

Система контроля качества данных SDP Data Quality

Руководство пользователя

Содержание

1. Руководство пользователя	
1.1 UI Реестр требований	3
1.2 UI Конструктор проверок	10
1.3 UI Результаты проверок.....	16

1 Руководство пользователя

1.1 UI Реестр требований

Для того, чтобы получить всю информацию по интересующему требованию необходимо воспользоваться UI Реестр требований.

Реестр требований состоит из нескольких вкладок: заявки, выгрузка реестра и справочники.

Вкладка «Заявки» является основным инструментом разработчика для управления метаданными проверок. Вкладка состоит из двух уровней:

- Управление метаданными заявки;
- Управление метаданными требований в заявке.

Заявки

ID	Подразделение заказчика	ID Заявки	ФИО заказчика	Требований	Дата создания	Дата редактирования
38	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR1	1	26.11.2020	27.11.2020
43	Блок *GR, правовые вопрос...	TEST1	TEST2	1	26.11.2020	26.11.2020
81	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR555	1	27.11.2020	27.11.2020
82	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR555	1	27.11.2020	27.11.2020
83	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR555	0	27.11.2020	
84	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR1	0	27.11.2020	
85	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR1	0	27.11.2020	
86	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR1	0	27.11.2020	
87	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR1	0	27.11.2020	
88	Блок *GR, правовые вопрос...	CONSTRUTO...	CONSTRUTOR1	0	27.11.2020	

Вкладка «Выгрузка из реестра» представляет собой плоскую таблицу, содержащую в себе всю информацию по требованиям и проверкам качества данных.

Выгрузка реестра

Q* Поиск: Все поля Найти Действия *

ID	ID заявки	АС реализации проверки	Область данных	Подразделение заказчика	ФИО заказчика	ФИО Владельца данных	Система-источник	ID требования
1	CONSTRUT...	ОД	Реплика ФСК	Блок *GR, правовые вопрос...	CONSTRUTOR1		Витрина Статус-члено...	CONSTRUTOR1_1
2	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			
3	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			
4	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			
5	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			
6	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			
7	CONSTRUT...	ОД	Реплика CRM Кор...	Блок *GR, правовые вопрос...	CONSTRUTOR1		Витрина Статус-члено...	CONSTRUTOR123...
8	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			
9	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			
10	CONSTRUT...			Блок *GR, правовые вопрос...	CONSTRUTOR1			

Вкладка «Справочники» содержит в себе информацию по всем справочникам, значения которых используются при формировании проверки и их метаданных.

- Справочники ▾
- Область данных
- Оргструктура
- Показатели
- Система-источник
- Статус требования
- Сущности
- Атрибуты
- Типы требований
- Бизнес-задача

Атрибутный состав реестра требований описан в таблице ниже.

Наименование поля	Описание поля
ID заявки	Уникальный идентификатор заявки (группы требований к качеству данных)
Подразделение Потребителя заказчика	Наименование подразделения Заказчика требования к качеству данных
ФИО заказчика	ФИО Потребителя данных, в интересах которого реализуются проверки качества данных
ID требования	Уникальный идентификатор требования к качеству данных
Область данных	Наименование области данных, в которой хранятся проверяемые данные (реплика, витрина)
Бизнес-задача	Бизнес-задача решаемая Заказчиком, для целей которых создаются проверки качества данных и осуществляется мониторинг результатов
Подразделение Владельца данных	Наименование подразделения, ответственного за данные (Владельца данных)
ФИО Владельца данных	ФИО ответственного за данные

Система источник	Система, данные которой проверяются напрямую, либо в реплике
Бизнес-описание сущности	Бизнес-наименование проверяемой сущности (таблицы)
Бизнес-описание атрибута	Бизнес-наименование проверяемого атрибута (поля в таблице)
Наименование БД	Техническое наименование базы данных, данные которой проверяются
Наименование сущности в БД	Техническое наименование таблицы, данные которой проверяются
Наименование атрибута в БД	Техническое описание атрибута а проверяемой таблице
Тип требования	Тип проверки качества данных (обязательность, уникальность, ссылочная целостность и т.д.)
Бизнес-описание требования к качеству данных	Бизнес-описание алгоритма проверки
Условие отбора проверяемых записей	Условие, которое накладывается на проверяемую таблицу в целях сокращения выборки проверяемых данных
Показатель	Качественная или количественная оценка соответствия определенного набора данных критериям оценки качества данных.
Пороговое значение показателя	Значение показателя, при достижении которого данные считаются непригодными к использованию
Сценарий	Сценарий определения превышения пороговых значений показателей
Краткое описание требования (для dashboard)	Описание требования, которое отображается в отчетах
Статус требования	Статус реализации требования к качеству данных
Дата регистрации	Дата, когда данное требование к качеству данных поступило
Код проверки	Идентификатор проверки качества данных
Код показателя	Идентификатор метрики
Статус показателя	Статус метрики (включена, отключена)

Фильтрация по полям производится нажатием на поле, на которое необходимо наложить фильтр и указанием значения или его части для поиска.

Заявки

ID	Подразделение заказчика	ID Заявки	ФИО заказчика	Требований	Дата создания	Дата редактирования
38	Блок "GR, правовые вопрос...			1	26.11.2020	27.11.2020
43	Блок "GR, правовые вопрос...			1	26.11.2020	26.11.2020
81	Блок "GR, правовые вопрос...	CONSTR		1	27.11.2020	27.11.2020
82	Блок "GR, правовые вопрос...	CONSTRUTOR5553	CONSTRUTOR555	1	27.11.2020	27.11.2020
83	Блок "GR, правовые вопрос...	CONSTRUTOR55533	CONSTRUTOR555	0	27.11.2020	
84	Блок "GR, правовые вопрос...	CONSTRUTOR13	CONSTRUTOR1	0	27.11.2020	

Есть два способа получения информации по интересующему требованию: через «Заявки» и через «Выгрузка из реестра».

