

ЛОГИЧЕСКАЯ ВИТРИНА ДАННЫХ на базе продуктов семейства АрхиГраф

Сергей Горшков

✉ serge@trinidata.ru

👉 trinidata.ru

триниData

onto
pro

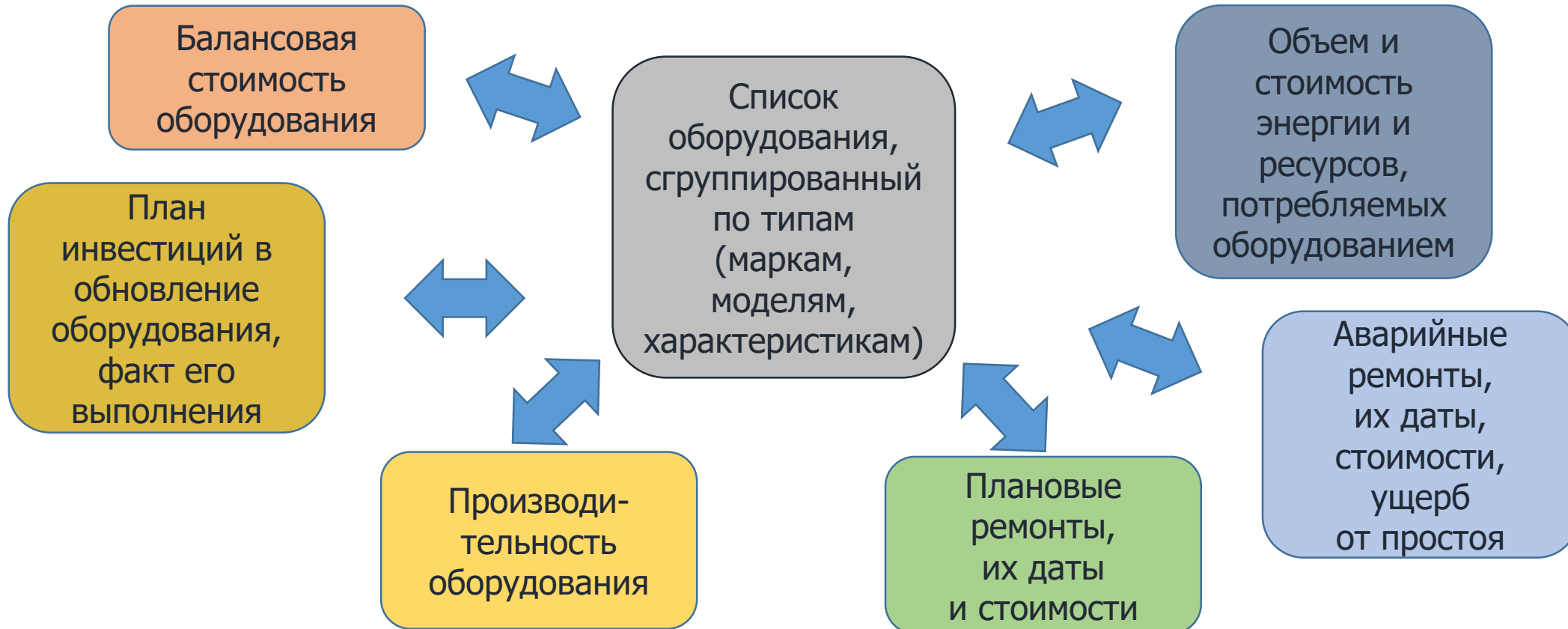


АрхиГраф.MDM

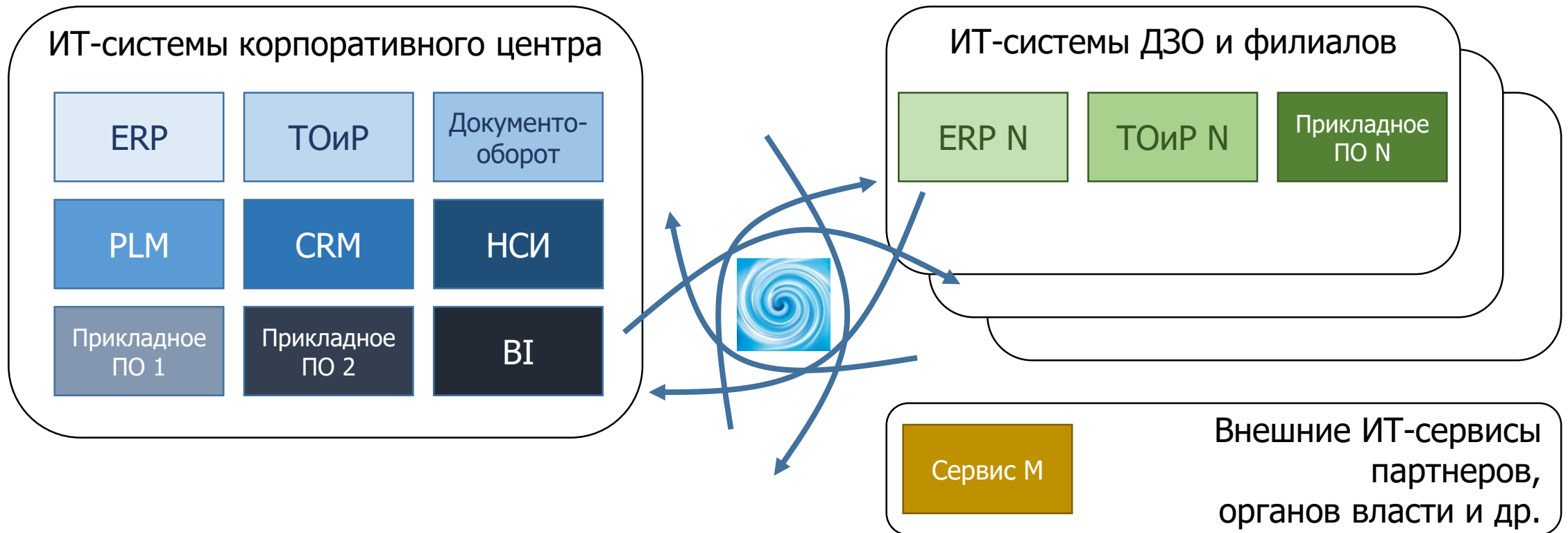


АрхиГраф.CY3

Представим себе большую промышленную корпорацию. В ней работает аналитик, которому нужно проанализировать **эффективность инвестиций** в обновление оборудования. Какие данные потребуются ему для решения этой задачи?



Построив список требуемых данных, аналитик идет в ИТ-отдел...
ИТ-системы корпорации при этом выглядят так:

