

M W

MTC
WEB
SERVICES

S

**КАК МЫ
ИСПОЛЬЗУЕМ
AI АГЕНТОВ
ДЛЯ ОПИСАНИЯ ДАННЫХ
В МТС?**



О чем будем говорить

- Проблема с качеством описания данных
- Как работают с описаниями данных в больших компаниях?
- Как работать с описаниями данных решили мы
- Особенности использования LLM
- Способы улучшения качества
- Показатели продукта
- Возможности по разворачиванию

Традиционный путь

2 шт

Таблицы описывается в день

300+ млн

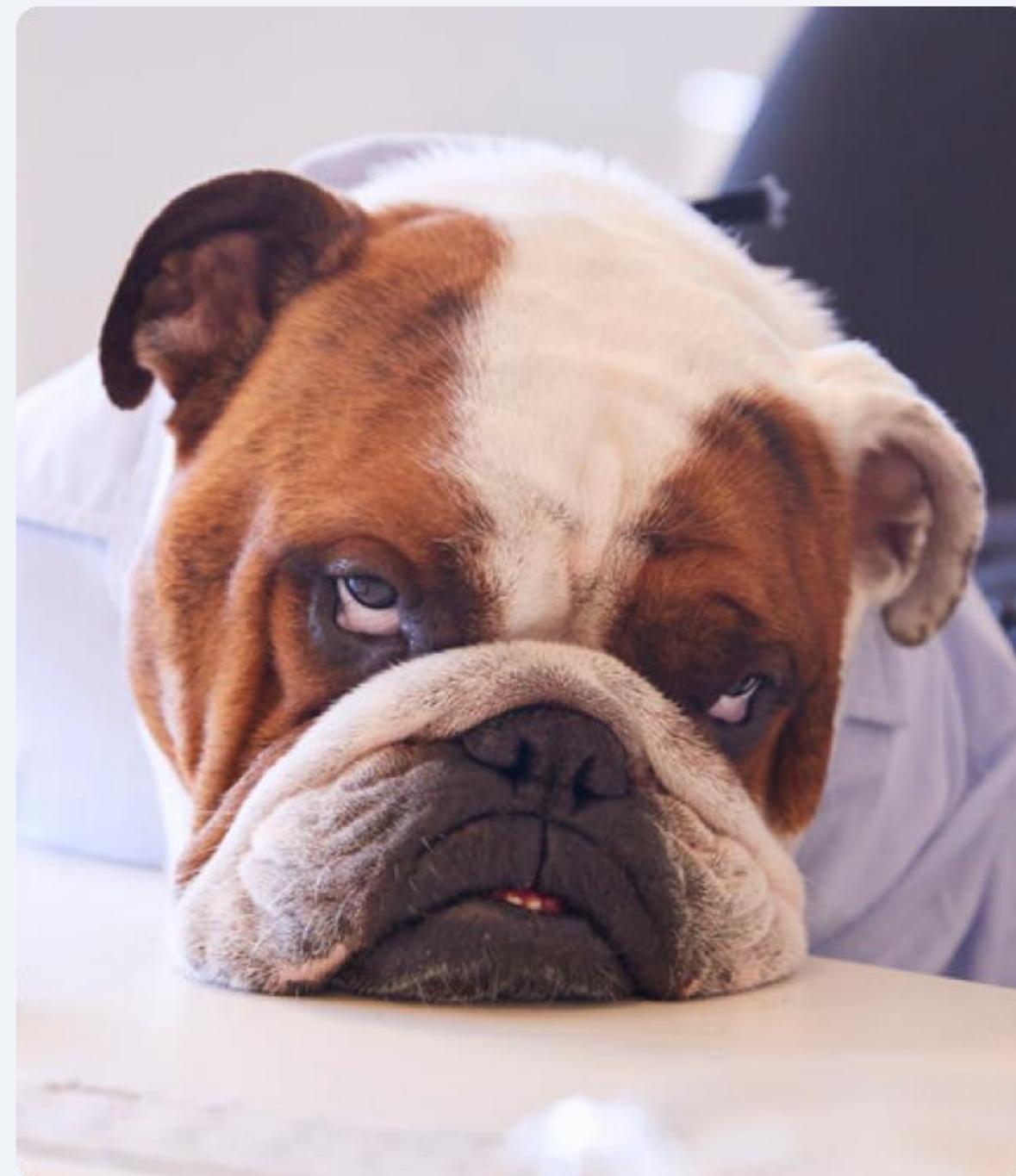
Затраты на аналитиков
для описания данных



