

Экосистема АрхиГраф

технологии и методики создания интеллектуальных систем
с использованием онтологических моделей

Сергей Горшков

✉ serge@trinidata.ru

👉 trinidata.ru

триниData

onto
pro



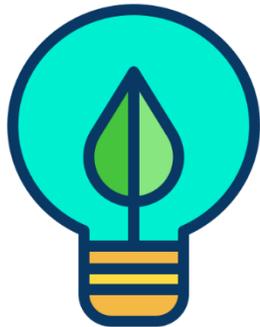
АрхиГраф.MDM



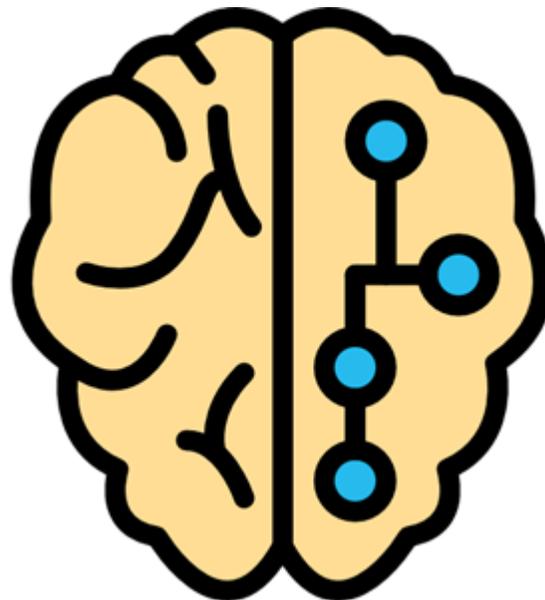
АрхиГраф.CY3

Решение интеллектуальных задач в области экономики и безопасности подразумевает выполнение системой функций, которые традиционно считаются прерогативой человеческого разума.

Интеллектуальная автоматизированная система должна **имитировать** те или иные аспекты мышления.



Эвристическое мышление и восприятие имитируются при помощи технологий **машинного обучения**, которые находят неявные корреляции между условиями задач и решений. Эти технологии проще в применении, но не способны обосновать выдаваемые решения.



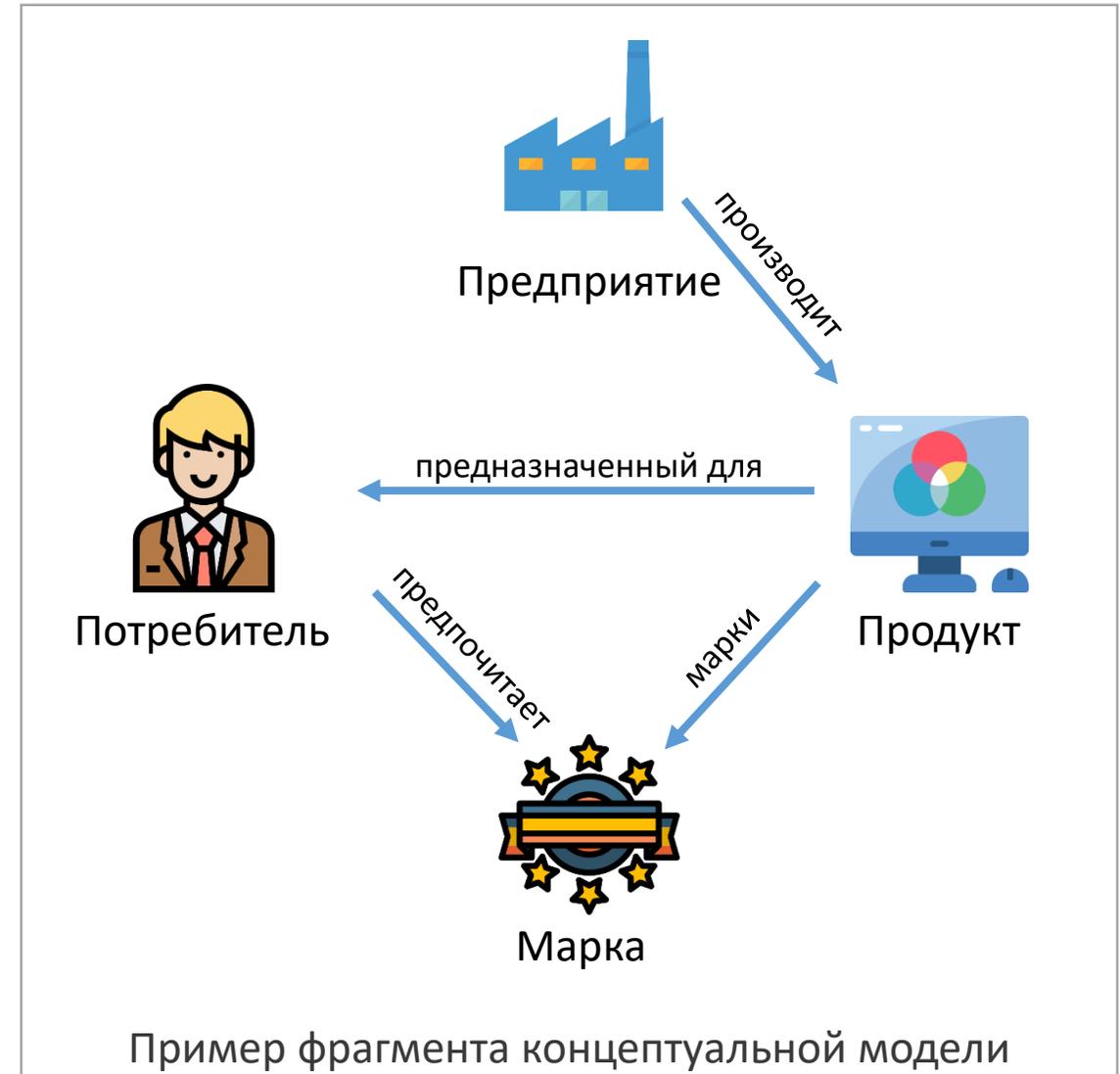
Логическое мышление имитируется путем создания концептуальной модели предметной области и представления информации в терминах такой модели. Этого позволяют достичь технологии **онтологического моделирования**.