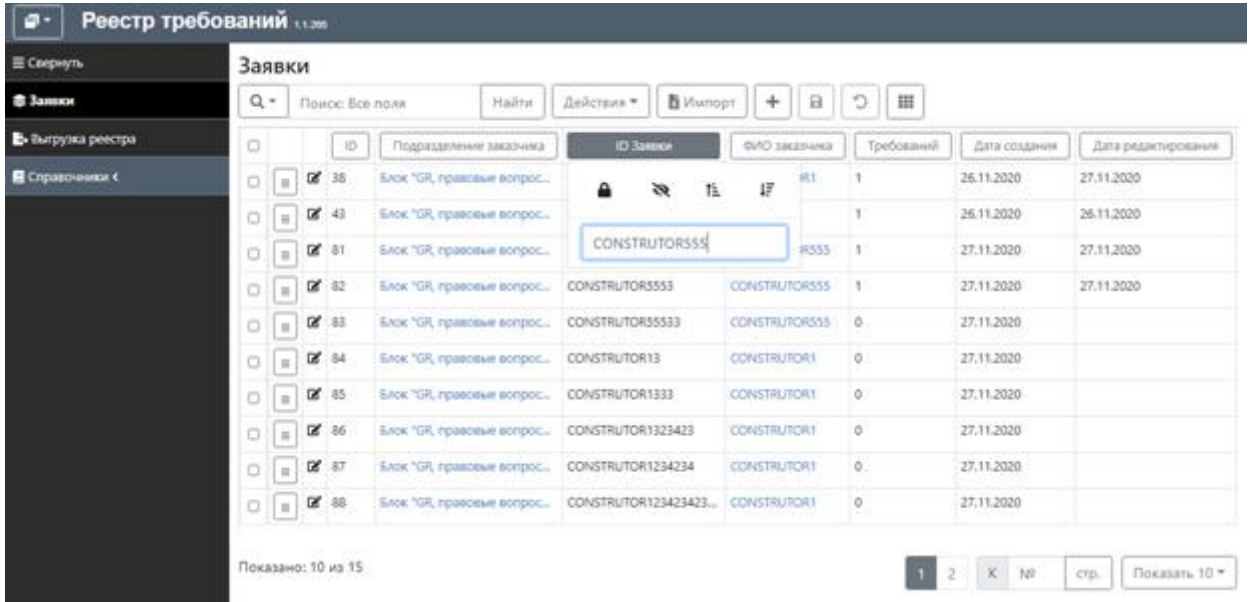
Рассмотрим оба способа.

Первый способ – через вкладки «Заявки» и «Детали по заявке».

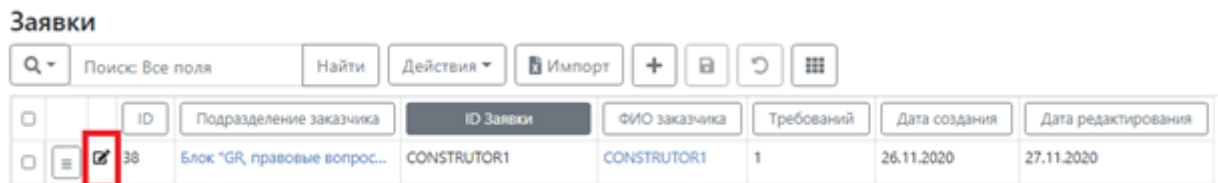
Для того, чтобы найти информацию по определенному требованию, необходимо знать ID заявки, в которую входит это требование. Зная ID требования можно просто определить ID заявки. ID требования формируется по следующему принципу:

[ID заявки]_[Номер требования п/п внутри заявки]

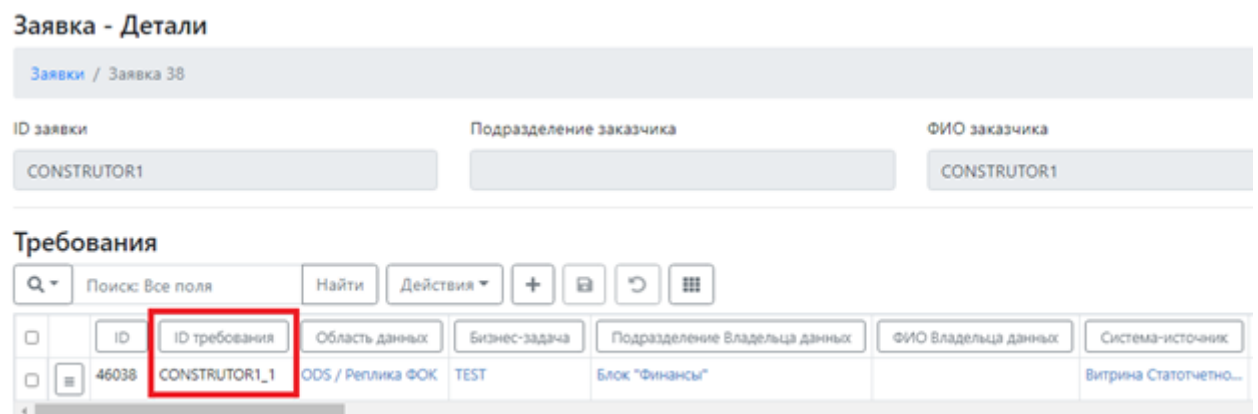
1. Заходим на вкладку «Заявки».
2. Накладываем фильтр на поле «ID заявки» с указанием ID, которое интересует.



