

**Консолидация и
повышение
доступности
данных**



1 | Проблемы консолидации и доступности данных

Основные проблемы



Разрозненность данных

Данные могут быть распределены по различным источникам внутри и вне организации, в разных форматах и структурах.



Сложность доступа к данным

Доступ к данным может быть затруднен из-за проблем:

- с безопасностью, разрешениями,
- интерфейсами API
- или даже недостатком технической экспертизы.

Кроме того, в разных системах разные критерии поиска, а часть информации может быть не структурирована.



Доверие к данным

Одни и те же атрибуты и свойства в разных источниках имеют разное качество.

Данные могут быть недостаточно точны или не согласованы, в результате чего их использование может быть ограничено.



Сложность поддержки

Данные и метаданные должны обновляться в соответствии с изменениями в бизнес-процессах и внешнем окружении.

В результате, добавление/изменение новых данных в периметр организации выполняется через сложный процесс, либо компания регулярно сталкивается с вопросом доверия к данным.

2 | Customer 360

Customer 360

Сбор уникального досье

Интеграция информации из различных источников, таких как CRM, транзакционные системы, системы ведения коммуникаций, интернет-банкинг и т.д.

Соединение данных

Сопоставление данных между источниками информации. Подготовка описания данных и правил их мапинга.

Очистка данных

Очистка данных от ошибок, дубликатов, неактуальной информации. Построения процессов поддержки данных в актуальном состоянии.

Анализ и наполнение досье

Формирование аналитических признаков, таких как шаблоны поведения и сегментация (высокодоходный клиент, лидер мнений, вероятность оттока и т.д.)



! Отсутствие универсального хранилища, способного **гибко управлять набором доменов и атрибутов.**



! **Признаки**, требуемые для анализа гипотез, регуляторных требований и запросов **бизнеса должны добавляться быстро.**



! Чаще всего Customer 360 строится на **базе CRM системы**, либо некой inhouse-разработки.



! **Скорость внесения** подобных изменений в решениях такого класса обычно невысокая или сопряжена со сложными организационными процессами.

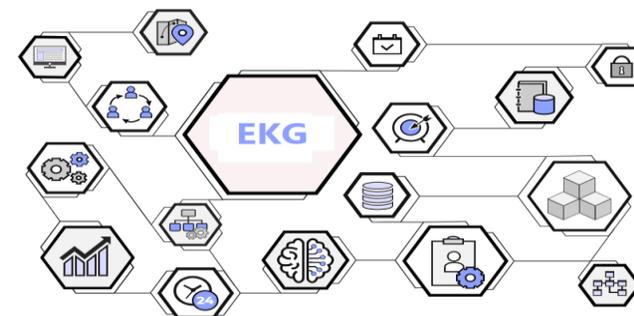
3 | Корпоративный граф знаний (ЕКГ)

Критерий	Граф	Таблица
Где хранится структура данных	В самом графе	Отдельно, в описании таблицы
Где хранится логика обработки	В правилах в графе	В хранимых процедурах и коде
Типизация сущностей	Поиск связей и зависимостей, сложные запросы	Массовые операции с однотипными объектами
Редактирование структуры данных	Сущность может иметь много типов – входить в классы	Каждая сущность принадлежит к одному типу – таблице
Многоязычность значений	Делает аналитик. Структура данных легко меняется	Делает программист. Код зависит от структуры
Многоязычность значений	Метка языка у каждой строки	Отдельный столбец для каждого языка

Комментарий

Представление корпоративных знаний и информации, структурированное в форме графа. **Он представляет собой сеть сущностей и связей между ними.**

Описание данных в виде графа позволяет визуализировать и анализировать более сложные взаимосвязи и зависимости между бизнес-объектами.



4 | Применение EKG

Методы и технологии

- ▶ EKG предоставляет методы и технологии для описания структуры данных.
- ▶ (Методы – онтологическое моделирование, технологии – спецификации W3C).
- ▶ Управление структурой данных из создания полей превращается в управление правилами и их применение.
- ▶ На уровне разработки описывается порядок работы с типами данных (различные СУБД, брокеры очередей и т.д.), работа со структурой данных и правилами описывается на уровне моделей.

