



Краткое описание ПО
BI.Qube MetaMasterData

© 2023 ООО «АйТи Про»

КРАТКОЕ ОПИСАНИЕ ПО BI.QUBE METAMASTERDATA

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
ГЛОССАРИЙ.....	3
1. ЦЕЛИ И НАЗНАЧЕНИЕ METAMASTERDATA	5
2. КРАТКОЕ ОПИСАНИЕ METAMASTERDATA	6
2.1. Описание архитектуры	6
2.2. Типовой сценарий использования.....	8
2.3. Функциональные возможности	9

ВВЕДЕНИЕ

MetaMasterData относится к классу MDM-систем и представляет собой решение по управлению основными данными, которое предоставляет надежный способ управления потоком данных через корпоративные ИТ-системы. Главная цель которого – обеспечить единство представления массивов данных в информационных системах. Кроме того, такой тип решений позволяет решить проблемы несоответствия, дублирования и несопоставимости данных.

Управление основными данными включает в себя действия, предпринимаемые организацией для нахождения и определения не транзакционных списков данных с целью компиляции управляемых главных списков. Основные данные отличаются от транзакционной большей стабильности, меньшими объемами и усложненной структурой.

Часто основные данные называют мастер-данные или эталонными данными, а также нормативно-справочной информацией.

В документе приведено описание компонента и принципы работы с ним.

Изучение данного документа позволит понять принцип работы компонента.

ГЛОССАРИЙ

1.	MetaMasterData	Инструмент VI.Qube, предназначенный для создания «идеального источника» данных, используя который, можно строить качественные аналитические срезы и повышать эффективность бизнес-процессов.
2.	Мастер-данные	Основные данные организации со схожим составом и атрибутами, хранящиеся в виде линейных или иерархических справочников.
3.	Master Data Management	Управление мастер-данными.
4.	НСИ	Нормативно-справочная информация – это фиксированные, исходно наполняемые и изменяемые только в редких случаях справочники (общероссийские или внутренние классификаторы, справочники стран, регионов, валют и т. д.).
5.	MDM-система	Это совокупность процессов и инструментов для непрерывного определения и управления мастер-

		данными и нормативно-справочной информацией в организации
6.	ИС	Информационные системы.
7.	ETL	Extract Transform Load – конвейер.
8.	Домен	Сущность с множеством допустимых значений.
9.	Золотая запись	— тезис в сфере управления мастер-данными. Он представляет собой формирование единственного, в наибольшей степени достоверного, истинного, непротиворечивого понятия об объектах предметной области. Это могут быть клиенты, контрагенты, чаще всего в сфере ритейл — это товары. Золотая запись часто именуют «единой версией правды», к которому могут по требованию обращаться все заинтересованные лица, если хотят убедиться, что находящиеся у них данные являются верными.
10.	Транзакционные данные	это данные, получаемые посредством CRUD операций.
11.	Структурированные данные	это данные, имеющие формально-определенную структуру.
12.	Полуструктурированные данные	это данные, не имеющие строго определенной структуры, однако, обладающие формальным описанием в виде тегов и/или маркеров.
13.	Неструктурированные данные	это данные не имеющие структуры и тегов/маркеров.
14.	Метаданные	это данные, используемые для описания какой-либо дополнительной информации о мастер-данных, к примеру конфигурационные данные.
15.	Историчность	это привязка модели к определенному времени, на протяжении которого данная модель не изменялась не изменялись.
16.	Версионность	модель, действующая в определенный период времени, сохраняющая модель данных или отдельных объектов данных: атрибутов, сущностей, справочников, связей.

17.	Модель данных	это абстрактное, самодостаточное, логическое определение объектов, с которой взаимодействует пользователь.
18.	Кроссплатформенность	способность программного обеспечения работать с несколькими аппаратно-программными комплексами
19.	Система управления базой данных (СУБД)	система, обеспечивающая контроль и управление данными, позволяющая выполнять различные административные операции.
20.	CRUD операции	операции по созданию, чтению, обновлению и удалению данных.
21.	Бизнес-правила	правила, используемые для валидации данных в соответствии с бизнес-логикой.
22.	Data Vault	Набор связанных между собой нормализованных таблиц, ориентированных на хранение детализированной информации с возможностью отслеживания происхождения данных и поддерживающих одну или несколько областей бизнеса.