3. Нажимаем на кнопку редактирования, чтобы переместиться в область «Детали по заявке».



4. Находим поле «ID требования» накладываем фильтр или же в строку поиска вводим значение для поиска.



Второй способ – через вкладку «Выгрузка их реестра».

1. Заходим на вкладку «Выгрузка из реестра».
2. Находим поле «ID требования» накладываем фильтр или же в строку поиска вводим значение для поиска.

Выгрузка реестра

Одним из компонентов модуля "Реестр требований" являются "Справочники".

В Сервисе Контроля качества данных справочники необходимы для создания проверок и конфигурирования их метаданных.

Переход на страницу с тем или иным справочником осуществляется через UI Реестр требований. В UI Реестр требований в левом углу экрана есть вкладка "Справочники". При нажатии на нее раскроется список со всеми доступными справочниками.

Наполнение, редактирование и актуализацию справочников разработчики проверок качества данных делают самостоятельно.



Корректность заполнения справочников напрямую связана с работоспособностью проверок, правильному формированию метаданных по проверкам, получению правильных результатов по проверкам качества данных.

Существует 9 справочников:

1. Область данных;
2. Оргструктура;
3. Показатели;
4. Система-источник;
5. Статус требования;
6. Сущности;
7. Атрибуты;
8. Типы требований;
9. Бизнес-задача.

Для добавления новой строки в справочник нужно нажать на кнопку "Добавить". Появится новая, выделенная цветом, строка. Те ячейки, которые выделены красным уголком - обязательны к заполнению. Заполнение значений происходит либо выбором значения из справочника, либо свободным текстом, в зависимости от специфики поля.

Для редактирования содержимого ячейки необходимо два раза нажать на нее.

Для сохранения внесенных изменений нужно нажать кнопку "Сохранить" . Если Вы хотите сбросить не сохраненные изменения нажмите на кнопку "Сбросить" чё.

Рассмотрим содержание и назначение каждого справочника в отдельности.

Справочник "Область данных"

Справочник Области данных содержит в себе список баз данных и их описания, атрибуты которых планируется проверять. Наличие записи по области данных, атрибуты которой планируется проверять, в текущем справочнике является обязательным условием для создания проверки качества данных, а также корректного создания новых записей в справочнике "Сущности".

Наименование поля	Описание поля	Тип заполнения
Наименование области данных	Бизнес-наименование базы данных	Свободное
Дата редактирования	Дата редактирования записи	Автоматическое
Дата создания	Дата создания записи	Автоматическое

Наименование АС	Наименование хранилища данных	Свободное
Соединение	Соединение к хранилищу	Свободное
Код области данных	Уникальный идентификатор области данных	Свободное

Справочник "Оргструктура"

Справочник "Оргструктура" представляет собой справочник подразделений, которые являются участниками процесса контроля качества данных. Заполнение данного справочника является обязательным условием для корректного создания метаданных по проверке качества данных.

Наименование поля	Описание поля	Тип заполнения
ID	Уникальный идентификатор записи в справочнике	Автоматическое
Сокращенное наименование	Аббревиатура подразделения	Свободное
Полное наименование	Полное наименование подразделения	Свободное
Дата создания	Дата создания записи	Автоматическое
Дата редактирования	Дата редактирования записи	Автоматическое

Справочник "Показатели"

Справочник содержит в себе типы показателей и их описание. По сути, показатель - это правило расчета метрики качества данных. По умолчанию, тип показателя - доля. Т.е. метрика считается, как доля корректных записей от общего числа проверенных записей.

Наименование поля	Описание поля	Тип заполнения
ID	Уникальный идентификатор записи в справочнике	Автоматическое
Показатель	Тип показателя (метрики), например, доля или количество	Свободное
Описание	Описание типа показателя	Свободное
Дата создания	Дата создания записи	Автоматическое
Дата редактирования	Дата редактирования записи	Автоматическое

Справочник "Система-источник"

Справочник содержит в себе список источников данных и их описания, атрибуты которых планируется проверять в базах данных. Наличие записи по источнику, атрибуты которой планируется проверять в базе данных, в текущем справочнике является обязательным условием для создания проверки качества данных, а также корректного создания новых записей в справочнике "Сущности".

Наименование поля	Описание поля	Тип заполнения
Код источника	Уникальный идентификатор системы-источника данных	Свободное
Наименование источника	Бизнес-наименование системы-источника данных	Свободное
Дата создания	Дата создания записи	Автоматическое
Дата редактирования	Дата редактирования записи	Автоматическое

Справочник "Статус требования"

Справочник содержит в себе статусы реализации требований к качеству данных.

Наименование поля	Описание поля	Тип заполнения
-------------------	---------------	----------------

ID	Уникальный идентификатор записи в справочнике	Автоматическое
Наименование статуса	Наименования статуса реализации требования к качеству данных	Свободное
Описание статуса	Описание статуса	Свободное
Дата создания	Дата создания записи	Автоматическое
Дата редактирования	Дата редактирования записи	Автоматическое

Справочник "Сущности"

Справочник содержит наименования таблиц с их бизнес-описанием, а также связи этих таблиц с областями данных и системами-источниками. Наличие информации и корректное ее заполнение в данном справочнике, является обязательным условием для создания проверки, а также для корректного формирования метаданных этой проверки. Отсутствие или некорректно введенная информация приведет к неуспешному выполнению проверки, либо к некорректным ее результатам.

