


Интеграция данных

Краснопольский Павел, Архитектор решений
Qlik СНГ

A decorative graphic consisting of overlapping, semi-transparent wavy shapes in shades of purple, blue, green, and brown, set against a background of a blue sky with white clouds. The shapes create a layered, mountain-like effect.

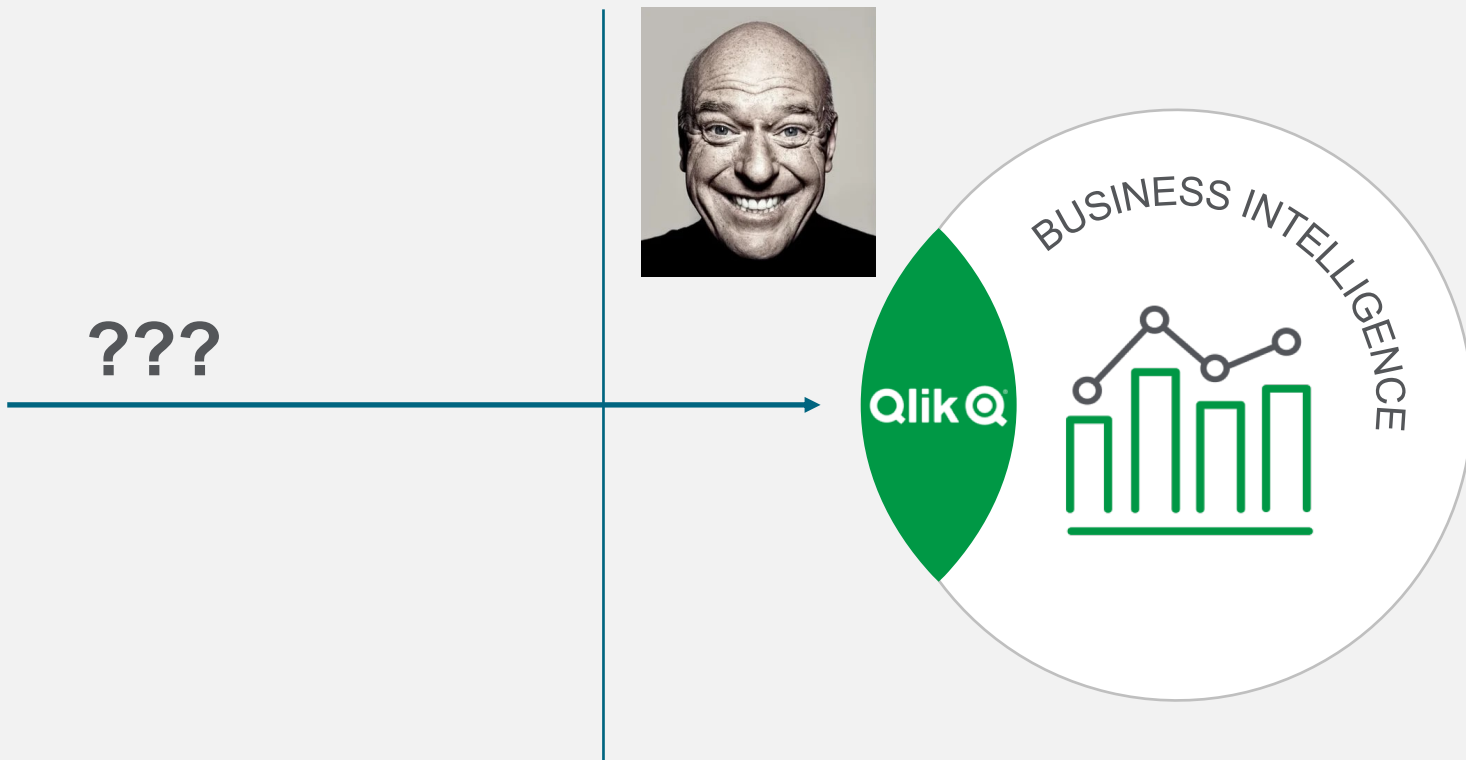
Welcome



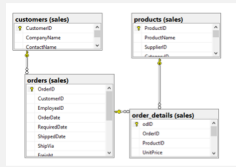
Qlik работает с данными



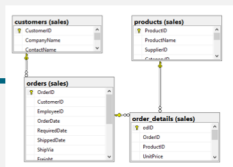
Qlik работает с данными, но откуда берутся данные?



