

Arenadata Catalog: управление данными современного предприятия

Единый взгляд на данные

Бизнес-гlossарий

- Показатели и Бизнес-объекты
- Бизнес-процессы и ИТ – системы
- Бизнес требования к качеству данных
- Контракты на поставку данных/ SLA
- Сущности/атрибуты

- Поиск по всем объектам
- Инструменты коммуникации
- История изменений
- Демократизация данных

Поиск данных и социализация



Интеграция метаданных

- Сбор технических метаданных
- Автоматическое построение data lineage: ETL продукты, DBT, Airflow/Ni-Fi
- Захват и разбор SQL запросов

- Контроль качества данных
- Захват данных о качестве из смежных систем
- Механизм расчета правил контроля качества в СУБД

Качество данных

AI Data Steward

- Генерация описаний
- Обнаружение связей и классификаций
- Гипотезы о логическом дублировании объектов

- Поточные проверки качества данных
- Собственный Domain Specific Language с low-code редактором
- Дополнительные запросы данных

Data Quality Framework

Итоги ADC 2025: Единая безопасная среда для управления данными

Поддержка Reference Data Management Реестра справочников (v.0.8.4)

Мы закрыли важную потребность в управлении и использовании референс-данных в описании объектов и наполнении бизнес-гlossария. Теперь пользователи могут создавать, вести и связывать справочники (например, почтовые индексы, коды валют) напрямую с объектами glossария. Это создает единый "золотой источник" даже для самых маленьких, но критически важных данных.

Безопасность и управление доступом на новом уровне (v.0.8.2):

- **Безопасный поиск:** Пользователи видят только то, что им разрешено. Поиск в glossарии и каталоге теперь учитывает права доступа вплоть до уровня отдельных атрибутов объекта.
- **Доступ по атрибутам:** Администраторы получили тонкий инструмент настройки политик доступа, позволяющий скрывать или отображать конкретные поля (например, примеры данных) для разных ролей.
- **Сертификация ФСТЭК:** Arenadata Catalog (ADC) был включен в Государственный реестр сертифицированных средств защиты информации и получил сертификат соответствия требованиям безопасности информации по 4 уровню доверия ФСТЭК России (сертификат № 4939) 28 мая 2025 года.

Итоги ADC 2025:

Интеллектуальное управление качеством и контекстом

Arenadata Catalog Data Quality Framework

Управление качеством на уровне бизнес-терминов:

- **Связь тестов DQ с глоссарием (v.0.8.1)**
Результаты технических проверок качества данных теперь консолидируются прямо в карточке объекта глоссария. Пользователь видит, насколько "чисты" данные бизнес-объекта, которыми он оперирует.
- **Визуализация влияния качества (v.0.8.5)**
В Data Lineage появился режим просмотра качества. Мы наглядно показываем, как "проваленные" тесты в одном объекте влияют на всех получателей данных.

Умные процессы работы с данными:

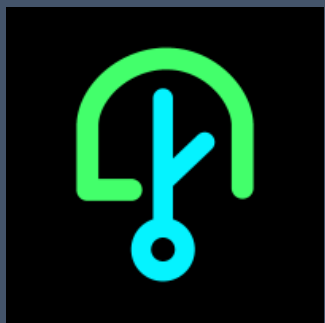
- **Правила ввода и согласований (v.0.8.1, v.0.8.5)**
Появилась возможность настраивать гибкие правила уникальности и заполнения атрибутов и требовать согласования изменений только для критически важных полей, не запуская тяжелый процесс при каждой правке.

Итоги ADC 2025:

Масштабирование и бесшовная интеграция в ИТ-ландшафт

Мы не только добавляли новый функционал, но и серьезно поработали над производительностью, архитектурой и расширением поддерживаемых источников, чтобы ADC мог расти вместе с вашим бизнесом.

- **Архитектурная гибкость (v.0.8.2, v.0.8.4):** Реализована поддержка развертывания в средах Kubernetes и с помощью Bare Metal инсталляций, что дает нашим заказчикам свободу выбора инфраструктуры.
- **Новые коннекторы метаданных:** Мы расширили периметр сбора метаданных, добавив поддержку критически важных для корпоративных клиентов систем: 1C, SAP ERP, Гармония MDM, Polymatica Dashboards, Greengage. А также углубили существующие коннекторы, добавив захват хранимых процедур (ADB, PostgreSQL, MS SQL, Oracle).



Итоги ADC 2025:

Новые проекты и расширение сообщества ADC



FSI

Системообразующий федеральный банк

Внедрение корпоративного бизнес-гlossария, повышение эффективности Data Governance

Универсальный Топ-10 банк РФ

Внедрение корпоративного бизнес-гlossария и каталога метаданных

Широкопрофильный финансовый холдинг

Автоматизация управления данными

Топ-20 банк РФ

Комплексное внедрение DG

Региональный банк

Комплексное внедрение DG

Крупнейшая страховая компания

Национальная перестраховочная компания

Retail

Крупнейшая сеть ресторанов быстрого питания

Автоматизация управления данными для предприятий сети и франчайзи

Производитель и дистрибутор ИТ-оборудования

Автоматизация управления данными и повышения достоверности аналитики

Лидирующая сеть гипермаркетов DIY

Автоматизация управления данными

Национальный оператор лотерей

Корпоративное хранилище данных построенное по методологии DG

Media

Медиа холдинг национального масштаба

Повышение достоверности, прозрачности расчета финансовых показателей

Manufacturing

Металлургический холдинг полного цикла

Создание хранилища данных финансового департамента, внедрение Data Governance

Лидер рынка Property tech

Каталог данных для data office

FMCG

Крупнейший производитель

Миграция хранилища с одновременным внедрением Data Governance

Производитель продуктов питания

Внедрение Data Governance

48+

клиентов





ADC MCP-сервер

AI Agents

AI DATA
STEWARD

AD Advanced RAG

ADC features

AI Data Steward

семейство ИИ-агентов, предназначенное для повышения эффективности работы с Arenadata Catalog

