

Цифровая лаборатория Т1 ИИ



Лаборатория данных



Лаборатория данных Т1 может настраиваться силами бизнес-команды заказчика, эксплуатироваться в продуктиве и приносить видимый эффект от правильной работы с данными

Включает необходимые компоненты платформы данных:

- Представление и использование
- Стандартизация и нормализация
- Локализация и хранение
- Интеграция источников

Лаборатория ИИ



Платформа Сайбокс выступает оркестратором экосистемы ИИ, интегрируется с лабораторией данных, позволяет экспериментировать с ИИ-агентами, внедрять ИИ-приложения

Экосистема на платформе Сайбокс включает:

- Сайбокс.Ядро
- Сайбокс.Конструктор

Особенности работы с данными



Для дискретной промышленности (машиностроения)



Высокая разрозненность данных

- + Данные не централизованы между площадками/направлениями производства ни по структуре, ни по гранулярности, ни по классификаторам
- + Большое количество локальной специфики



Недостаточная автоматизация бизнес-процессов

- + Большое количество учетных процессов не автоматизированы в ERP и аналогичных L4 системах, не существует надежных источников
- + Большое количество локальных данных



Высокая разнородность классов данных

- + В качестве источников могут выступать: исторические неструктурированные документы (в т.ч. сканы), сырые данные результаты моделирования и испытаний, данные с оборудования и MES, структурированные данные



Использование сложного мат. аппарата

- + Работа с данными включает сложные математические обработки временных рядов, решения оптимизационных задач, модели физических процессов. Более сложные чем во многих других областях



Ограниченное применение отчетности

- + Отчетность формируется в основном в простой табличной форме по установленным требованиям
- + Self-service аналитика и самостоятельное исследование данных слабо распространено



Отсутствие описанных наборов данных

- + Большинство наборов данных владеют отдельные специалисты/отделы, описание и доступ к этим наборам данных отсутствует
- + Все используют локальные «свои» данные



Замкнутый цикл работы с данными



Невозможно доказать ценность без внедрения, невозможно внедрить без обоснования ценности



Решение. Лаборатория данных T1



Лаборатория данных T1

Представление и использование

- + Компонент для просмотра и редактирования данных в таблицах любого подключенного источника
- + Выгрузка в Excel заданного шаблона для локальной обработки
- + Объяснение данных с помощью ИИ



Стандартизация и нормализация

- + Набор типовых обработок для анализа статистики по данным и исправления типовых ошибок и их нормализации в т.ч. с использованием ИИ
- + SDK для разработки собственных обработок
- + Ведение справочников и баз знаний



Локализация и хранение

- + Реляционная база знаний Postgres SQL с преднастроенными зонами хранения
- + Объектное хранилище MiniO для конфигураций и рабочих файлов



Интеграция источников

- + Spark и его возможности для загрузки
- + Компонент загрузки данных из Excel
- + Airflow для оркестрации и нестандартных оркестраций
- + Интеграция данных источников в части модели данных и контента



От лаборатории к платформе данных



1

Быстрый запуск

Старт процесса выстраивания работы с централизованными данными, определение новых процессов управления данными и работы с новым типом данных, определение общих нормализованных данных

**3-5 месяцев, 5-7 млн. руб*

2

Адаптация

Ранжирование потребностей пользователей в части работы с данными, определение наиболее полезных компонентов платформы данных. Появление «живых» ролей в процессах управления данными

3

Замена компонентов

Замена встроенных компонентов, показавших свою полезность, но близких к лимиту эффективного использования на более производительные продукты и миграция данных. Сохранение пользовательских путей работы

Замена компонентов лаборатории данных



Лаборатория данных T1 может быть эффективно преобразована в различные системы работы с данными



Архитектура лаборатории данных

Готовые решения / технологии

Проектная работа / стандартная компонента



Блоки демонстрации



Базовые компоненты

Работа с источниками данных

Регистрация внешних баз данных, создание таблиц и проверок данных, загрузка и выгрузка excel по шаблонам, объектное хранилище системы

Работа с локальными данными

Настраиваемый пользовательский интерфейс для ввода, редактирования, фильтрации и проверки данных в системе

Self Service интеграция

Возможность настройки интеграционных потоков для перегрузки данных с учетом различных профилей нагрузки (от отдельных записей до миллионов строк)

Self Service обработка данных

Возможность использовать самостоятельно разрабатываемые преобразования и код на уровне базы. Виды обработок

Недостаточно просто положить все данные в одну базу: чтобы данными можно было пользоваться для отчетности и принятия решений, сначала их нужно привести к общему знаменателю

Нормализация данных

Классификация данных

Использование локальных классификаторов, моделей ИИ и алгоритмов для распределения записей по таксономиям

Извлечение атрибутов

Определение стандартов, материалов и нормализованных наименований значений с использованием обработок на платформе (алгоритмических и ИИ)

Кодификация

Пример процесса кодификации материала, формирование уникального идентификатора записи

Обогащение справочников

Пример процесса дополнения справочников системы записями из процесса загрузки данных

Демо