Qlik работает с данными, но откуда берутся данные?



Oracle



MS SQL



SAP

Qlik Q

BUSINESS INTELLIGENCE

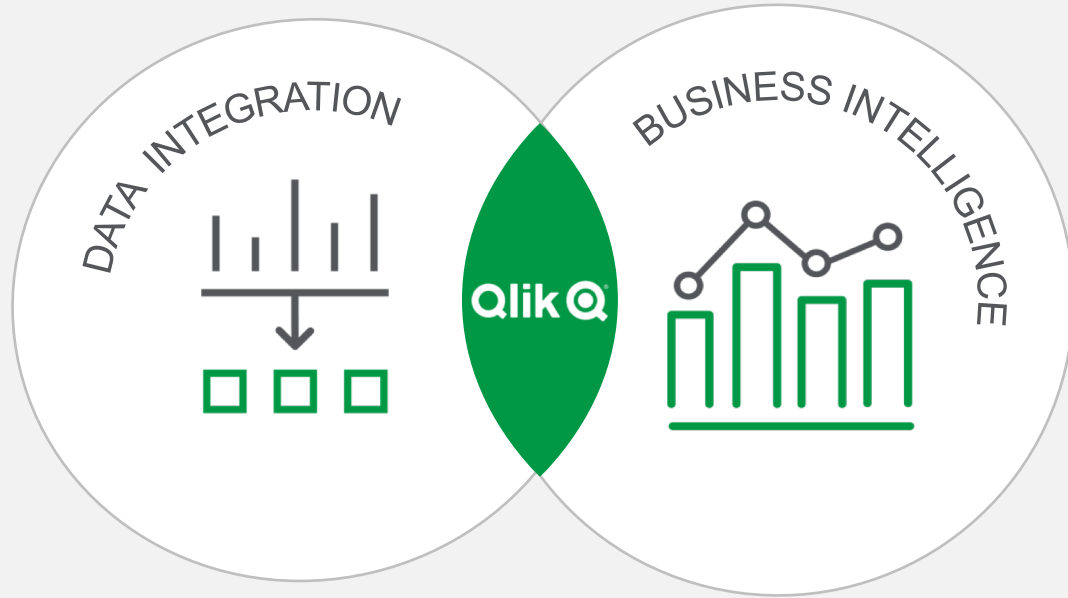
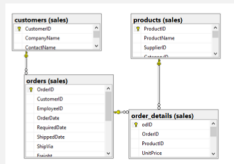
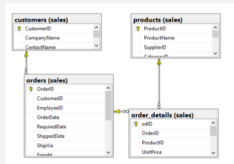
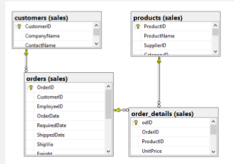


Какие задачи стоят на пути загрузки данных?

