

*LOW-CODE ПЛАТФОРМА  
ДЛЯ СОЗДАНИЯ ETL-  
ПАЙПЛАЙНОВ LEGATO  
DATA*

Руководство пользователя

Москва 2023

## Оглавление

<b>Аннотация .....</b>	<b>3</b>
<b>Список используемых определений.....</b>	<b>4</b>
<b>Список используемых терминов и сокращений .....</b>	<b>6</b>
<b>1. Введение .....</b>	<b>7</b>
<b>1.1 Область применения .....</b>	<b>7</b>
<b>1.2 Краткое описание возможностей.....</b>	<b>7</b>
<b>1.3 Уровень подготовки пользователя.....</b>	<b>7</b>
<b>1.4 Перечень эксплуатационной документации.....</b>	<b>8</b>
<b>2. Назначение и условия применения .....</b>	<b>9</b>
<b>2.1 Виды деятельности, функции.....</b>	<b>9</b>
<b>2.2 Программные и аппаратные требования к Системе .....</b>	<b>10</b>
<b>3. Подготовка к работе .....</b>	<b>12</b>
<b>3.1 Порядок запуска Платформы .....</b>	<b>12</b>
<b>3.2 Вход в Платформу .....</b>	<b>12</b>
<b>4. Описание операций .....</b>	<b>14</b>
<b>4.1 Функции управления профилем пользователя.....</b>	<b>14</b>
4.1.1 Функция просмотра и редактирования профиля пользователя.....	14
4.1.2 Функция выхода из учетной записи .....	16
<b>4.2 Функции управления отображения пользовательского интерфейса.....</b>	<b>16</b>
4.2.1 Функция смены темы оформления пользовательского интерфейса .....	16
4.2.2 Функция смены языка оформления пользовательского интерфейса.....	17
<b>4.3 Функции настройки области отображения пайплайна и вспомогательных     сущностей</b>	<b>18</b>
4.3.1 Функция просмотра и редактирования перечня Компаний .....	18
4.3.2 Функция просмотра и редактирования общей информации для Компании .....	19
4.3.3 Функция просмотра и редактирования Подключений в рамках Компании .....	21
4.3.4 Функция просмотра и редактирования перечня Проектов в рамках Компании .....	23

4.3.5	Функция просмотра и редактирование информации о Проекте .....	24
4.3.6	Функция настройки расписания Пайплайнов по-умолчанию в рамках Проекта .....	26
4.3.7	Функция настройки оркестрации Пайплайнов по-умолчанию в рамках Проекта .....	31
4.3.8	Функция просмотра и редактирования перечня Пайплайнов в рамках Проекта .....	33
4.3.9	Функция создания Пайплайна .....	34
4.3.10	Функция просмотра и редактирования информации о Пайплайне .....	35
<b>4.4</b>	<b>Функции запуска и настроек запуска Пайплайна .....</b>	<b>36</b>
4.4.1	Функция настройки оркестрации Пайплайна .....	36
4.4.2	Функция настройки расписания Пайплайна .....	36
4.4.3	Функция постановки Пайплайна на исполнение .....	36
<b>4.5</b>	<b>Функции работы с задачами загрузки/выгрузки данных .....</b>	<b>37</b>
4.5.1	Функция создания набора данных для задач типа «Загрузка» .....	37
4.5.2	Функция редактирования набора данных для задач типа «Загрузка» .....	40
4.5.3	Функция удаления набора данных для задач типа «Загрузка» .....	41
4.5.4	Функция создания задачи типа «Загрузка» .....	41
4.5.5	Функция редактирования задачи типа «Загрузка» .....	43
4.5.6	Функция удаления задачи типа «Загрузка» .....	45
4.5.7	Функция редактирования набора данных типа «Выгрузка» .....	45
4.5.8	Функция создания задачи типа «Выгрузка» .....	45
4.5.9	Функция редактирования задачи типа «Выгрузка» .....	47
4.5.10	Функция удаления задачи типа «Выгрузка» .....	49
<b>4.6</b>	<b>Функции работы с задачами трансформации данных .....</b>	<b>49</b>
4.6.1	Функция создания задачи типа «Трансформация» .....	49
4.6.2	Функция редактирования задачи типа «Трансформация» .....	51
4.6.3	Функция удаления задачи типа «Трансформация» .....	52
<b>4.7</b>	<b>Функции мониторинга .....</b>	<b>53</b>
4.7.1	Мониторинг выполнения задач .....	53

## **Аннотация**

Документ содержит руководство пользователя автоматизированной платформой управления процессами работы с данными «Legato Data» (далее – Платформа).

