ✓ Мониторинг процесса построения отчетов.



ИНФРАСТРУКТУРНЫЕ КОМПОНЕНТЫ

ИНФРАСТРУКТУРНЫЕ ПЛАТФОРМЫ



Фабрика данных использует в своей работе инфраструктурные платформы, позволяющие обеспечить оптимальное выполнение задач обработки данных.



Все инфраструктурные компоненты разворачиваются и управляются через единую среду управления Enterprise Manager.



Фабрика данных управляет использованием инфраструктурных компонентов, необходимых для решения конкретной задачи

НАЗНАЧЕНИЕ КОМПОНЕНТОВ



DIGITAL Q.NIFI

Инфраструктурная платформа для интеграции данных на базе open-source решения Apache NIFI. Содержит более 300 коннекторов к различным источникам.



Используется фабрикой данных для доступа и транспортировки данных между различными системами.



СУБД DIGITAL Q.DATABASE

Гибридная СУБД, основанная на синергии технологий хранения данных в реляционных, аналитических и in-memory СУБД.



Используется фабрикой данных для хранения и обработки информации. Подходит для задач построения хранилищ данных, формирования отчетности и дашбордов.

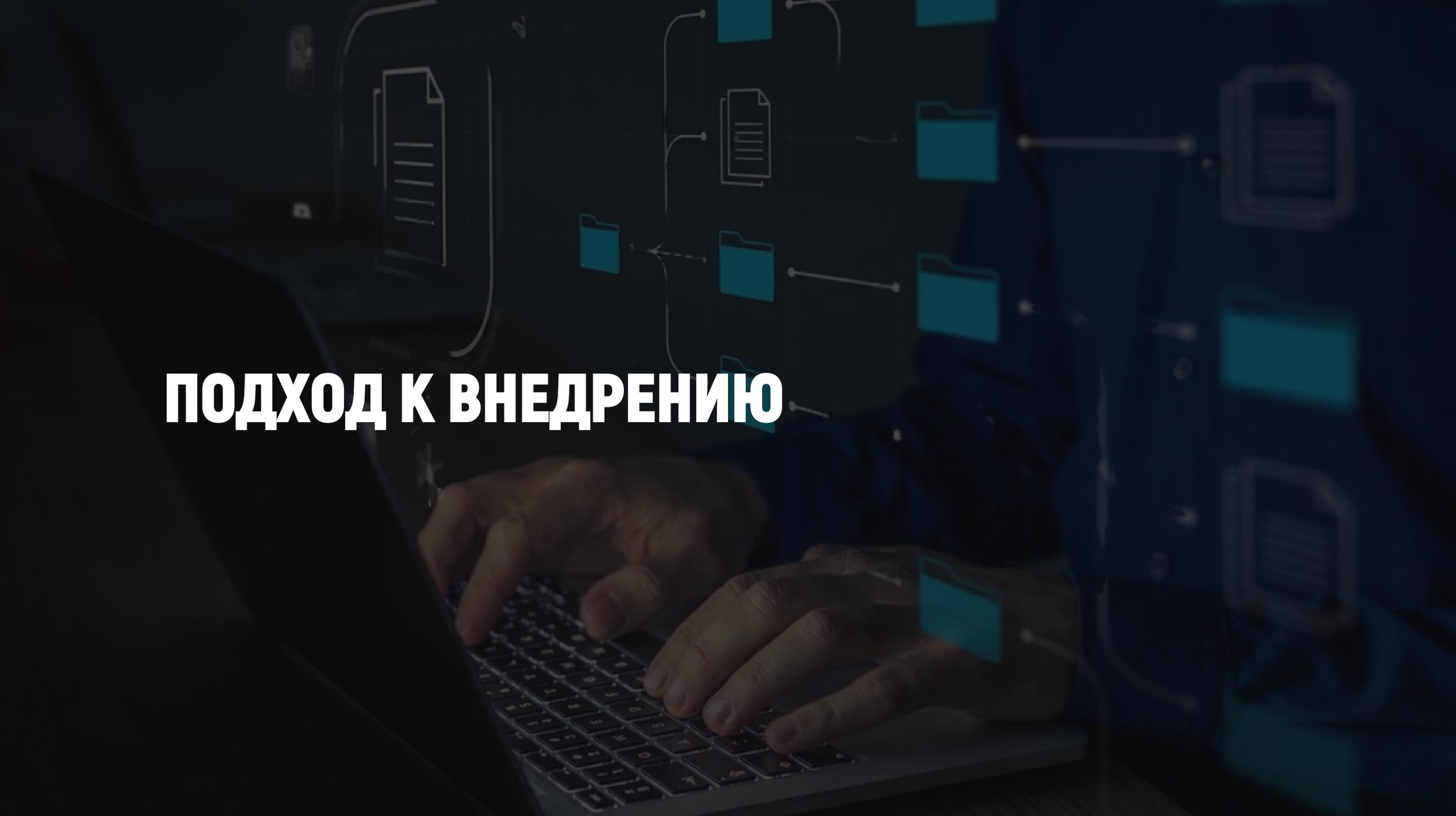


РАСПРЕДЕЛЕННАЯ ОБРАБОТКА ДАННЫХ

Хранение данных: Apache Hadoop или S3 Object Store
Обработка данных в оперативной памяти: Apache Spark.



Используется фабрикой данных для хранения и обработки информации. Подходит для задачи реализации озер данных и LakeHouse архитектур хранилищ данных и задач обработки данных искусственным интеллектом.

The image shows a person's hands typing on a laptop keyboard. The background is dark blue with a digital overlay of a file system. The overlay consists of several teal-colored folder icons and document icons connected by thin white lines, suggesting a hierarchical structure. The text 'ПОДХОД К ВНЕДРЕНИЮ' is written in white, bold, uppercase letters across the middle of the image.

ПОДХОД К ВНЕДРЕНИЮ

КАК СДЕЛАТЬ ДАННЫЕ ПОЛЕЗНЫМИ

ФАБРИКА ДАННЫХ ПОЗВОЛЯЕТ
РЕАЛИЗОВАТЬ ГИБКИЙ ПОДХОД К