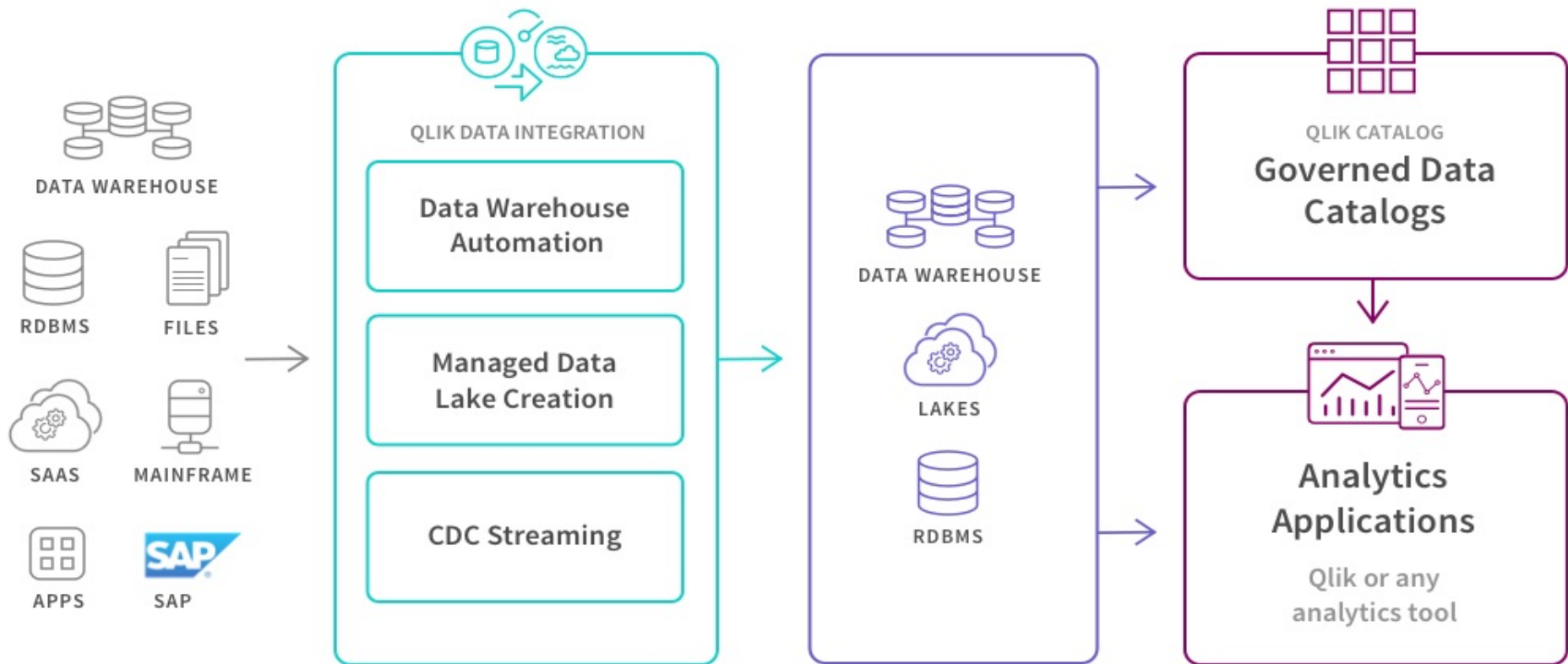
Требования

- Иметь данные, которым можно доверять
- Быстрое изменение и отображение новых данных
- Разгрузить ИТ-отдел
- Без агентов и посредников захват и доставка данных – репликация в режиме реального времени
- Автоматическое сопоставление, создание целевых таблиц и наполнение данными
- Разнообразие сценариев
- Быстрый Time-To-Market
- Не нагружать продуктивную среду
- Синхронизация структуры источников и приёмников
- Доступность аналитики в режиме реального времени
- Гибкая ролевая модель
- Безопасность

Qlik работает с данными



Аналитический конвейер данных Qlik



Начинаем с репликации данных



Qlik Replicate

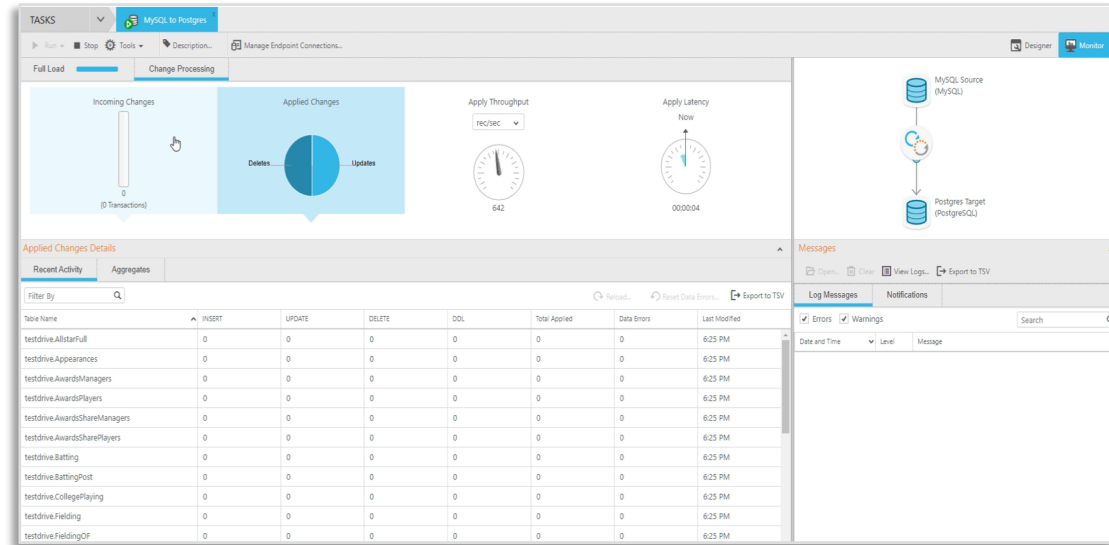
Replicate | CDC Поточковая репликация on-line