Правила

- ▶ Позволяет описать правила консолидации, очистки и валидации данных.
- ▶ Если в любом решении на базе реляционной СУБД потребуется разработка запрограммированных процедур извлечения данных, **в EKG это описано в виде правил, которые могут быть достаточно сложными.**
- ▶ Редактирование правил требует меньше усилий, выполняется и проверяется быстрее, чем программирование процедуры.

Точки зрения

На один и тот же объект можно смотреть по-разному.

В зависимости от процесса или роли в процессе, один и тот же объект воспринимается участниками по-разному.

Клиент для сотрудника отдела продаж не то же самое, что клиент для отдела управления рисками.

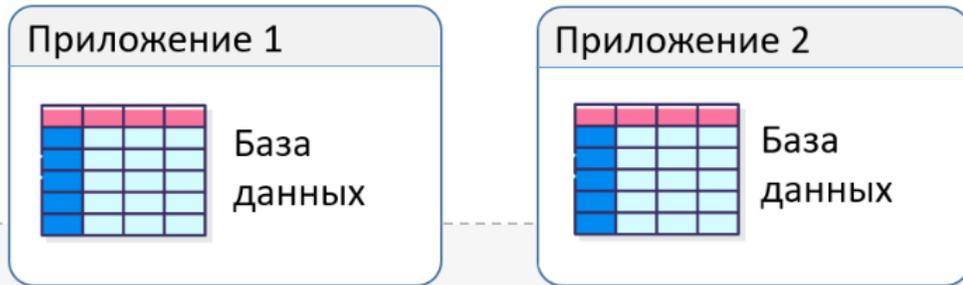
Инструменты поиска

Применение EKG облегчает поиск информации. Запрос выглядит понятнее, более интуитивно, а также будет более производительным.

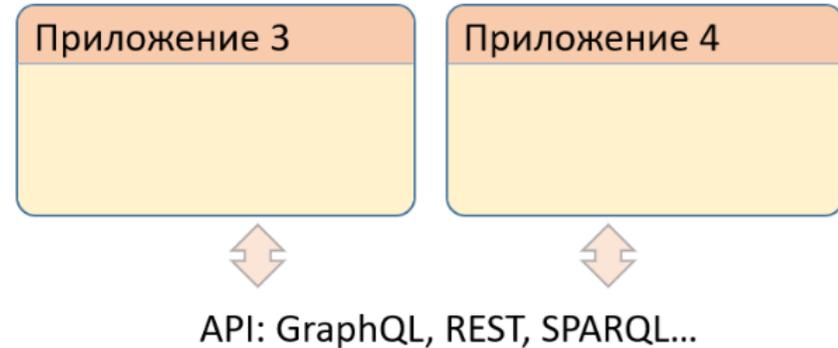


5 | Внедрение в условиях работающего бизнеса

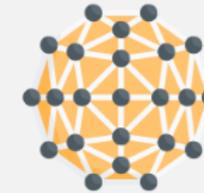
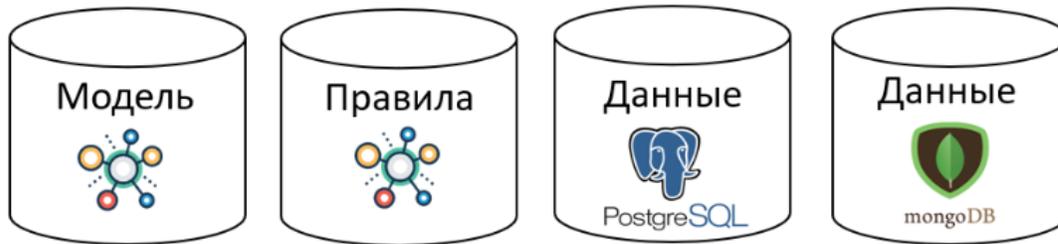
Унаследованные приложения



Датацентричные приложения



Логическая витрина данных



Виртуальный корпоративный граф знаний

Корпоративная платформа виртуализации данных

6 | Преимущество EKG



7 | Общий пример применения NLU, как тренда будущего периода



Работа с неструктурированными данными

- ! Извлечение информации из документов. Из договора можно извлечь много формализованной информации. Участники договора, их роли, предмет договора и т.д.