Онтологические модели позволяют:

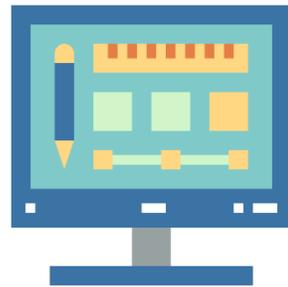
- зафиксировать набор **терминов** (концептов) для любой предметной области;
- записать любую имеющуюся информацию в этих терминах;
- построить **правила** получения логических выводов (умозаключений);
- применить эти правила для обработки любой поступающей информации;
- **обосновать** (доказать) каждый полученный результат.

Онтологическая модель состоит из двух уровней:

- **Концептуального уровня**, где определены термины для типов объектов и отношений,
- **Фактографического уровня**, на котором в этих терминах сообщаются факты о конкретных объектах.

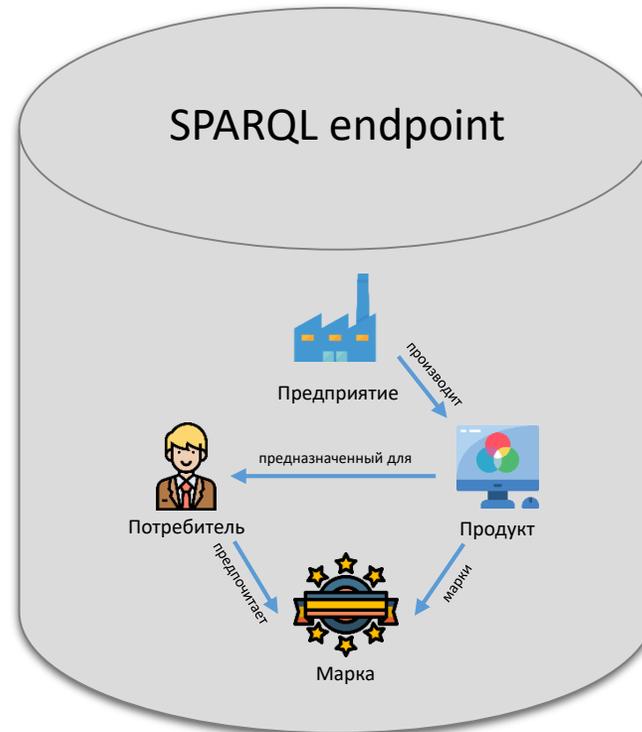


Онтологические модели создаются аналитиками-онтологами при помощи специальных редакторов и размещаются в графовых базах данных, доступных по протоколу SPARQL.