Наименование поля	Описание поля	Тип заполнения
ID	Уникальный идентификатор записи в справочнике	Автоматическое
Наименование БД	Техническое наименование базы данных, в которую входит таблица	Свободное
Наименование сущности в БД	Техническое наименование таблицы	Свободное
Бизнес-описание сущности	Бизнес-наименование таблицы	Свободное
Наименование источника	Бизнес-наименование системы, которая является источником данных для таблицы	Выбор из списка
Наименование области данных	Бизнес-наименование базы данных (области данных)	Выбор из списка
Дата создания	Дата создания записи	Автоматическое
Дата редактирования	Дата редактирования записи	Автоматическое

Справочник "Атрибуты"

Справочник содержит наименования атрибутов с их бизнес-описанием, а также связи этих атрибутов с таблицами. Наличие информации и корректное ее заполнение в данном справочнике, является обязательным условием для создания проверки, а также для корректного формирования метаданных этой проверки. Отсутствие или некорректно введенная информация приведет к неуспешному выполнению проверки, либо к некорректным ее результатам.

Наименование поля	Описание поля	Тип заполнения
ID	Уникальный идентификатор записи в справочнике	Автоматическое
Наименование атрибута в БД	Техническое наименование атрибута	Свободное
Бизнес-описание атрибута	Бизнес-наименование атрибута	Свободное
Идентификатор сущности	Конкатенация технического и бизнес-наименования таблицы, которая содержит данный атрибут	Выбор из списка

Дата создания	Дата создания записи	Автоматическое
Дата редактирования	Дата редактирования записи	Автоматическое

Справочник "Типы требований"

Справочник содержит типы требований к качеству данных и их описание. Отсутствие значений в справочнике сделает невозможным создание проверки качества данных.

Наименование поля	Описание поля	Тип заполнения
ID	Уникальный идентификатор записи в справочнике	Автоматическое
Тип требования	Тип требования к качеству данных	Свободное
Описание	Описание типа требования	Свободное
Дата создания	Дата создания записи	Автоматическое
Дата редактирования	Дата редактирования записи	Автоматическое

Справочник "Бизнес-задача"

Справочник содержит бизнес-задачи и их описание, в рамках решения которых реализовываются проверки качества данных. Отсутствие значений в справочнике сделает невозможным создание проверки качества данных.

Наименование поля	Описание поля	Тип заполнения
ID	Уникальный идентификатор записи в справочнике	Автоматическое
Бизнес-задача	Наименование бизнес-задачи	Свободное
Описание бизнес-задачи	Краткое описание бизнес-задачи	Свободное
Дата создания	Дата создания записи	Автоматическое

1.2 UI Конструктор проверок

В Сервисе Контроля качества данных есть модуль Конструктор проверок, который позволяет пользователям создавать и редактировать проверки качества данных.

Переход в UI Конструктор проверок осуществляется выбором в списке соответствующего UI.

Модуль "Конструктор проверок" состоит из двух вкладок:

- Список проверок
- Очередь пакетов

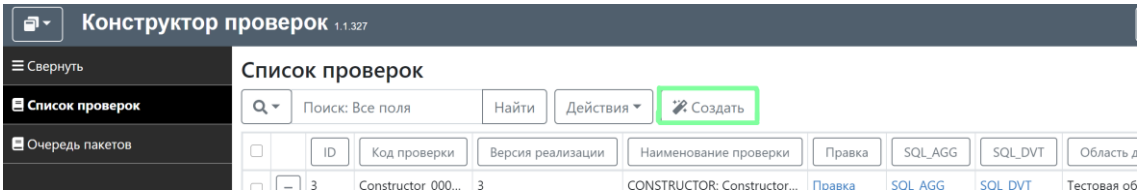
Вкладка "Список проверок" является основной рабочей вкладкой по созданию проверки качества данных. На странице представлена информация по разработанным проверкам качества данных. С этой страницы осуществляется переход непосредственно в мастер создания новых проверок качества данных, также есть возможность создания проверки корректируя ранее разработанную.

Создать проверку качества данных с помощью конструктора очень просто и не требует каких-то дополнительных знаний. Но если вы хотите создать проверку, которая показывает валидный результат, Вам необходимо знать и понимать структуру базы данных, таблицы и атрибуты которой Вы хотите проверять, а также понимать бизнес-процессы, чтобы учесть все их нюансы при создании проверки.

Вкладка "Очередь пакетов" является информационной и показывает состояние созданной проверки в очереди и информация по связи с пакетами проверок.

Рассмотрим создание новой проверки более детально.

Создание новой проверки качества данных



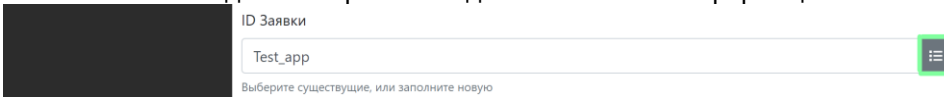
Для входа в мастер создания проверки необходимо нажать на кнопку "Создать".
Мастер создания проверки состоит из 7 шагов, которые необходимо последовательно пройти для получения итогового результата - проверки качества данных.


Шаг 1. Выбор заявки

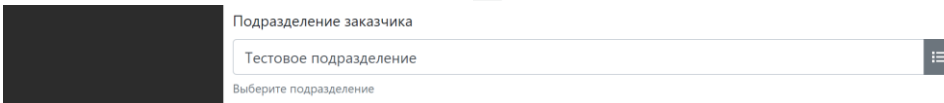


На первом шаге необходимо привязать проверку к одной из существующих заявок или создать новую заявку.

Для создания новой заявки необходимо вписать ее наименование в поле "ID заявки". Обратите внимание, что если Вы укажете наименование заявки совпадающее с ранее созданной, то автоматически подтянется ранее созданная заявка и информация по ней.