Аналитики выполняют рутинную работу, анализируя десятки источников и витрин, вручную пытаюсь разобраться в структуре данных

Кейс МТС:

Исследовали систему Resource Inventory. 7 человек на протяжении 3х месяцев. Совокупно обработали 310 витрин



Что решили сделать мы?

Разработали AI Агента для автоматизации работы с данными на базе ИИ, которая интегрируется с популярными Дата Каталогами



Автоматически генерируем описание схем, таблиц, столбцов



Находим критические данные



Строим физическую и логическую модели данных (ER диаграммы)

Описание таблицы actlinkauto

AI+ Таблица автоматических ссылок на действия связей Actions

Колонки

Имя	Тип	Тег	AI+ Описание
autoservername	VARCHAR(255)		Имя сервера действия
recordid	NUMBER		Идентификатор записи действия.
actlinkid	NUMBER		Идентификатор ссылки действия
actionindex	NUMBER		Порядокиндексации действий.

Тег	Статус	Уровень применения	Ответственный	Дата
Критические данные - Phone number	Активный	Глобальный		01/23/2025, 14:49:25
Критические данные - Email	Активный	Глобальный		01/30/2025, 11:42:45
Критические данные - ФИО	Активный	Глобальный		01/30/2025, 11:42:55
Критические данные - Password	Активный	Глобальный		01/30/2025, 11:43:04

СОЗДАТЬ ТЕГ ЗАПУСТИТЬ РАЗМЕТКУ ДЛЯ ДАННОГО ИСТОЧНИКА ЗАПУСТИТЬ РАЗМЕТКУ ДЛЯ ВСЕХ ИСТОЧНИКОВ

LLM решит все проблемы?

Если «просто закинуть» в LLM метаданные, то ответ будет некачественным

01 Базовая информация даёт базовые описания

02 Отсутствие ключевых слов/терминов приводит к снижению качества

03 Чем менее понятный запрос, тем больше вероятность галлюцинаций



Классификация таблиц

Классификация по способу использования.

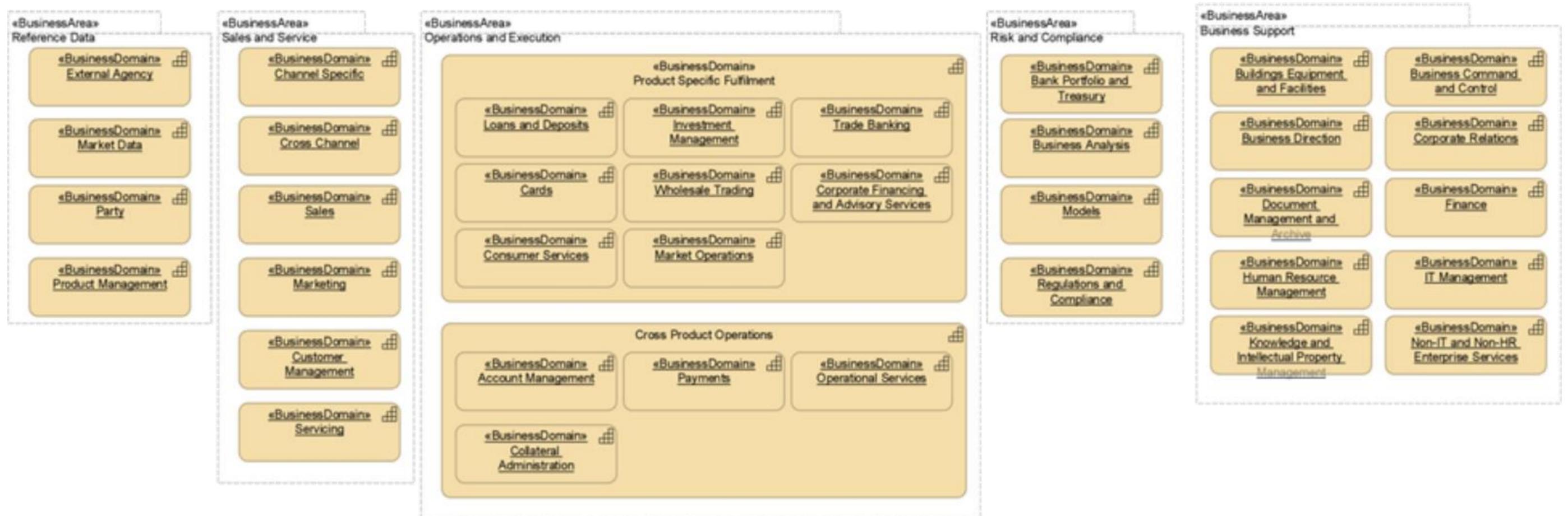
СМ или классические методы машинного обучения