Чат-бот для поиска данных

- Популяризация каталога данных в организации
- Упрощение доступа к программному обеспечению
- Сокращение T2D

Анализ логического дублирования объектов

Предложения по связям метаданных с глоссарием

- Снижение трудозатрат при ведении каталога
- Улучшение качества метаданных
- Повышение актуальности наполнения каталога

Генерация и регенерация описаний физической модели данных



Предложения по разметке классификаторами

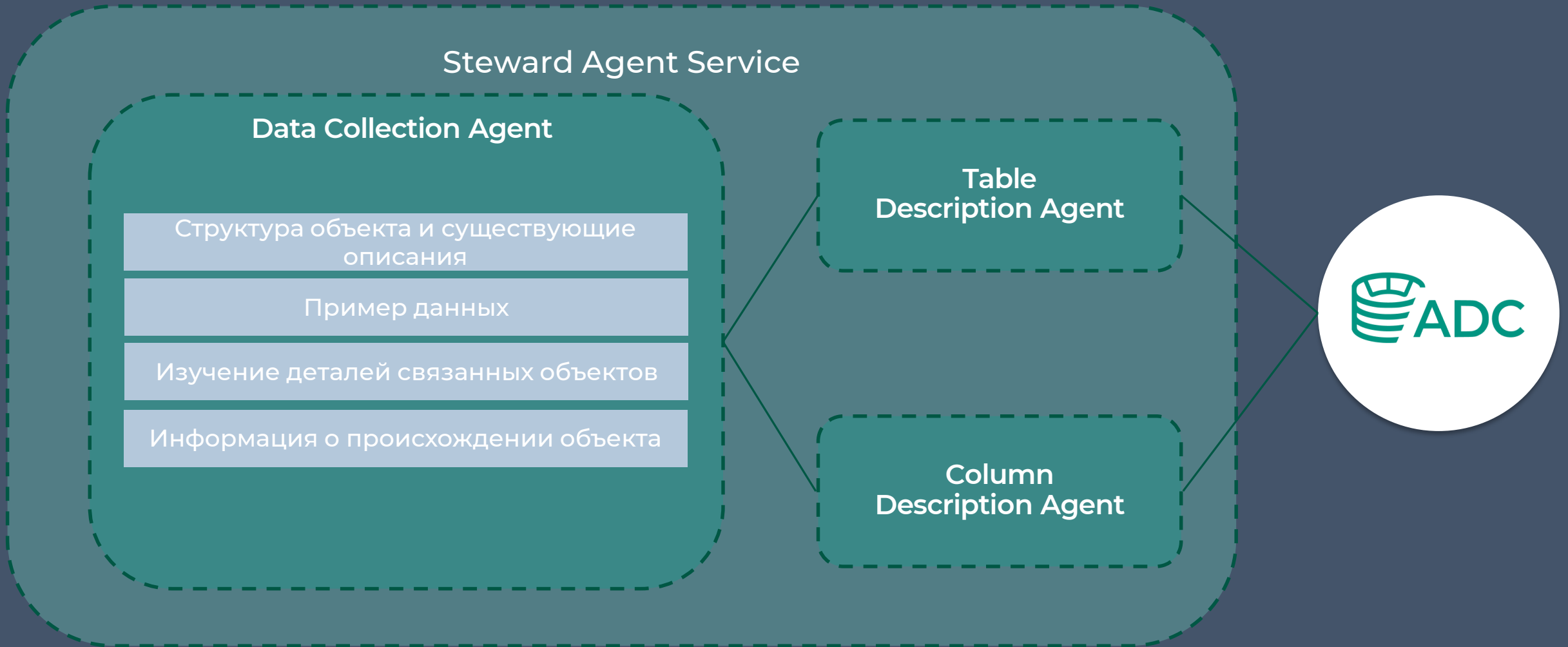


Генерация SQL-тестов для оценки качества данных

Text-2-SQL – ad-hoc запросы к СУБД на естественном языке

- Демократизация данных
- Оптимизация процессов
- Улучшение качества данных

Генерация описаний для физической модели данных



Результат выполнения: сгенерированная заявка на изменение/ заполнение описания объекта и элементов его структуры



Генерация описаний для физической модели данных

Пример запроса:

```
{  
  "openmetadata_jwt_token": "",  
  "table_fqns": [  
    "gr_test.postgres.adcbugs_2825.book"  
  ],  
  "category_names": [  
    "Категория персональных данных"  
  ]  
}
```

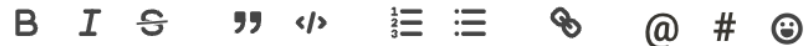
Пример результата:

Запросить пользовательское описание для table opti_db

Назначенные лица:  

Создано:  chat

Таблица используется для хранения информации об оптимизационных сценариях и результатах, связанных с планированием и управлением маркетинговыми акциями на производственных и логистических процессах, обеспечивая интеграцию данных в реальном времени из систем SCADA2 с данными о промо-кампаниях для повышения эффективности цепочки поставок и точности прогнозирования спроса.

B I 

Используйте знак @, чтобы упомянуть пользователя или команду.

Используйте знак #, чтобы сослаться на объект каталога данных или глоссария.

Закреть

Принять рекомендацию



КОНТАКТЫ

i.novoselov@arenadc.io

Иван Новосёлов

Генеральный директор ООО «ДатаКаталог»

a.nesterchuk@arenadc.io

Артем Нестерчук

Владелец продукта Arenadata Catalog

<https://arenadc.io>