После того, как Вы указали ID новой заявки, необходимо заполнить поле "Подразделение заказчика". Значение выбирается из справочника подразделений. Справочник значений вызывается нажатием кнопки .



Далее необходимо заполнить поле "ФИО Заказчика" - это поле со свободным вводом текста.

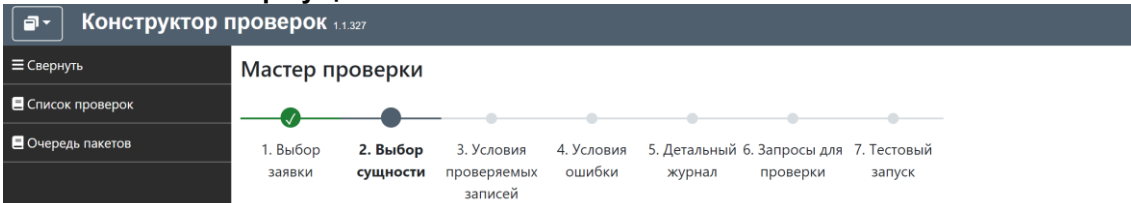


Если же, Вы хотите привязать проверку к ранее созданной заявке, то в поле для ввода ID заявки нажмите на кнопку  и выберите необходимую заявку в раскрывшемся списке.

Значения в поля "Подразделение Заказчика" и "ФИО Заказчика" подтянутся автоматически и будут недоступны для редактирования.

После того, как все поля на первом шаге заполнены, нажмите на кнопку "Следующий шаг".

Шаг 2. Выбор сущности



На втором шаге необходимо выбрать сущность, атрибуты которой планируется проверять и указать путь до этой таблицы.


АС реализации проверки - в поле указывается хранилище, в котором находятся области данных (витрины, реплики), атрибуты которых планируется проверять в рамках текущей проверки.

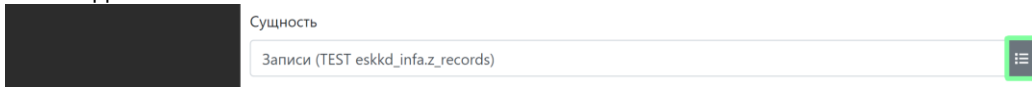


Область данных - в поле указывается схема, в которой находятся таблицы, атрибуты которых планируется проверять в рамках текущей проверки.

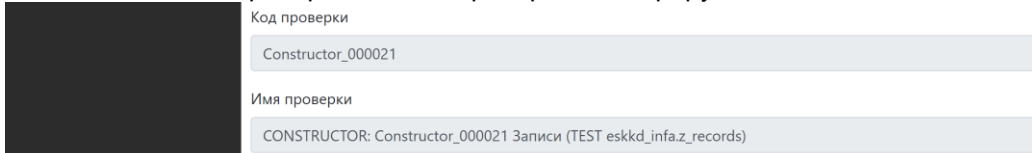


Сущность - таблица, атрибуты которой планируется проверять в рамках текущей проверки.

Заполнение всех значений на втором шаге осуществляется выбором из справочников. Для того, чтобы открыть значения справочников, нужно нажать на кнопку  и выбрать из списка необходимое значение.

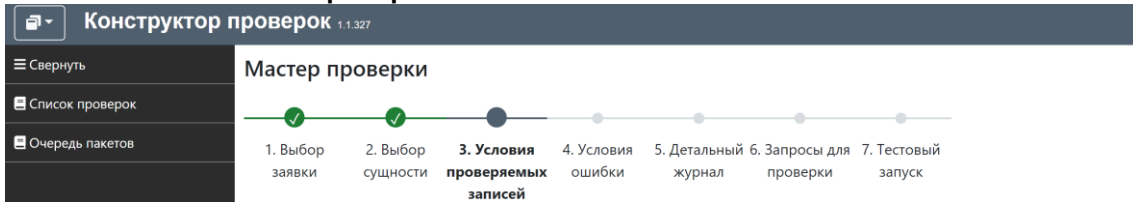


Поля "Код проверки" и "Имя проверки" сгенерируются автоматически.



После того, как все поля на втором шаге заполнены, нажмите на кнопку "Следующий шаг".


Шаг 3. Условия проверяемых записей

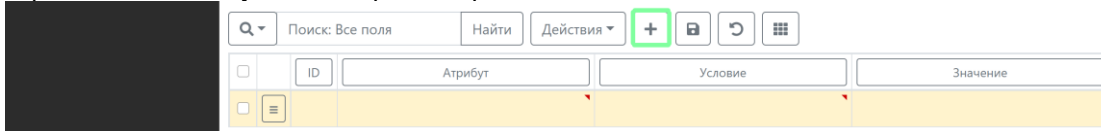


На третьем шаге создаются условия отбора проверяемых записей, т.е. те условия, которые ограничивают выборку проверяемых данных.

Внимание!!! На данном шаге выбираются именно ограничение набора проверяемых записей, а не создаются условия на ошибку.

Внимание!!! Создание условий отбора проверяемых записей не является обязательным для заполнения. Если Вы хотите проверять всю таблицу, тогда просто пропустите это шаг.

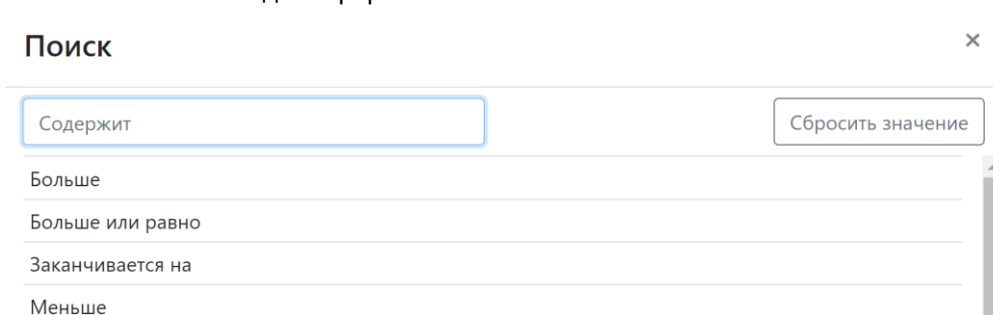
Для добавления условия отбора необходимо нажать на кнопку . После нажатия кнопки 1 строка станет доступной для редактирования.