Преимущества

- Без агентов и посредников захват и доставка данных – репликация в режиме реального времени.
- Автоматическое сопоставление, создание целевых таблиц и наполнение данными
- Разнообразие сценариев

Выгоды

- Гарантия консистентности данных
- Без влияния на продуктивные системы
- Простота администрирования, Dr&Dr
- Быстрый Time-To-Market
- Низкая стоимость владения

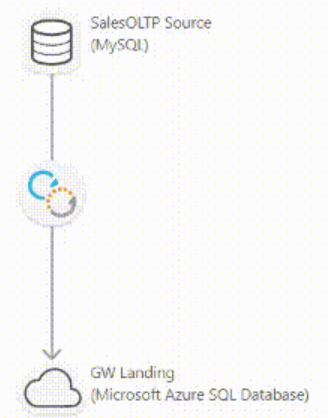


TASKS

[+ New Task...](#)
[Open](#)
[Delete...](#)
[Run](#)
[Skip](#)
[Import Task](#)
[Export Task](#)
[Export to TSV](#)
[Manage Endpoint Connections...](#)
[Details](#)

Filter Tasks

Name	Description	Type	State	Reason
C4DL Sales - Delta Lake demo		Unidirectional	Stopped	User-initiated or t...
C4DL Sales - HDI2		Unidirectional	Stopped	User-initiated or t...
C4DL SFDC - HDI2		Unidirectional	Stopped	User-initiated or t...
C4DW Sales - Azure SQLDW2		Unidirectional	Stopped	User-initiated or t...
C4DW Sales MySQL - AzureDB		Unidirectional	Stopped	User-initiated or t...
C4DW Sales MySQL - Snowflake2		Unidirectional	Stopped	User-initiated or t...
Gen Hosp MySQL - AzureDB		Unidirectional	Stopped	User-initiated or t...
Sales - Databricks Complete		Unidirectional	Stopped	User-initiated or t...
SalesDemo MySQL SQLDW		Unidirectional	Stopped	User-initiated or t...
SFDC - MySQL2		Unidirectional	Stopped	User-initiated or t...



Messages

[Open...](#)
[Clear All](#)
[View Logs...](#)
[Export to TSV](#)

[Log Messages](#)
[Notifications](#)

Errors
 Warnings
 Search

Date and Time	Level	Message
0 Log Messages		

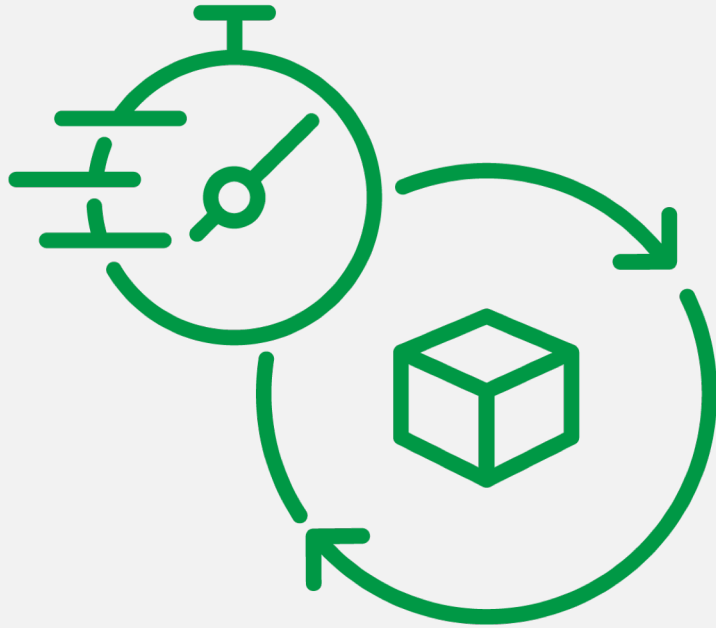


Qlik Replicate

Особенности

- Нет агентов и посредников
- Не нагружает продуктивную среду
- Интуитивно понятный интерфейс
- Реплицирует данные в реальном времени
- Быстрый Time-To-Market
- Не требует глубоких знаний в написании скриптов
- Удобное управление и мониторинг

