




Проект:	«OTRi.DG»
Документ:	Описание функциональных характеристик
Дата:	16.07.2024
Версия:	1.0


«OTRi.DG»

Описание функциональных характеристик

	Проект:	«OTRi.DG»
	Документ:	Описание функциональных характеристик
	Дата:	16.07.2024
	Версия:	1.0

СОДЕРЖАНИЕ

1	Термины и сокращения.....	3
2	Функциональные характеристики	3
2.1	Управление метаданными	3
2.2	Контроль качества данных	5
3	Список изменений.....	8

	Проект:	«OTRi.DG»
	Документ:	Описание функциональных характеристик
	Дата:	16.07.2024
	Версия:	1.0

1 Термины и сокращения

Термины и сокращения	Определения
CASE-средство	Методы и технологии, которые позволяют проектировать различные информационные системы (в частности, базы данных) и автоматизировать их создание
Data lineage	Процесс отслеживания перемещения данных от их источника до конечных точек, включая все промежуточные этапы трансформации и обработки данных
OTRi.DG	Программное обеспечение, предназначенное для управления метаданными и осуществления контроля качества данных
SQL	Декларативный язык программирования, применяемый для создания, модификации и управления данными в реляционной базе данных, управляемой соответствующей системой управления базами данных
Бизнес-гlossарий	Словарь терминов, связанных с цифровыми данными, которыми оперирует организация, описанный на понятном языке
Метаданные	Зарегистрированная информация о данных
Модель метаданных	Множество логически связанных объектов метаданных
Объект модели метаданных	Метаданные объекта внешней системы (базы данных, таблицы, дашборда и т.д.)
Парсинг	Синтаксический анализ данных с последующим их структурированием в необходимый вид


2 Функциональные характеристики

ПО «OTRi.DG» обеспечивает выполнение функций, перечисленных в разделах 2.1 – 2.2 настоящего документа.

2.1 Управление метаданными

В части управления метаданными ПО «OTRi.DG» обеспечивает выполнение следующих функций:

- поддержка ведения бизнес-гlossария, включая:
 - а) ведение бизнес-гlossария с необходимыми атрибутами;
 - б) настройку дополнительных атрибутов бизнес-гlossария администратором (пользователем модуля, имеющим соответствующие права);
 - в) предоставление доступа на просмотр бизнес-гlossария через веб-интерфейс работникам на основе ролевой модели доступа;

	Проект:	«OTRi.DG»
	Документ:	Описание функциональных характеристик
	Дата:	16.07.2024
	Версия:	1.0

г) ведение метаданных таблиц и отдельных атрибутов, включая метаданные, вводимые пользователем;

- настройка администратором структуры и состава метаданных для объектов;
- настройка администратором через интерфейс рабочего процесса на изменение метаданных: добавление, изменение и удаление метаданных для таблиц и атрибутов и предоставления интерфейса согласования изменений для лиц, участвующих в рабочем процессе;

- ведение и просмотр моделей метаданных, включая модель данных БД;
- поиск по модели метаданных;
- просмотр lineage-диаграмм (цепочек происхождения данных) по модели метаданных;

- просмотр из интерфейса ПО исходных кодов процедур на языке SQL, которые формируют модель данных, для визуализации логики расчета, а также иных объектов баз данных;

- отслеживание корректности отработки алгоритмов загрузок, включая:

- а) фиксацию атрибутов элементов модели данных, которым не сопоставлены термины из бизнес-гlossария;

- б) настройку уведомлений пользователю в случае срабатывания проверок (атрибуты элементов модели метаданных, включая модель данных БД, не сопоставлены терминам бизнес-гlossария, сработали проверки управления качеством данных);

- просмотр проверок качества данных, связанных с объектами модели метаданных;
- просмотр информационных сообщений и новостей, относящихся к объектам модели метаданных;

- обеспечение контроля состояния объектов:


- а) сведения о текущем статусе объектов;

- б) сведения о последнем обновлении объекта.

Графический интерфейс пользователя ПО «OTRi.DG» позволяет выполнять следующие действия:

- просмотр структуры данных БД (включая структуры схем, таблиц, колонок) с возможностью фильтрации неразмеченных объектов БД;

- просмотр бизнес-гlossария;

	Проект:	«OTRi.DG»
	Документ:	Описание функциональных характеристик
	Дата:	16.07.2024
	Версия:	1.0

- просмотр сведений о проверках данных и переход к соответствующим проверкам;
- просмотр data lineage (цепочка происхождения данных);
- просмотр информационных сообщений и новостей, относящихся к объектам модели метаданных.