Значение в поле "Атрибут" - выбор значения из справочника. Отображаются только атрибуты, привязанные к выбранной на предыдущем шаге таблице.



Значения в поле "Условие" выбирается из списка, а значение в поле "Значение" заполняется в свободной форме.




Описание вариантов условий и примеры значение представлены в таблице ниже.

Условие	Описание условия (синтаксис SQL)	Условие заполнения "Значения"
Содержит	schema.table.attr LIKE "%VALUE%"	Текст, значения, символы

Заканчивается на	schema.table.attr LIKE "%VALUE"	Текст, значения, символы
Равно	schema.table.attr = "VALUE"	Текст, значения, символы
Больше	schema.table.attr > "VALUE"	Текст, значения, символы
Больше или равно	schema.table.attr >= "VALUE"	Текст, значения, символы
Пустое	schema.table.attr IS NULL OR schema.table.attr = ""	НЕ ЗАПОЛНЯЕТСЯ
Меньше	schema.table.attr < "VALUE"	Текст, значения, символы
Меньше или равно	schema.table.attr <= "VALUE"	Текст, значения, символы
Ссылочная целостность	join schema.table1 t on t.attr1=schema.table.attr	schema.table1.attr1
Не содержит	schema.table.attr NOT LIKE "%VALUE%"	Текст, значения, символы
Не заканчивается на	schema.table.attr NOT LIKE "%VALUE"	Текст, значения, символы
Не равно	schema.table.attr <> "VALUE"	Текст, значения, символы
Не пустое	schema.table.attr IS NOT NULL OR schema.table.attr <> ""	НЕ ЗАПОЛНЯЕТСЯ
Не начинается с	schema.table.attr NOT LIKE "VALUE%"	Текст, значения, символы
Начинается с	schema.table.attr LIKE "VALUE%"	Текст, значения, символы
Содержит	schema.table.attr LIKE "%VALUE%"	Текст, значения, символы
Заканчивается на	schema.table.attr LIKE "%VALUE"	Текст, значения, символы

Условия отбора проверяемых записей соединяются по логическому "И" (AND).

После того, как условия отбора созданы нажмите кнопку . Если Вы хотите сбросить не сохраненные изменения нажмите на кнопку .

Для того, чтобы произвести действия (дублировать, удалить) с уже созданной строкой или строкой, которая на текущий момент редактируется нажмите на кнопку  и в открывшемся контекстном меню выберите действие, которое хотите совершить со строкой.



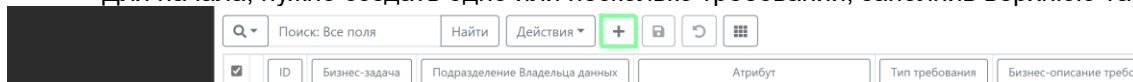
После того, как Вы закончили создавать условия отбора и текущий список Вас устраивает, тогда нажмите на кнопку "Следующий шаг".

Шаг 4. Условия ошибки





На четвертом шаге создаются непосредственно сами требования к качеству данных, их описания, условия на ошибку.


Для начала, нужно создать одно или несколько требований, заполнив верхнюю таблицу.



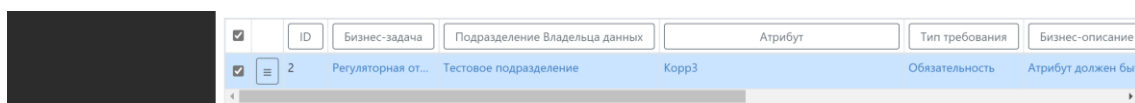
Список полей, их описание и принцип заполнения описаны в таблице ниже.

Наименование поля	Описание поля	Принцип заполнения
Бизнес-задача	Наименование бизнес-задачи, для решения которой создается текущая проверка качества данных	Выбор из справочника
Подразделение Владельца данных	Наименование подразделения, ответственного за исправление ошибок качества данных в случае их выявления	Выбор из справочника
Атрибут	Атрибут таблицы, которые планируется проверять в рамках текущего требования к качеству данных	Выбор из справочника
Тип требования	Тип проверки	Выбор из справочника
Бизнес-описание требования к качеству данных	Подробное описание того, что проверяется в данной проверке, т.е. максимально подробно сформулированное требование к качеству данных	Заполняется в свободной форме
Краткое описание требования (для dashboard)	Описание требования, которое отображается в отчете "Результаты проверок"	Формируется автоматически, но можно редактировать в свободной форме
Показатель	Тип показателя (по умолчанию. доля (доля корректных записей))	Выбор из справочника
Пороговое значение	Пороговое значение показателя, при превышении которого данные считаются неудовлетворительными.	По умолчанию проставляется 90.0, но можем быть отредактировано на любое числовое, положительное значение

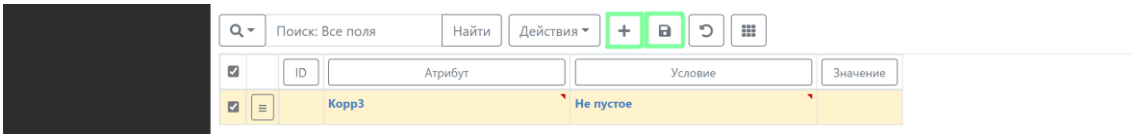
После того, как требования и их описания созданы нажмите кнопку . Если Вы хотите сбросить не сохраненные изменения нажмите на кнопку .