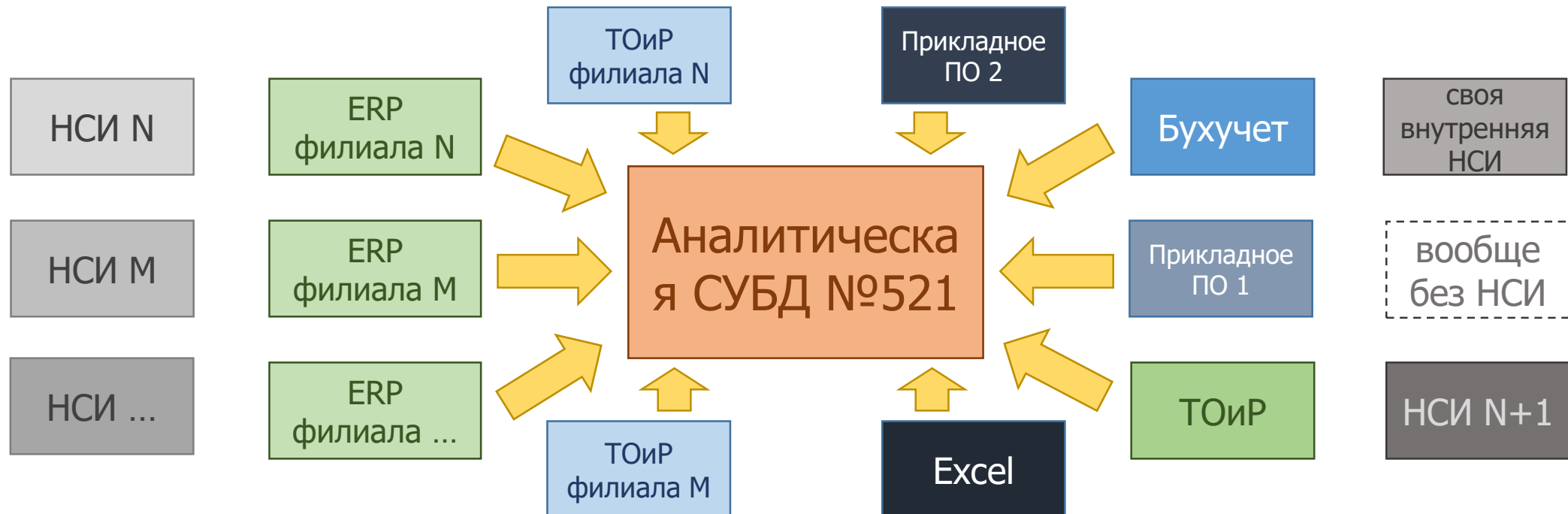


Такая ситуация складывается в ходе естественного развития ИТ-систем, догоняющего потребности бизнеса. Реализовать централизованную схему интеграции, перейти к единому набору НСИ – в таких случаях, как правило, невозможно. Слишком велики требуемые затраты и организационные усилия.

ИТ-специалисты составляют **техническое задание** на:

- создание базы данных, в которую будет помещена вся необходимая информация,
- выгрузку данных из источников (включая взаимодействие с ДЗО и филиалами),
- приведение их к единому набору НСИ (списку оборудования, как минимум),
- загрузку результата в получившуюся общую базу.

После этого аналитик сможет работать с нужными данными в своей BI-системе...



И так – каждый раз.

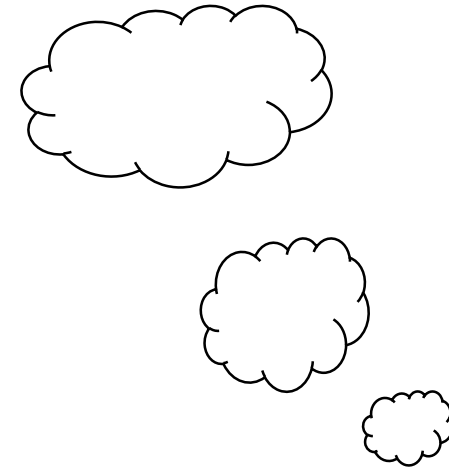
Для решения каждой новой аналитической задачи нужно собирать данные, которые:

- Разрозненны,
- Плохо доступны,
- Имеют постоянно меняющуюся структуру,
- Имеют разные наборы НСИ,
- ...

У меня есть мечта: однажды я смогу получить любые данные за пару минут.

У меня есть мечта: я смогу сделать это сам, не обращаясь в ИТ-отдел.

У меня есть мечта: четверо моих подчиненных будут заниматься аналитикой, а не поиском данных.



Сформулируем требования к витрине данных:

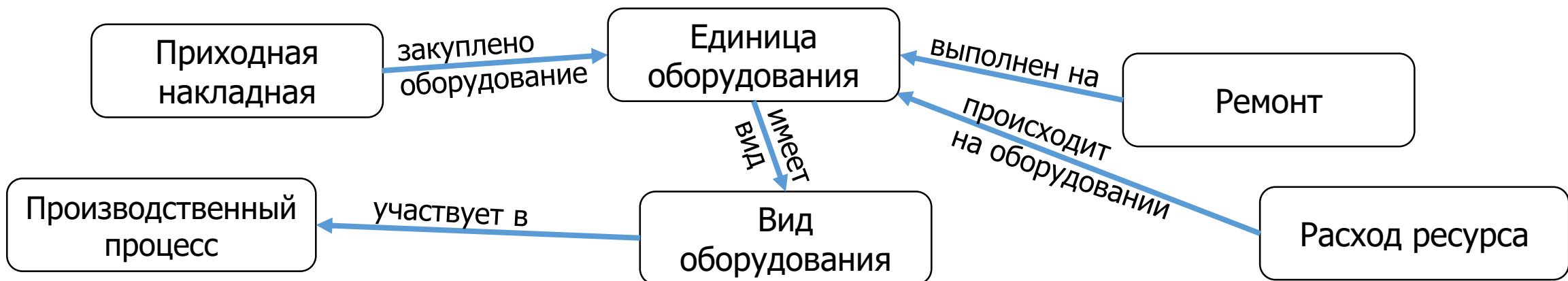
- Она должна уметь извлекать любые данные из удаленных источников по запросу.
- Пользователь может сам строить запросы к любым данным.
- Пользователь не обязан знать, откуда взяты те или иные данные, какова была их исходная структура, какому набору НСИ они соответствовали.
- Витрина не должна хранить все данные в одном хранилище – это невозможно из-за их огромного объема.