Эффективное использование реплицированных данных



Qlik Compose

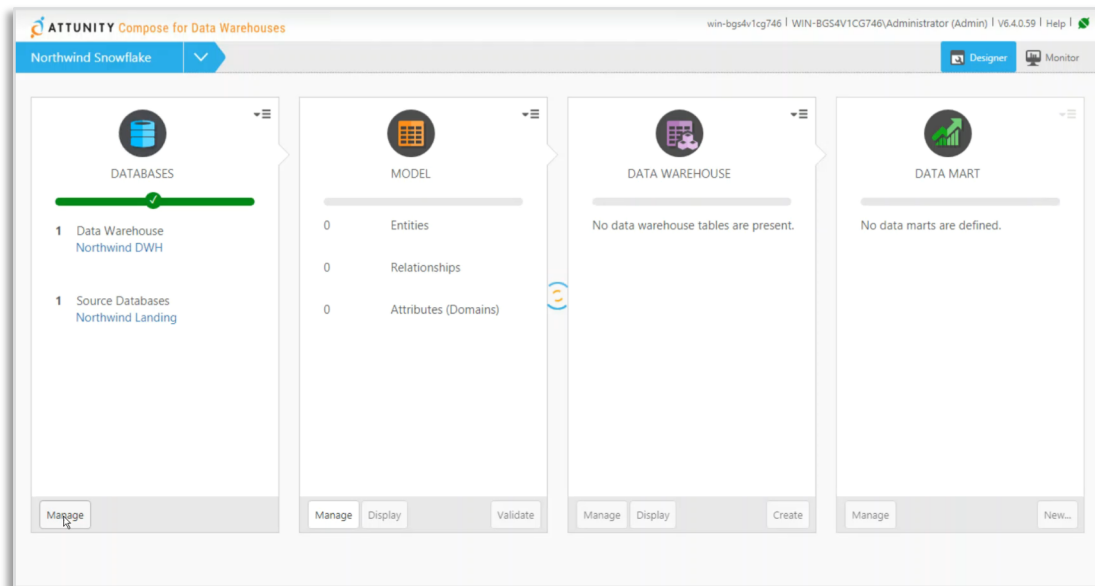
Compose | DWH Автоматизация хранилищ

Преимущества

- Полная автоматизация создания модели данных, её имплементации и документирования.
- Автоматическое сопоставление, создание целевых таблиц и наполнение данными
- Непрерывная on-line репликация

Выгоды

- Минимизация риска, сокращение времени и затрат – минимум кодирования и скриптов.
- КХД строится за час, обновляется минуты.
- Открытый API