Например:

- ! Передача долгов коллекторским агентствам
- ! Решения судебных органов
- ! Работа с обращениями клиентов
- ! Нормативно-правовые акты



Диалоговые помощники

- ! Причина обращения клиента?
- ! Каким НПА надо соответствовать при открытии счета клиенту-нерезиденту?
- ! В каких договорах предметом залога являются объекты недвижимости площадью более 100 кв.м?

За счет применения **NLU** можно обогатить граф знаний извлеченными знаниями. **Оттуда они уже будут доступны потребителям (BI, любые другие прикладные системы)**

8 | Дальнейшие тренды развития

Тренды



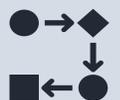
Продукты
прикладного
применения

- На основе корпоративного графа знаний уже появляются нишевые коробочные продукты, и их количество будет расти



Взаимодействие
NLP с EKG

- Технологии обработки естественного языка может использовать EKG для улучшения понимания текста, в свою очередь, благодаря NLP можно получать структурированные данные из неструктурированных для извлечения смысла и расширения графа.
- Граница между структурированными и неструктурированными данными стирается, что очень хорошо с точки зрения потребителя, т.к. получается что потребитель может работать с любой нужной ему информацией как инструментами обработки структурированных, так и неструктурированных данных.

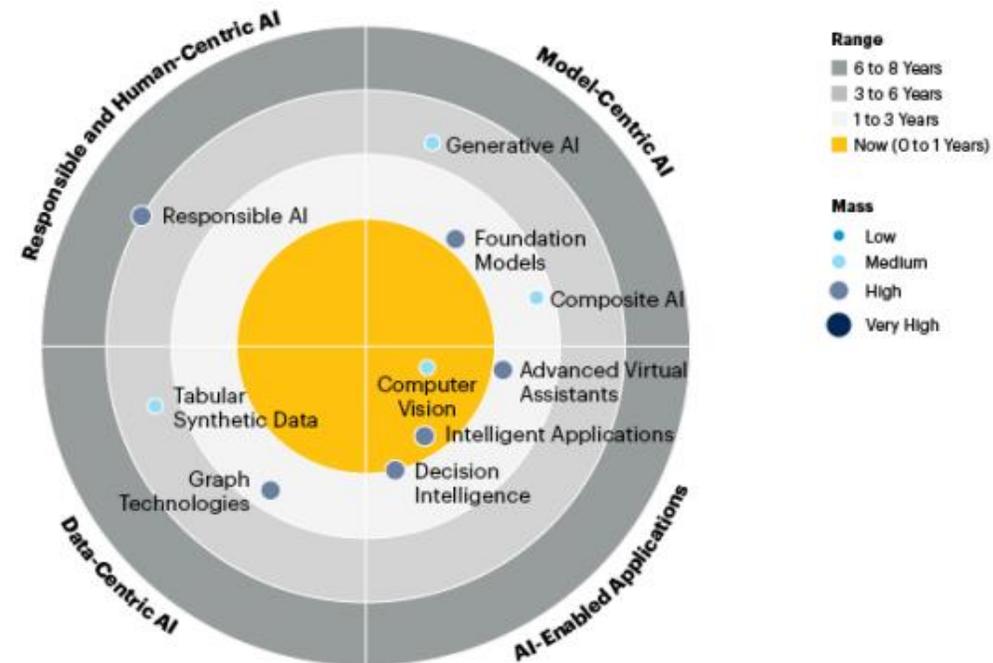


Графовые
технологии

- Графовые технологии станут основной для сложных комплексных продуктов, таких как Composite AI (подход объединения различных моделей и алгоритмов)

Gartner

Impact Radar for Artificial Intelligence in Banking



Source: Gartner
781293_C

Спасибо за внимание!