Редактор онтологий

Позволяет аналитикам-онтологам создавать как концептуальную, так и фактографическую часть модели, но чаще всего используется для решения только первой задачи



Графовая СУБД (RDF triple store)



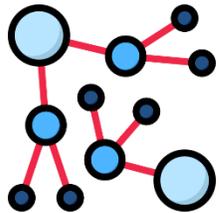
```
1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX owl: <http://www.w3.org/2002/07/owl#>
3 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 PREFIX my: <http://trinidadata.ru/my/>
6
7 SELECT * WHERE { my:Предприятие my:Производит ?product }
8
9
```

При помощи этого интерфейса любые программные компоненты могут считывать и изменять содержимое как концептуальной, так и фактографической части модели.

Графовые СУБД очень мощны, но имеют и недостатки.

Из-за этого хранить в них всю фактографическую информацию в промышленных применениях невозможно.

Преимущества



- Удобно описывать и искать связи между объектами
- Структура данных не отделена от содержания
- Можно менять структуру данных по ходу работы

Чтобы использовать все преимущества хранения информации в графовой структуре, но избежать недостатков графовых СУБД, мы создали платформу АрхиГраф.

Недостатки

- Медленно работают



АрхиГраф содержит инструменты работы с данными, **структура которых соответствует онтологии.**

При этом **хранение** данных может осуществляться практически **в любом noSQL-хранилище или реляционной СУБД.**

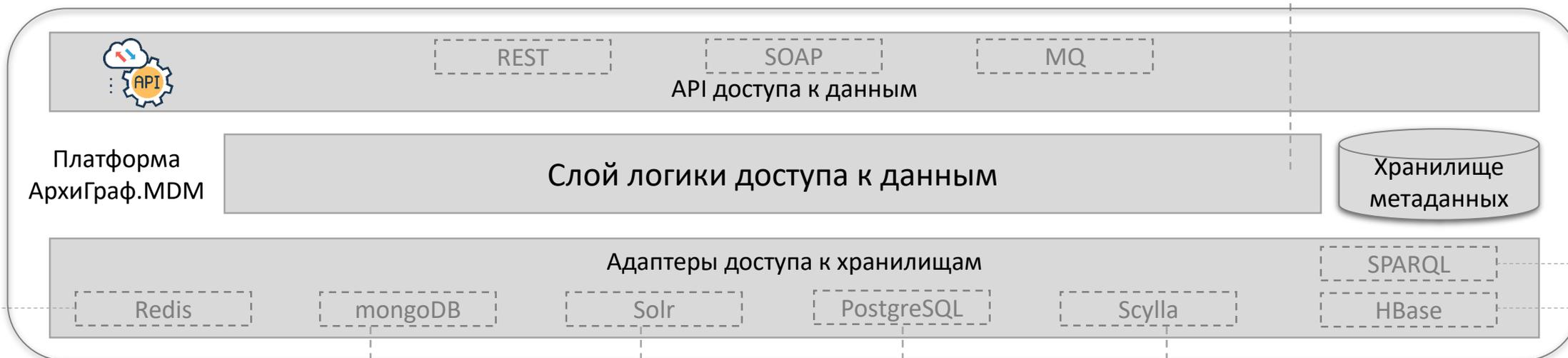
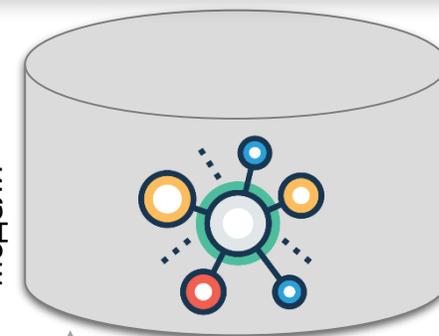
Виды запросов:

- поиск объектов по любым условиям
- изменение объектов
- групповые операции
- геопоиск
- полнотекстовый поиск