ВНЕДРЕНИЮ АНАЛИТИЧЕСКИХ
СИСТЕМ

AGILE – ПОДХОД К ОБРАБОТКЕ ДАННЫХ

- Разбиваем на небольшие бизнес-задачи
- Используем только те данные, которые нужны для решения
- Получаем пользу, переходим к следующей задаче, тиражируя и используя накопленные разработки
- Фабрика данных обеспечивает использование целевых технологий и процессов обработки данных начиная с первого пилота

AGILE-ПОДХОД К РАБОТЕ С ДАННЫМИ: БЫСТРЕЕ И ГИБЧЕ

AGILE

РЕЗУЛЬТАТ: МАЛО, НО ЧАСТО

Практика показывает, что при длительном использовании agile-подход дает на порядок больше, чем waterfall

Децентрализация ответственности по бизнес-областям, построенная на общих принципах и общей платформе

WATERFALL

РЕЗУЛЬТАТ: МНОГО, НО РЕДКО

Польза для бизнеса будет в конце, после решения инфраструктурных задач

Специализированная **централизованная** команда по развитию системы

Подход к процессу
производства

Масштабируемость



AGILE-ПОДХОД К РАБОТЕ С ДАННЫМИ: БЫСТРЕЕ И ГИБЧЕ

ПРЕИМУЩЕСТВА ПЛАТФОРМЫ



ВЫБОР ТЕХНОЛОГИЙ ДЛЯ РЕШЕНИЯ БИЗНЕС ЗАДАЧИ

Платформа поддерживает различные технологии обработки данных и позволяет выбрать оптимальные для решение конкретной бизнес задачи



СНИЖЕНИЕ СТОИМОСТИ РАЗРАБОТКИ

Возможности фреймворка построения конвейеров обработки данных позволяют ускорить разработку и снизить уровень требований к квалификации дата-инженеров для разработки на платформе



ПОДХОДИТ ДЛЯ РАЗЛИЧНЫХ БИЗНЕС ЗАДАЧ

Все компоненты платформы платформы взаимодействуют и позволяют настроить конвейер обработки данных для разных бизнес задач: оперативная онлайн отчетность, построение дашбордов, построение хранилища данных



ЕДИНАЯ ОТВЕТСТВЕННОСТЬ

Ответственность за все компоненты решения от прикладных сервисов до инфраструктурных платформ, поставляемых компанией

НАЧНИТЕ ПОЛУЧАТЬ ПОЛЬЗУ ОТ ДАННЫХ С ПОМОЩЬЮ ФАБРИКИ ДАННЫХ!

Россия, 127018, Москва
ул. Полковая, д. 3, стр. 14
Т: +7 (495) 780 7575; 789 9339
info@diasoft.ru, www.diasoft.ru