Для того, чтобы произвести действия (дублировать, удалить) с уже созданной строкой или строкой, которая на текущий момент редактируется нажмите на кнопку  и в открывшемся контекстном меню выберите действие, которое хотите совершить со строкой.

Теперь необходимо заполнить сами условия на ошибку. Для этого, к каждому созданному Вами требованию, необходимо привязать хотя бы одно (и методологически правильно, чтобы было одно условие на ошибку по одному требованию) условие на ошибку.



Выделяем требование и заполняем нижнюю таблицу. Заполнение нижней таблицы происходит аналогично с третьим шагом, только здесь определяются именно условия на ошибку.



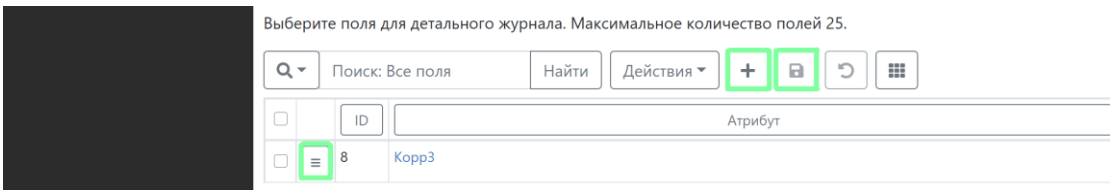
После того, как все условия на ошибку заполнены, нажмите кнопку "Следующий шаг".

Шаг 5. Детальный журнал



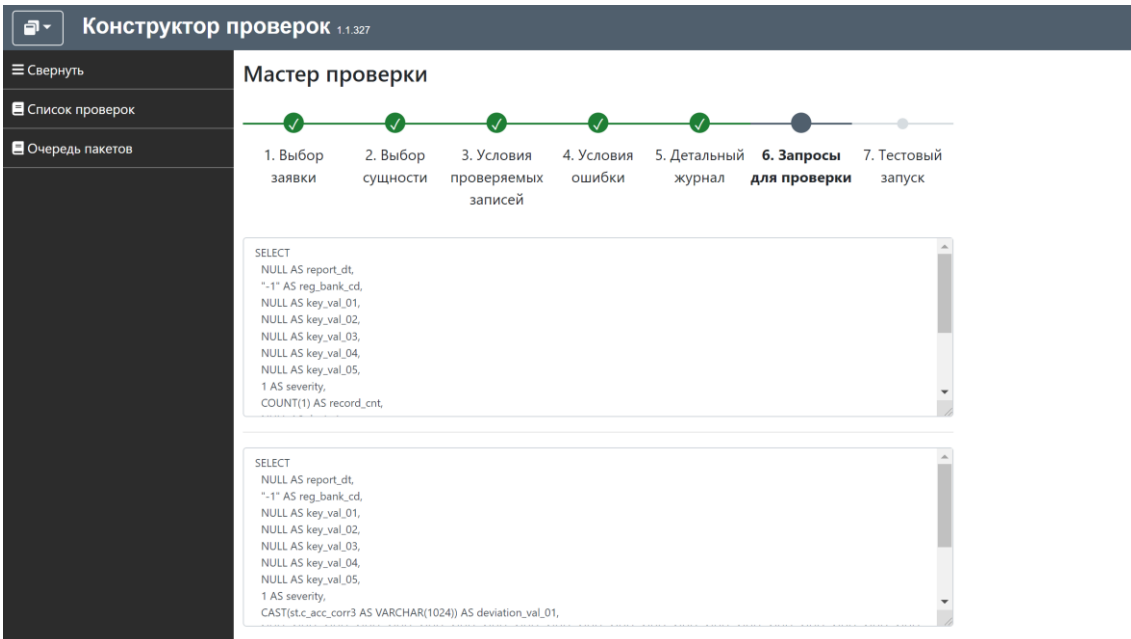
На пятом шаге определяется набор атрибутов, которые будут выводиться в детальный журнал ошибочных записей. В детальный журнал необходимо выводить те атрибуты таблицы, которые позволяют однозначно идентифицировать ошибочную запись, например, ID. По умолчанию, на пятом шаге мастера, выводятся все атрибуты проверяемой таблицы, которые есть в справочнике атрибутов. Для того, чтобы редактировать список (дублировать, удалить) нажмите на кнопку и в открывшемся контекстном меню выберите действие, которое хотите совершить со строкой. Для добавления атрибута (если он был ошибочно удален) необходимо нажать на кнопку .

После того, как список готов нажмите кнопку . Если Вы хотите сбросить не сохраненные изменения нажмите на кнопку .



Далее нажмите кнопку "Следующий шаг".

Шаг 6. Запросы для проверки



Шестой шаг является информационным. На нем представлено два окна с sql-кодами получившихся проверок.

Верхнее окно - это sql-код получения агрегированных результатов проверки, т.е. результатом его выполнения будут являться количество проверяемых записей и количество ошибочных записей по каждому созданному на 4 шаге условию на ошибку.