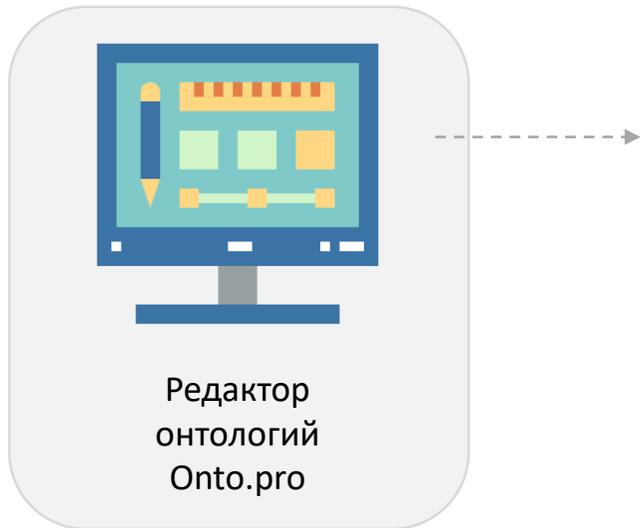
Основные функции:

- управление распределением по хранилищам
- разграничение прав доступа
- подписка на изменения данных
- запросы на изменение с подтверждением
- протоколирование доступа

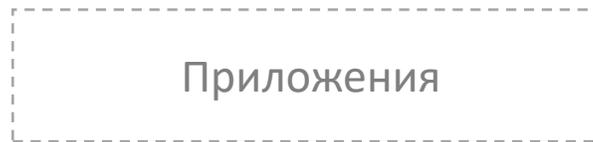
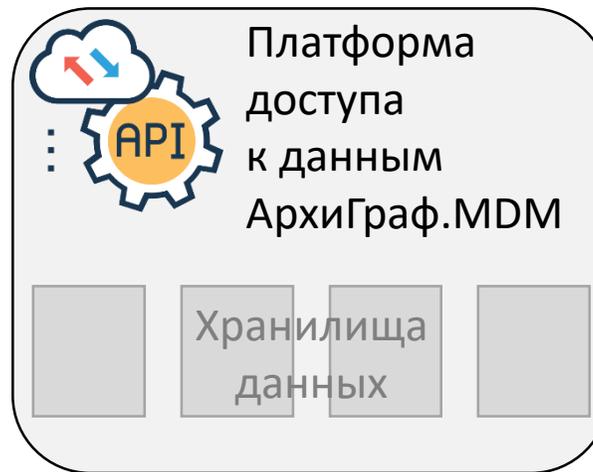
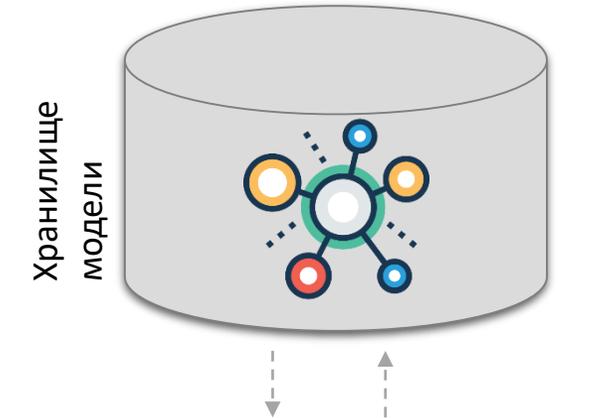
Хранилище модели



Инструменты аналитика



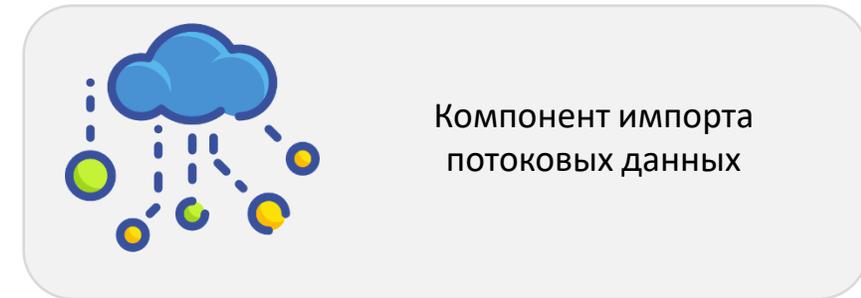
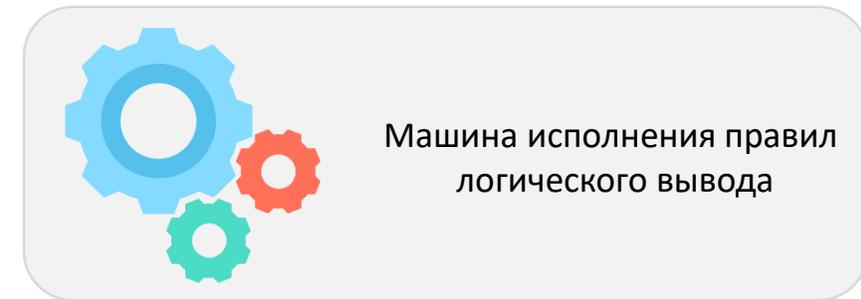
Инструменты хранения данных