Фактически, пользователь должен работать с витриной, как с единой точкой доступа ко всей корпоративной информации, которая объединяется таким образом в единый виртуальный массив.

Часть 1: Концептуальная модель

Пользователь будет строить запросы в терминах **КОНЦЕПТУАЛЬНОЙ МОДЕЛИ** предметной области.

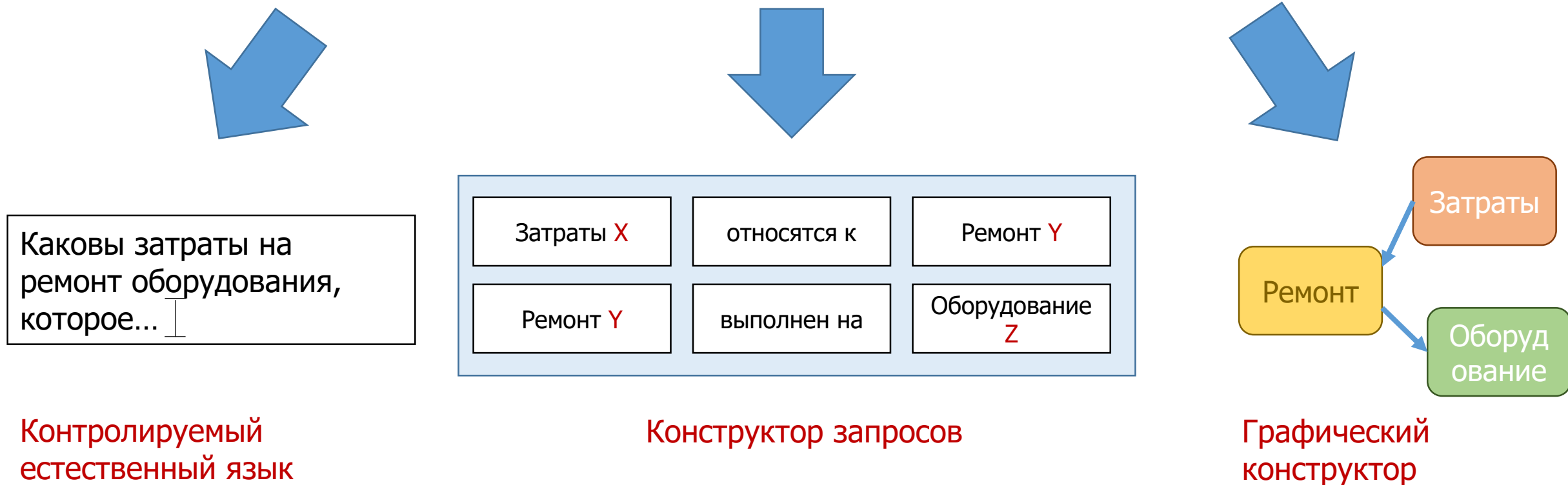
Эта модель представляет собой тезаурус, содержащий определения терминов, в которых бизнес описывает свою деятельность, и связи между ними. Это – структура корпоративных знаний.



Концептуальная модель – это только структура знаний! Фактической информации здесь еще нет.