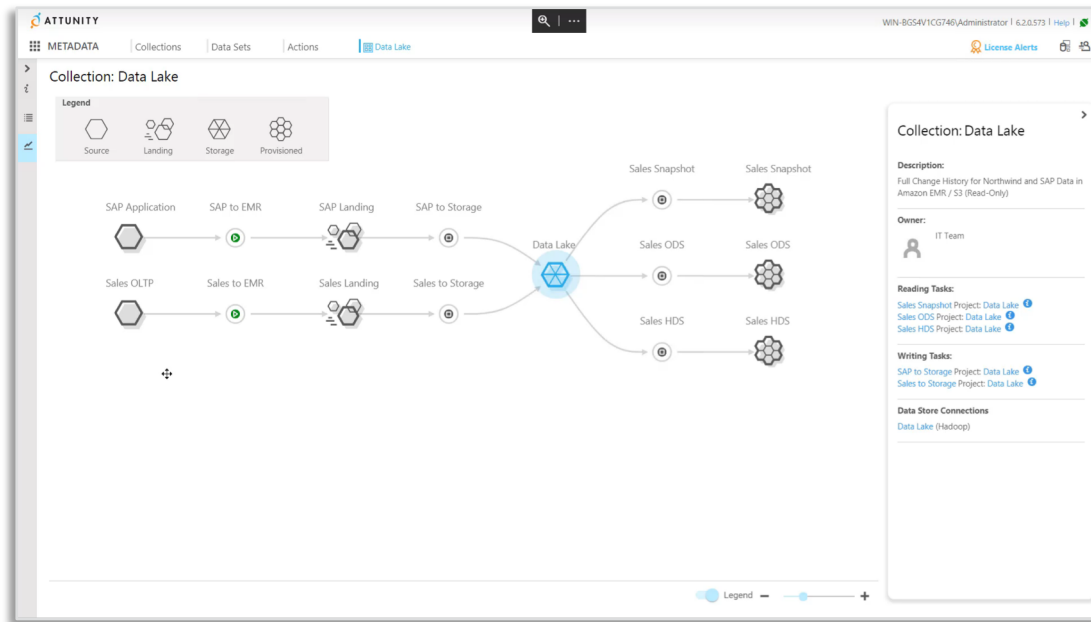
Compose | Data Lake: Создание озер данных

Преимущества

- 100% автоматическое создание таблиц, модели и взаимосвязей в управляемом озере данных
- Синхронизация структуры источников и приёмников
- Непрерывная репликация в режиме реального времени

Выгоды

- Быстрое и простое создание гибких data pipelines
- Отсутствие сложного и дорогого кодирования
- Доступность аналитики в режиме реального времени



DATABASES

- 1 Data Warehouse
EDW (bjonessqldw)
- 1 Source Databases
Sales Source

Manage

MODEL

- 0 Entities
- 0 Relationships
- 0 Attributes (Domains)

Manage Display Validate

WAREHOUSE

These tables are present.

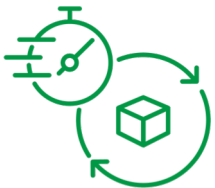
Manage Display Create

DATA MART

No data marts are defined.

Manage New...

- Manage
- Display
- Validate
- Discover...
- Add Date and Time Entities
- Import From ERwin...
- Manage Attributes Domain...
- Reusable Transformations...
- Global Mappings...



Qlik Compose

Особенности

- Быстрое создание DWH и Data Lake
- Автоматизация хранилищ
- Интуитивно понятный интерфейс
- Не требует углубленных навыков в написании скриптов
- Автоматическое создание моделей данных с использованием лучших практик
- Переиспользование имеющихся сущностей
- Быстрое внесение изменений
- Проверка данных на ошибки
- Подготовка данных для Data Mart

Магазин данных

Qlik  **Q** **Catalog**

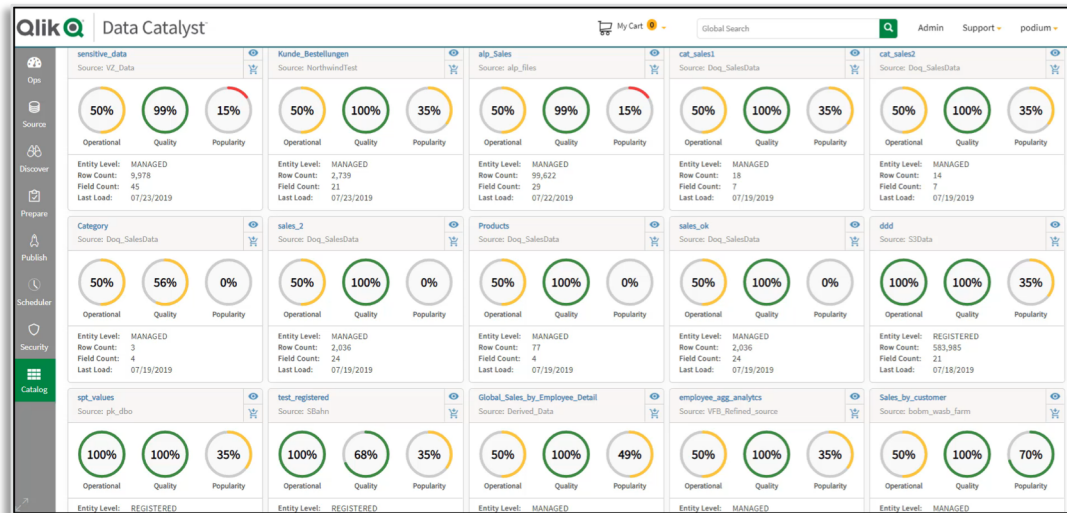
Catalog | Обогащение источников и данных