1. ЦЕЛИ И НАЗНАЧЕНИЕ METAMASTERDATA

Компонент MetaMasterData, входит в состав системы аналитического корпоративного хранилища BI.Qube и предназначен для быстрого построения и управления мастер-данными или нормативно-справочной информацией (НСИ) в организации.

Главная цель компонента — обеспечить единство представления массивов данных во всех информационных системах, посредством создания «золотой записи», то есть целостного и всестороннего представления о мастер-сущности и взаимосвязях, эталона мастер-данных, который используются всем предприятием, а иногда и между предприятиями для упрощения обмена информацией.

Компонент MetaMasterData предназначен для автоматического создания новых сущностей интегрируемых в существующую реляционную базу данных, с сохранением имеющейся модели данных. Компонент реализует стандартную функциональность по добавлению новых данных,

редактированию и удалению данных из базы данных, что позволит контролировать данные в базе данных, обогащать их новыми данными, связывать с имеющиеся данные с новыми.

Основным назначением является возможность создавать правила обработки данных: изменение типов данных, создание новых данных на основе имеющихся (новые поля, сущности), создание масок ввода и отображения данных, создания золотой записи. При этом сохраняется все история произведенных преобразований, что позволят всегда вернуться к исходному состоянию редактируемой базы данных.

Компонент MetaMasterData включен в качестве расширения в компонент MetaVault и без него автономно работать не может.

2. КРАТКОЕ ОПИСАНИЕ METAMASTERDATA

2.1. Описание архитектуры

В основу построения системы положена трехуровневая архитектура, которая включает в себя отдельные сервисы.

Такая архитектура позволяет улучшить надежность и устойчивость системы, а также обеспечить взаимодействие с каждым сервисом по отдельности через REST-API.

Программный комплекс состоит из трех уровней: клиентская часть, серверная часть, часть хранилища данных.

Клиентская часть представляет собой веб-приложение, которое взаимодействует с серверной частью при помощи REST-API. Для разработки клиентской части использовались следующие технологии: TypeScript, библиотека React js, Ant Design, HTML и CSS.

Серверная часть программного комплекса включает в себя приложение-сервер, которое является кроссплатформенным и может быть развернуто как на локальной машине, так и в облаке. Для разработки серверной части использовались технологии .NET 6 CORE.

Использование подхода хранения данных Data Vault (реализованного в компоненте MetaVault) имеет ряд преимуществ. Например, она обладает высокой масштабируемостью и способностью поддерживать сложные бизнес-правила. Также она обеспечивает быстрый доступ к данным, что важно для систем, где происходит обработка больших объемов информации.

Еще одно преимущество заключается в том, что подход хранения данных Data Vault обеспечивает надежность хранения данных, поскольку она предусматривает использование специальных таблиц для хранения истории

изменений данных. Это позволяет восстанавливать данные в случае ошибок или сбоев в системе.

Взаимодействие с компонентом MetaVault через NuGet пакет также обеспечивает простоту управления и мониторинга системы. Все необходимые инструменты для администрирования и настройки системы уже включены в состав NuGet пакета, что упрощает работу с ним.

Кроме того, NuGet пакет обеспечивает совместимость с другими платформами и языками программирования, что расширяет возможности использования системы.

В целом, использование стороннего программного комплекса Data Vault для хранения мастер-данных и метаданных в программном комплексе интеграции и управления мастер-данных является эффективным и надежным решением, которое обеспечивает высокую отказоустойчивость и масштабируемость системы. Кроме этого, отпадает необходимость интеграции мастер данных в хранилище данных, так как интеграция реализована уже на уровне ядра хранилища.