2.2 Контроль качества данных


В части контроля качества данных ПО «OTRi.DG» обеспечивает выполнение следующих функций:

- регулярная массовая автоматизированная проверка данных в процессе их загрузки из систем-источников, обработки и/или передачи согласно настроенным процедурам и разработанным контролям с:

- а) выявлением отклонений;
- б) возможностью передачи сведений о результатах проверки путём отправки уведомлений по электронной почте;
- в) постановкой задач на расследование пользователям с соответствующими компетенциями в случае ошибок загрузки/проверки данных;
- г) управлением из графического веб-интерфейса с помощью настраиваемого CASE-средства.

В части подключения источников в ПО «OTRi.DG» реализованы следующие функции:

- возможность подключения системы-источника данных (для анализа качества данных);
- создание, изменение, удаление специальных функций для использования в процедурах проверки качества данных, включая возможность использования настроечных параметров, словарей и применения следующих функций: математических, логических, статистических, агрегирования, нечеткого поиска, композиции функций (например, функции сравнения распределений данных);
- создание, изменение, удаление процедуры проверки качества данных посредством встроенного CASE-средства и с помощью документированного языка программирования/запросов для целей использования для проверки качества данных, предоставляемых сторонним системам, с возможностями в рамках процедуры проверки качества данных:

	Проект:	«OTRi.DG»
	Документ:	Описание функциональных характеристик
	Дата:	16.07.2024
	Версия:	1.0

- а) применения условий ограничения (фильтра) для формирования выборок данных;
- б) осуществления декомпозиции данных на составные части по заданным или встроенным правилам (парсинг);
- в) ограничения состава атрибутов, необходимых для проверок качества данных;
- г) анализа в рамках одной проверки качества данных нескольких объектов метаданных (например, проверка заполненности сразу нескольких колонок таблицы);
- д) одновременного (параллельного) выполнения нескольких проверок качества данных на одной выборке данных;
- е) использования в процедуре заранее созданных специальных функций для настройки проверки качества данных.

Настройка запуска процедур проверки качества данных включает в себя следующие возможности:


- настройка периодичности/расписания/условий (по флагу готовности в таблице, по факту появления файла в папке, по определенному результату SQL-запроса) запуска проверок качества данных;
- настройка последовательности запуска процедур проверок качества данных относительно процесса импорта данных источников.

Запуск и выполнение процедур проверки качества данных включают в себя следующие возможности:

- автоматический запуск проверок качества данных с указанием настроек (например, условия отбора данных: за отчетный период, за конкретный день, по одному контролю, пакету контролей и пр.);
- запуск проверок качества данных с возможностью параллельной работы процедур, мониторинг функционирования и отработки проверок качества данных;
- автоматическое информирование о сбоях запуска/отработки и результата отработки процедур;
- сохранение результатов проверок качества данных в реляционной базе данных.

Детальная информация о процедуре проверки качества данных включает в себя:

- просмотр основной информации о процедуре проверки качества данных;
- просмотр предыдущих состояний процедуры проверки качества данных;

	Проект:	«OTRi.DG»
	Документ:	Описание функциональных характеристик
	Дата:	16.07.2024
	Версия:	1.0

- просмотр и выгрузку информации о процедурах проверки качества данных;
- переход к просмотру детальных результатов работы процедуры проверки качества данных.

Просмотр списков всех инцидентов качества данных осуществляется с возможностями:

- отображения большого количества инцидентов наглядно фрагментарно с возможностью навигации между фрагментами (например, скролл, пагинация или иные механизмы);
- перехода к просмотру детальной информации об инциденте качества данных.

Просмотр детальных результатов работы процедур проверки качества данных в табличном виде происходит с детализацией до:

- даты/времени выполнения проверки;
- проверяемого объекта с возможностью:
 - а) отражения статуса исправления ошибки, определенной по инциденту качества данных;
 - б) отражения метрик качества данных по каждой проверке;
 - в) перехода к связанным с результатом проверки качества данных уведомлениям, рекомендациям, заявкам, задачам.


Просмотр агрегированных результатов работы процедур проверки качества данных в текстовом/табличном виде осуществляется с детализацией до атрибутов.

Автоматическое информирование пользователя (по подписке) о событиях запуска/отработки и результата отработки процедуры проверки качества данных происходит с возможностью:

- отключения от/подключения к подпискам на различные события на уровне администратора, пользователя или группы пользователей;
- настройки регламента уведомления по подписке;
- настройки каналов получения информации в зависимости от видов событий, пользователей и проверок качества данных, включая корпоративную электронную почту.

В рамках автоматического информирования пользователя реализована возможность получения уведомления о:

- завершении процесса по подписанной проверке качества данных;
- возникновении ошибки при выполнении проверки качества данных.

	Проект:	«OTRi.DG»
	Документ:	Описание функциональных характеристик
	Дата:	16.07.2024
	Версия:	1.0

3 Список изменений

Версия	Дата	Внесенные изменения	Исполнитель