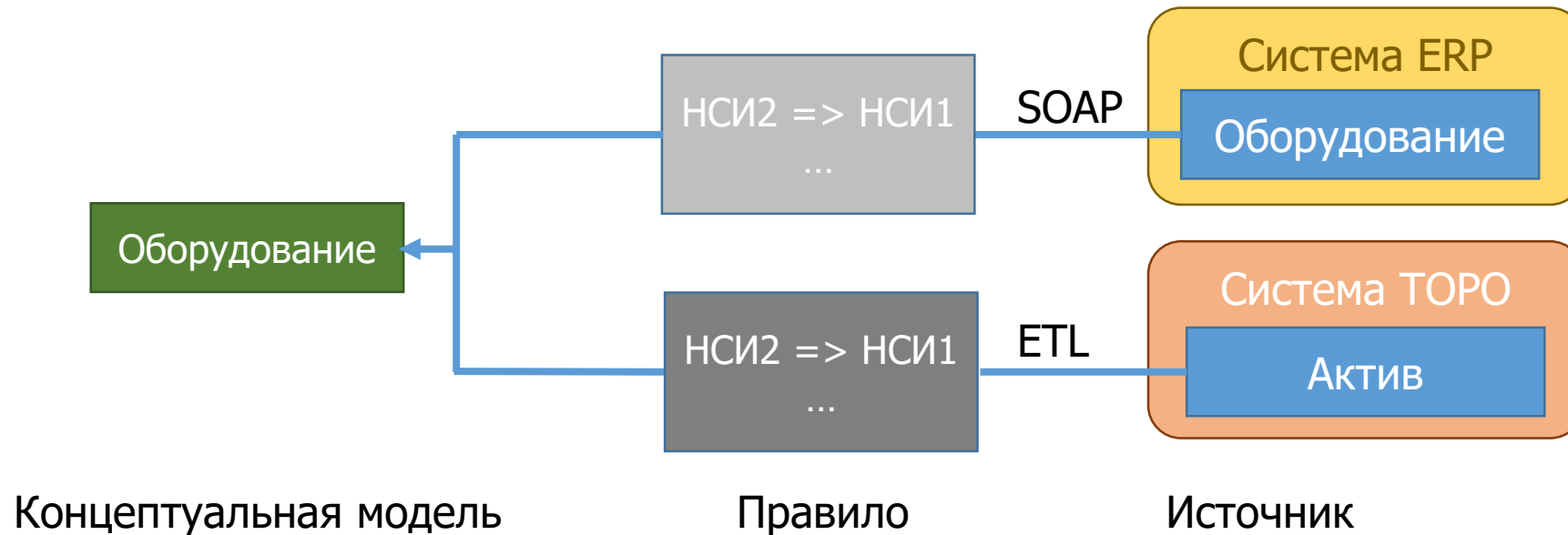
Часть 2: Интерфейс пользователя

Пользователю предоставляется несколько интерфейсов для создания запросов.



Часть 3: Правила извлечения и преобразования

В административном интерфейсе происходит привязка источников данных к элементам концептуальной модели, настройка правил их извлечения и трансформации.



Часть 4: Механизм запросов к данным

После того, как связи настроены, пользователь может делать запросы. Выполняются они так:



Реализация архитектуры витрины

Наша компания обладает полным спектром ПО для реализации логических витрин данных:

- Редактор модели – [Onto.pro](#);
- Редактор правил – [АрхиГраф.СУЗ](#);
- Конструктор запросов – [АрхиГраф.СУЗ](#);
- Хранилище модели – [АрхиГраф.MDM](#);
- Шина передачи данных – Apache Synapse или наши разработки;
- Готовые адаптеры обмена данными для ряда платформ.

Все используемые продукты – наши собственные разработки, или свободно распространяемое ПО. Мы оказываем полный спектр услуг по проектированию и реализации витрины, включая создание концептуальной модели и настройку правил.

Часть 5: Представление результатов

Результат запроса собирается из сведений, полученных от разных источников, и помещается в:

- базу данных BI-системы,
- OLAP-куб,
- экспортируется в виде файла, и т.д.

Благодаря внедрению витрины, бизнес получает:

- Расширение видов доступной аналитики.
- Радикальное сокращение времени и затрат на поиск и преобразование данных в процессе выполнения анализа.
- Возможность принимать обоснованные решения, опирающиеся на знания.
- Возможность использовать разностороннюю информацию в качестве основы для построения оптимизационных, прогнозных моделей.

Спасибо за внимание!

✉ serge@trinidata.ru

👉 trinidata.ru

👉 serge-gorshkov.ru

📞 +7 (343) 2-110-256