Рассмотрим подсистемы подробнее.

- **Подсистема безопасности** – позволяет управлять пользователями, группами, а также их правами.
- **Подсистема работы с иерархическими моделями** – позволяет создавать, редактировать и удалять домены, модули и сущности, а также их атрибуты, правила валидации. Состоит из следующих модулей.
- **Подсистема версионирования** – позволяет версионировать структуру сущностей, атрибутов и данных, а также отменять уже примененные изменения.
- **Подсистема работы с данными: Модуль управления данными** – позволяет управлять данными сущностей, включая работу с атрибутами на доменную сущность.
- **Подсистема работы с метаданными: Модуль управления метаданными** – позволяет управлять метаданными профилей, доменов, сущностей, представлений, а также атрибутов.
 - **Подсистема валидации данных** – позволяет управлять правилами валидации данных посредством визуального

интерфейса и строить правила валидации по принципу ЕСЛИ-ТО-ИНАЧЕ, а также подключать существующие CUSTOM правила с хранилища; позволяет валидировать внесенные данные в сущность посредством созданных бизнес-правил.

- **Подсистема работы с «Золотой записью»** – позволяет настраивать приоритет источников данных, а также получать «золотой» элемент посредством слияния данных об одной записи с разных источников; слияния записей осуществляется при помощи специального алгоритма, в котором находятся записи об одной сущности и сливаются в «золотую» запись.

Схематично Архитектура программного комплекса управления мастер-данными изображена на рисунке ниже.



Рисунок 1. Организационная схема компонента MetaMasterData

2.2. Типовой сценарий использования

Современные требования к MDM-системам превышают их стандартные функциональные возможности, на сегодня требуется система, которая бы предоставила средства управления справочниками, получаемыми из имеющихся информационных систем, и попутно позволила бы привести все справочники к единому формату. После такой обработки стало бы возможным использовать единые средства интерпретации и анализа данных.

Сценарий использования MMD:

1. Формируем эталонные справочники.
2. Внести данные.

3. Получить данные справочников из системам-источников. Для этого используем средства ETL из инструмента BI.Qube.
4. Производим маппинг, т.е. устанавливаем соответствие записей, полученных от систем-источников с записями эталонных справочников.
5. Предоставляем системам-потребителям средства для получения данных из эталонных справочников и маппингов.

2.3. Функциональные возможности

1. Обогащать и вносить изменения в загружаемые из систем-источников данные с использованием визуального интерфейса:
 - a. Визуальный интерфейс для оператора - представляет из себя страницу на которой видна таблица с возможностью обнулять, изменять и удалять данные из нее.
 - b. Визуальный интерфейс для администратора - представляет из себя страницу, на которой он может изменять структуру данных, изменять атрибутивный состав таблиц, создавать бизнес-правила и настраивать валидацию данных.
2. Создание бизнес-правил:
 - a. Отдельный интерфейс настройки бизнес-правил
 - b. Обогащение новыми данными (добавление нового источника, ввод вручную) - к справочникам, которые добавлены из других источников, необходимо зайти в визуальный интерфейс администратора и добавить новые поля, в дальнейшем при работе оператора, он сможет обогащать справочник ново добавленными атрибутами.
3. Создание золотой записи:
 - a. Создается в полуавтоматическом режиме необходимо настраивать связи между справочниками.
4. Управление правами доступа:
 - a. Для разграничения возможностей системы (для примера, рядовой сотрудник не должен иметь доступа к редактированию метаинформации).
5. Создание уникальных наборов данных для пользователей:
 - a. Разграничение видимости справочников по профилям, конкретные профили имеют общую видимость либо конкретно привязываются к пользователю. Администратор или Пользователь сам настраивает вхождение сущностей в профиль
6. Смотреть историю изменения данных, управлять версионностью данных:
 - a. Просмотр полной истории изменений гарантирует прозрачность ведения справочников и контроля изменений данных.
7. Возможность вносить изменения в справочники через подключение Microsoft Excel + поддержка сводных таблиц.