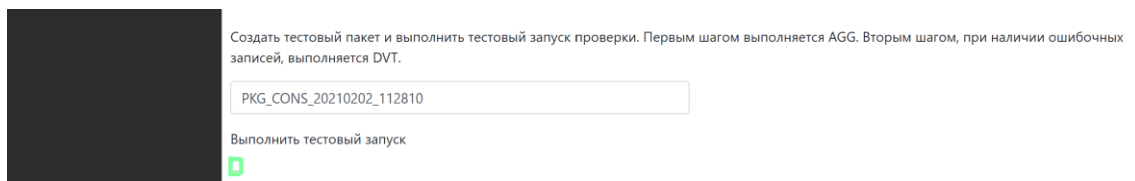
Нижнее окно - это sql-код получения детального списка ошибочных записей.

После того, как Вы ознакомились с получившимися кодами проверок, нажмите кнопку "Следующий шаг".

Шаг 7. Тестовый запуск



На седьмом шаге предлагается произвести тестовый запуск созданной проверки, либо закончить создание проверки без тестового запуска.

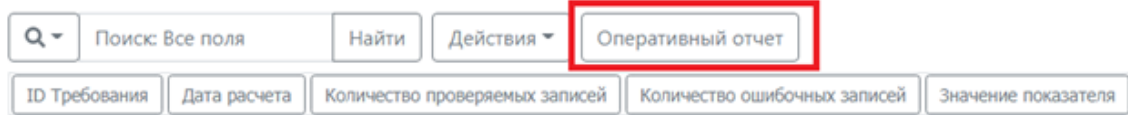


1.3 UI Результаты проверок

За осуществление мониторинга качества данных отвечает модуль Визуализация и отчетность. Модуль реализуется пользовательским интерфейсом – Результаты проверок.

Модуль имеет два уровня детализации результатов: оперативный отчет и аналитический отчет.

Аналитический отчет



Результаты проверок качества данных в отчетах представлены в виде таблиц, содержащих в себе сами результаты проверок и метаданные, позволяющие идентифицировать принадлежность результатов проверок к тому или иному требованию к качеству данных.

Для удобства использования отчеты представлены в двух разрезах: dashboard для Потребителей данных и dashboard для Владельцев данных. Dashboard Переход между двумя разрезами осуществляется при нажатии на соответствующую кнопку.

Dashboard для Владельцев данных



Dashboard'ы отличаются друг от друга набором метаданных, позволяющих пользователям осуществлять поиск интересующих результатов проверок качества данных по специфичным для роли полям, например, подразделение Потребителя/Владельца данных.

Рассмотрим каждый отчет в частности.

Оперативный отчет показывает актуальные результаты проверок качества данных (результаты последних замеров качества данных по каждой проверке).

Dashboard'ы содержат следующие поля:

Наименование поля	Краткое описание	Наличие в Dashboard для Потребителей данных	Наличие в Dashboard для Владельцев данных
Подразделение заказчика	Наименование подразделения Заказчика требования к качеству данных	✓	-
Блок Владельца данных	Наименование подразделения, ответственного за данные (Владельца данных)	-	✓
Бизнес-задача	Бизнес-задача решаемая Заказчиком, для целей которых создаются проверки качества данных и осуществляется мониторинг результатов	✓	-
Система источник	Система, данные которой проверяются напрямую, либо в реплике	-	✓
Область данных	Наименование области данных, в которой хранятся проверяемые данные	✓	-
Сущность	Бизнес-наименование проверяемой сущности (таблицы)	✓	✓
Атрибут	Бизнес-наименование проверяемого атрибута (поля в таблице)	✓	✓
ID требования	Уникальный идентификатор требования к качеству данных	✓	✓
Описание требования	Бизнес-описание алгоритма проверки качества данных	✓	✓
Дата расчета	Дата выполнения проверки, результаты которой отображаются в этой строке	✓	✓
Количество проверяемых записей	Количество записей, которое было проверено	✓	✓
Количество ошибочных записей	Количество записей, в которых по результатам проверки была обнаружена	✓	✓

	ошибка качества данных		
Значение показателя	В зависимости от типа показателя может показывать: дол корректных записей из числа проверенных, долю ошибочных и пр.	✓	✓
Пороговое значение	Значение показателя, при достижении которого данные считаются непригодными к использованию	✓	✓
Показатель	Значение показателя, при достижении которого данные считаются непригодными к использованию	✓	✓
Тип требования	Тип проверки качества данных (обязательность, уникальность, ссылочная целостность и т.д.)	✓	✓

Аналитический отчет отображает результаты проверок качества данных за все время. В отличие от оперативного отчета, в аналитическом нет разделения между Потребителями и Владельцами данных. Отчёт содержит следующие поля:

Наименование поля	Краткое описание
ID требования	Уникальный идентификатор требования к качеству данных
Дата	Дата выполнения проверки качества данных
Количество проверенных записей	Количество записей, которое было проверено
Количество ошибочных записей	Количество записей, в которых по результатам проверки была обнаружена ошибка качества данных
Значение показателя	В зависимости от типа показателя может показывать: дол корректных записей из числа проверенных, долю ошибочных и пр.
Показатель	Значение показателя, при достижении которого данные считаются непригодными к использованию

Фильтрацию в отчетах можно осуществлять двумя способами: нажатием на наименование самого поля с последующем указанием интересующих значением этого поля, либо выбором из списка значений на панели фильтрации в левой части экрана. Фильтры, которые применяются на таблицу в текущий момент, указаны в левом верхнем углу экрана.

Подразделение заказчика

Блок "Корпоративные инвестиционные услуги"
Блок "Риски"
Блок "Розничный бизнес"
Блок "Технологии"
Блок "Финансы"

Бизнес-задача

Массовая персонализация
Не определено
Операционная деятельность
ППР Руководителей

Область данных

Витрина КФЛ
Данные ЕСС в детальном интегрированном слое
Данные ЕСС в детальном слое
ПКАП - Слой Детальных данных

Сущность

Аналитический счет
Бюджеты в СБОЛ