Экосистема АрхиГраф включает:

- редактор онтологий Onto.pro
- редактор правил логического вывода в составе АрхиГраф.СУЗ
- платформу доступа к данным АрхиГраф.MDM
- машину исполнения правил логического вывода
- компонент импорта потоковых данных

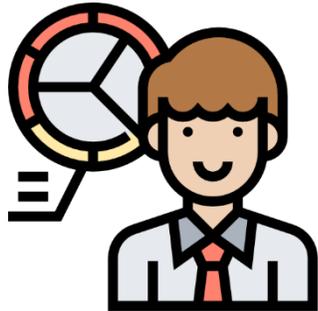
Инструменты обработки данных





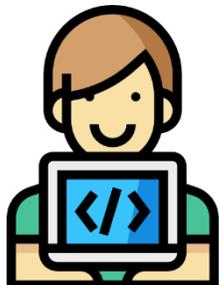
Для системного архитектора

- Перенести логику выполнения из программного кода в информационную модель
- Сократить сроки и затраты на внесение изменений
- Управлять алгоритмами работы программы без остановки и без вмешательства в код
- Менять компонентный состав приложения, не нарушая его работоспособности



Для аналитика

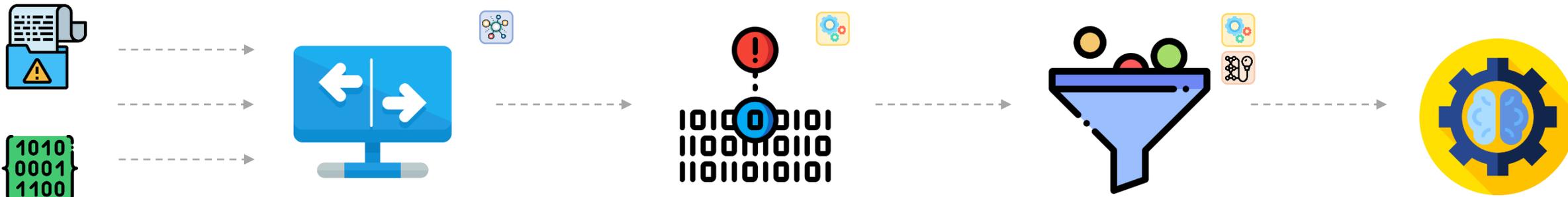
- Создать онтологическую модель предметной области
- Распределить данные между хранилищами, используя преимущества каждого из них
- Настроить импорт данных
- Создать правила логической обработки информации



Для разработчика

- Работать с данными, не заботясь о месте и форме их хранения
- Работать со структурой данных, изменяющейся по ходу выполнения
- Оптимизировать скорость работы с данными «родными» инструментами хранилищ

Один из сценариев использования АрхиГраф – обработка информации, поступающей в ситуационный центр территории или предприятия.



1) Из внешних источников поступают потоки данных, характеризующих ситуацию

2) Информация преобразуется в структуру, соответствующую информационной модели: из данных – в факты