Класс	Описание
Log	Используются для мониторинга и записи событий в системе.
Transactional	Фиксируют бизнес-транзакции или процессы ETL, записывая информацию о каждой транзакции. Этот класс также включает в себя stage таблицы.
Directory	Хранилище справочных данных, которые остаются неизменными с течением времени. Они включают в себя более-менее или совсем постоянные значения: типы, enums, константы, ед. измерения и их соответствующие описания.
Bridge	Хранит отношения many-to-many.
Summary	Хранят агрегированные статистические данные, полученные из других таблиц, такие как максимумы, минимумы, средние значения, итоги и другие сводные метрики.

Отраслевая и доменная классификация

Классификация выполняется для определения в какой домен попадает та или иная таблица. Нами проработаны BIAN и eTom модели бизнесов

BIAN Service landscape overview 7.0 - Business Domains



Уникальные промпты для каждого случая

	Log	Dcionary	Transactional	Bridge	...
Loans & Deposits	Prompt 101 v1.2	Prompt 102 v1.9	Prompt 103 v1.2	Prompt 104 v2.2	...
Consumer services	Prompt 201 v2.0	Prompt 202 v1.7	Prompt 203 v3.0	Prompt 204 v2.3	...
Customer domain	Prompt 301 v0.2	Prompt 302 v1.3	Prompt 303 v0.2	Prompt 304 v2.1	...
Credit card	Prompt 401 v3.1	Prompt 402 v2.1	Prompt 403 v3.1	Prompt 404 v1.1	...
Market data	Prompt 501 v1.2	Prompt 502 v1.0	Prompt 503 v1.2	Prompt 504 v2.1	...
Finance	Prompt 601 v2.4	Prompt 602 v1.6	Prompt 603 v1.4	Prompt 604 v1.5	...

...

...

...

...

...

...

Бизнес процесс



Характеристики системы



Скорость

Описание одной таблицы за **60 секунд**
в сравнении с **6 часами** работы аналитика



Точность

80% описаний не требуют корректировки
(постоянно ведем работу над улучшением качества)



Как описать все данные в компании в 5 раз быстрее с помощью InDatabase?

Работы по внедрению

Команда аналитиков и разработчиков поможет описать все данные

Готовые интеграции



Варианты поставки платформы



Варианты поставки оборудования

Минимум 1 GPU A100 80GB



Продажа GPU MWS

Аренда GPU MWS

Сервера компании

Спасибо за внимание

Напишите мне. Обсудим детали и предложим оптимальное решение под ваши задачи



**ВИТАЛИЙ
ЛЯПЕР**

Архитектор
ИИ-решений MWS

vsliap10@mts.ru

Отсканируйте QR-код,
чтобы сохранить контакт
на смартфон



Наш сайт

<https://mws.ru/>

Наш Telegram

<https://t.me/MTSWebServices>