Преимущества

- Автоматизированное обогащение данных бизнес мета-информацией.
- Каталогизация всех источников + QVD
- Безопасность
- Быстрая публикация
- Гибкая ролевая модель

Выгоды

- Быстрое формирование гипотез на основе данных
- Управление безопасностью
- Снижение нагрузки на IT подразделения.



DATABASES

- 1 Data Warehouse
EDW (bjonesqldw)
- 1 Source Databases
Sales Source

Manage

MODEL

- 10 Entities
- 8 Relationships
- 82 Attributes (Domains)

Manage Display Validate

DATA WAREHOUSE

18 Data Warehouse Tables Present.

Sales Source	1,095 instructions
Sales Source_CDC	1,154 instructions

Manage Display Validate









DATA MART

No data marts are defined.










Manage New...

QDI - Матрица поддержки платформ

ИСТОЧНИКИ

 <p>Базы данных</p>	 <p>Мэнфреймы</p>	 <p>SAP</p>
<ul style="list-style-type: none"> Oracle SQL Server DB2 iSeries DB2 z/OS DB2 LUW MySQL PostgreSQL Sybase ASE Informix ODBC 	<ul style="list-style-type: none"> DB2 z/OS IMS/DB VSAM COBOL Copybooks 	<ul style="list-style-type: none"> ECC ERP CRM SRM GTS MDG S/4HANA <p>(on Oracle, SQL, DB2, HANA)</p>
 <p>Хранилища данных</p>	 <p>Облака</p>	 <p>SaaS</p>
<ul style="list-style-type: none"> Exadata Teradata Netezza Vertica Pivotal 	<ul style="list-style-type: none"> Amazon RDS (SQL Server, Oracle, MySQL, Postgres) Amazon Aurora (MySQL) Amazon Redshift Azure SQL Server M1 (Q1) 	<ul style="list-style-type: none"> Salesforce (Q2)
 <p>Плоские файлы</p>	 <p>Плоские файлы</p>	<ul style="list-style-type: none"> XML JSON С разделителями (например, CSV, TSV)

ПРИЕМНИКИ

 <p>AWS</p>	 <p>Azure</p>	 <p>Google</p>
<ul style="list-style-type: none"> RDS (MySQL, Postgres, MariaDB, Oracle, SQL Server) Aurora (MySQL, Postgres) S3 EMR Kinesis Redshift <hr/> <ul style="list-style-type: none"> Snowflake (Q1) Databricks (Q2) 	<ul style="list-style-type: none"> DBaaS (SQL DB) DBaaS (MySQL, Postgres) ADLS BLOB HDInsight Event Hub SQL DW <hr/> <ul style="list-style-type: none"> Snowflake (Q1) Databricks (Q2) 	<ul style="list-style-type: none"> Cloud SQL (MySQL, Postgres) Cloud Storage Dataproc PubSub (Q2) Big Query (Q2)
 <p>Озёра данных</p>	 <p>Хранилища данных</p>	 <p>Базы данных</p>
<ul style="list-style-type: none"> Hortonworks Cloudera MapR Amazon EMR Azure HDInsight Google Dataproc 	<ul style="list-style-type: none"> Exadata Teradata Netezza Vertica Sybase IQ SAP HANA Microsoft PDW 	<ul style="list-style-type: none"> Oracle SQL Server DB2 LUW MySQL PostgreSQL Sybase ASE Informix MemSQL
 <p>Потоковые данные</p>	 <p>SAP</p>	 <p>Плоские файлы</p>
<ul style="list-style-type: none"> Kafka Amazon Kinesis Azure Event Hubs MapR Streams 	<ul style="list-style-type: none"> HANA 	<ul style="list-style-type: none"> С разделителями (например CSV, TSV)

Спасибо!