3) Выполняется валидация поступившей информации

4) Происходит группировка и фильтрация фактов

5) Формируется комплексное описание ситуации в виде фактов в терминах модели



6) Данные визуализируются в приложении для оператора

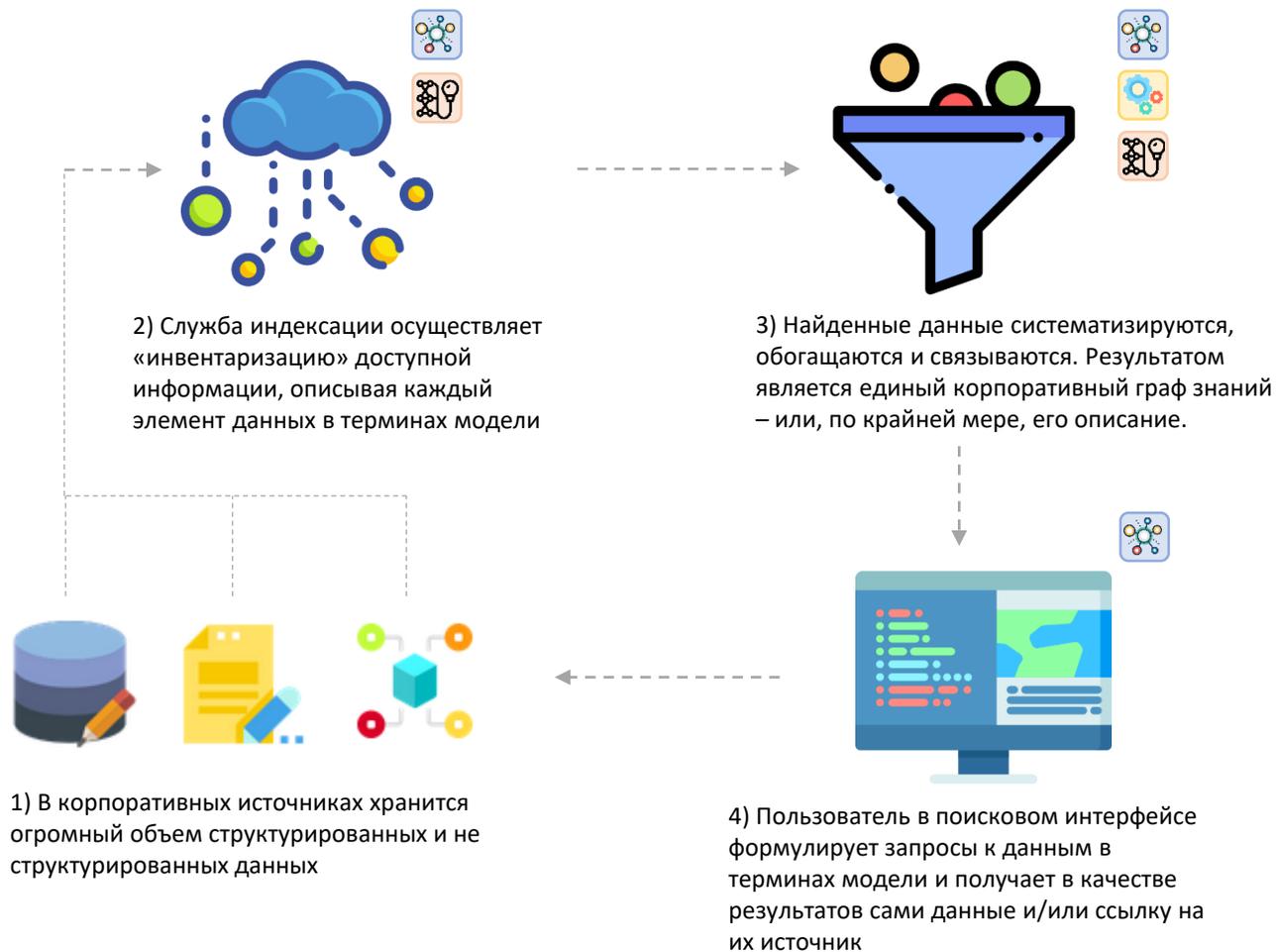
7) Система воспроизводит цепочки логических рассуждений для прогнозирования развития ситуации, выдачи рекомендаций

8) Формируются аналитические представления информации

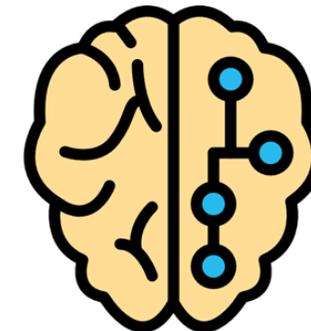
Знак  показывает, что логика алгоритма хранится в информационной модели, а  – что для обработки применяются правила логического вывода.

Знак  показывает, что на этом шаге автоматизированное получение логических выводов может сочетаться с использованием машинного обучения

Другой сценарий – обеспечение единой точки доступа к разрозненной корпоративной информации



В этом сценарии платформа АрхиГраф выполняет роль стюарда данных, выстраивая сложные связи между блоками данных и исполняя запросы на основе этих связей.



Знак  показывает, что логика алгоритма хранится в информационной модели, а  – что для обработки применяются правила логического вывода. Знак  показывает, что на этом шаге автоматизированное получение логических выводов может сочетаться с использованием машинного обучения

Спасибо за внимание!

✉ serge@trinidata.ru

👉 trinidata.ru

👉 serge-gorshkov.ru

📞 +7 (343) 2-110-256