## Список используемых определений

Перечень используемых определений представлен в Таблица 1:

№ п/п	Термин	Определение
1	GUI	(англ. Graphical User Interface) – разновидность пользовательского интерфейса, в котором элементы интерфейса (меню, кнопки, значки, списки и т.д.), представленные пользователю на дисплее, исполнены в виде графических изображений
2	JSON	(англ. JavaScript Object Notation) - стандартный текстовый формат для представления структурированных данных на основе синтаксиса объекта JavaScript
3	Вычислительный кластер	массив серверов (вычислительных узлов или так называемых нодов) объединенных коммуникационной сетью и размещенных в отдельной стойке. Вычислительный узел имеет несколько многоядерных процессоров, свою оперативную память и работает под управлением своей операционной системы.
4	ETL-пайплайн или ETL-процесс или Пайплайн	процесс объединения данных из нескольких источников в одном центральном хранилище, которое называется складом данных.
5	Мониторинг	Процесс Систематического и непрерывного сбора информации о параметрах сложного объекта или процесса
6	Low-code подход	метод проектирования и разработки приложений с помощью интуитивно понятных графических инструментов и встроенных функций, которые снижают традиционные (профессиональные) требования к написанию кода
7	API	(англ. Application Programming Interface) – программный интерфейс приложения. В

№ п/п	Термин	Определение
		контексте API слово «приложение» относится любому ПО с определенной функцией. Интерфейс можно рассматривать как сервисный контракт между двумя приложениями.
8	Компания	Компания (организация-заказчика), содержит в себе Проекты, используется для разграничения областей видимости Пайплайнов
9	Проект	Проект, содержит в себе Пайплайны, относится к Компании
10	Подключение	Подключение к внешней или внутренней системе – источнику данных/назначению данных. Термин используется для обозначения подключений к публичным и непубличным API, СУБД, брокерам сообщений.
11	Набор данных (датасет)	Сущность, содержащая информацию о Подключении со схемой хранения данных в системе, к которой осуществлено Подключение
12	Apache Spark	Открытое программное обеспечение для реализации распределённой обработки неструктурированных и слабоструктурированных данных, входящий в экосистему проектов Hadoop
13	Apache Airflow	Открытое программное обеспечение для создания, выполнения, мониторинга и оркестровки потоков операций по обработке данных
14	Нода	Ярлык задачи любого типа на рабочей области конструктора Пайплайна

Таблица 1

## Список используемых терминов и сокращений

Список используемых терминов и сокращений приведен в Таблица 2:

№ п/п	Термин	Определение
1	ETL	(англ. Extract, Transform, Load) — дословно «извлечение, преобразование, загрузка»)
2	АРМ	Автоматизированное рабочее место
3	БД	База данных
4	СУБД	Система управления базами данных

*Таблица 2*

# 1. Введение

## 1.1 Область применения

Автоматизированная платформа управления процессами работы с данными «Legato Data» предназначена для автоматизации настройки и управления процессами работы с данными с использованием low-code подхода, в том числе и для сбора, обработки и загрузки данных.

## 1.2 Краткое описание возможностей

Платформа предоставляет пользователю следующие возможности:

- подключение к различным СУБД, публичным и непубличным API, BI-системам;
- создание задач загрузки и выгрузки данных;
- создание задач трансформации данных;
- визуальный конструктор ETL-пайплайнов;
- управление данными;
- тестирование качества данных;
- настройка расписания выполнения ETL-пайплайнов;
- мониторинг работы ETL-пайплайнов.

Для решения данных задач Платформа предоставляет следующий функционал:

- Функции управления профилем пользователя
- Функции управления отображения пользовательского интерфейса
- Функции настройки области отображения пайплайна и вспомогательных сущностей
- Функции запуска и настроек запуска Пайплайна
- Функции работы с задачами загрузки/выгрузки данных
- Функции работы с задачами трансформации данных
- Функции мониторинга

## 1.3 Уровень подготовки пользователя

Пользователи данной Платформы должны пройти обучение работе с Платформой и обладать следующими навыками и знаниями:

- базовые навыки работы на персональном компьютере с графическим пользовательским интерфейсом (клавиатура, манипулятор типа «мышь», управление окнами и приложениями, файловая система);
- знание основ информационной безопасности; знание предметной области;

- знание всего комплекса задач, решаемых Платформой;
- навыки работы в операционной системе Windows;
- навыки работы с браузерами (Internet Explorer, Mozilla).

#### **1.4 Перечень эксплуатационной документации**

Пользователю необходимо ознакомиться со следующей рабочей документацией на Платформу:

- Текущее «Руководство пользователя Legato Data».

## 2. Назначение и условия применения

### 2.1 Виды деятельности, функции

Платформа предназначена для автоматизации следующих функций:

- Функция просмотра и редактирования перечня Компаний;
- Функция просмотра и редактирования общей информации для Компании;
- Функция просмотра и редактирования Подключений в рамках Компании;
- Функция просмотра и редактирования перечня Проектов в рамках Компании;
- Функция просмотра и редактирование информации о Проекте;
- Функция настройки расписания Пайплайнов по-умолчанию в рамках Проекта;
- Функция настройки оркестрации Пайплайнов по-умолчанию в рамках Проекта;
- Функция просмотра и редактирования перечня Пайплайнов в рамках Проекта;
- Функция создания Пайплайна;
- Функция просмотра и редактирования информации о Пайплайне;
- Функция настройки оркестрации Пайплайна;
- Функция настройки расписания Пайплайна;
- Функция постановки Пайплайна на исполнение;
- Функция создания набора данных для задач типа «Загрузка»;
- Функция редактирования набора данных для задач типа «Загрузка»;
- Функция удаления набора данных для задач типа «Загрузка»;
- Функция создания задачи типа «Загрузка»;
- Функция редактирования задачи типа «Загрузка»;
- Функция удаления задачи типа «Загрузка»;
- Функция создания задачи типа «Трансформация»;
- Функция редактирования задачи типа «Трансформация»;
- Функция удаления задачи типа «Трансформация»;
- Функция редактирования набора данных типа «Выгрузка»;
- Функция создания задачи типа «Выгрузка»;
- Функция редактирования задачи типа «Выгрузка»;
- Функция удаления задачи типа «Выгрузка»;
- Функция смены темы оформления пользовательского интерфейса;
- Функция смены языка оформления пользовательского интерфейса;
- Функция просмотра и редактирования профиля пользователя;

- Функция выхода из учетной записи;
- Мониторинг выполнения задач.

## 2.2 Программные и аппаратные требования к Системе

Для обеспечения функционирования программного и информационного обеспечения Платформы используется комплекс технических средств, включающий серверы и рабочие станции пользователей, а также сетевое оборудование, предназначенное для организации вычислительных сетей и информационного взаимодействия.

В Таблица 3 и Таблица 4 перечислены требования к программному и аппаратному обеспечению APM Legato Data.

№п/п	Компонент	Версия
1	Операционная Система	<ul style="list-style-type: none"> <li>• Windows 2003 Server (Standard Edition, Enterprise Edition) + Service Packs 2</li> <li>• Windows XP (все виды)</li> <li>• Windows 7 (все виды)</li> <li>• Windows Server 2008 (все виды)</li> <li>• Windows 8 или 8.1 (все виды)</li> <li>• Windows 10 (все виды)</li> <li>• macOS Big Sur (все виды)</li> <li>• macOS Catalina (все виды)</li> <li>• macOS Mojave (все виды)</li> <li>• macOS Big Sur (все виды)</li> <li>• Linux Ubuntu 20.10 и позднее</li> </ul>
2	Интернет-браузер	Google Chrome 119.0.6045.159/160 и позднее

Таблица 3

№п/п	Компонент	Минимальные требования	Рекомендуемые требования
1	CPU (процессор)	1ГГц	2ГГц
2	RAM (оперативная память)	2Гб	4Гб
3	HDD	500Мбайт	-
4	Видеокарта и монитор	с разрешением не менее 1024x768 точек и	с разрешением не менее 1024x768 точек

<b>№п/п</b>	<b>Компонент</b>	<b>Минимальные требования</b>	<b>Рекомендуемые требования</b>
		глубиной цвета не менее 32bit	и глубиной цвета не менее 32bit
5	Звуковая карта	Наличие	Наличие
6	Клавиатура	Наличие	Наличие
7	Манипулятор «Мышь» или «Трекпад» или «Тачпад»	Наличие	Наличие
8	UPS (источник бесперебойного питания)	Наличие	Наличие

*Таблица 4*

### 3. Подготовка к работе

#### 3.1 Порядок запуска Платформы

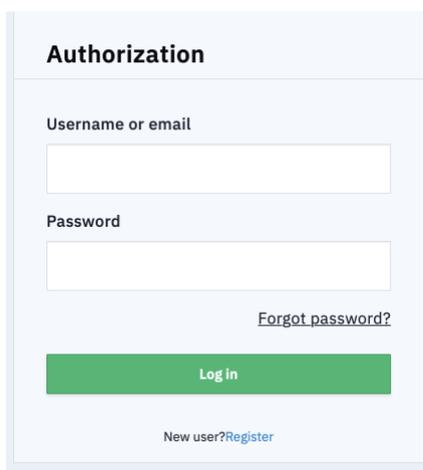
Серверная часть работает в постоянном режиме и не требует от конечного пользователя отдельных операций по запуску.

#### 3.2 Вход в Платформу

Вход в Платформу осуществляется с рабочего места (АРМ), оснащенного доступом к сети Интернет, через интернет-браузер.

Для входа в Систему необходимо выполнить следующие действия:

1. Запустить интернет-браузер. В случае успешного запуска на экране отобразится окно браузера;
2. В адресной строке интернет-браузера ввести адрес сервера приложений, предоставленного администратором. В результате операции Платформа выдаст окно с запросом на ввод следующей информации для доступа на сервер приложений (Рисунок 1):
  - a. Логин пользователя;
  - b. Пароль пользователя;
  - c. Кнопку «Log in».
3. Ввести логин и пароль пользователя, нажать кнопку «Log in». В случае правильно введенных данных Платформа предоставит пользователю окно Legato Data с перечнем доступных пользователю компаний (Рисунок 2).



The image shows a web form titled "Authorization". It contains two text input fields: "Username or email" and "Password". Below the "Password" field is a link labeled "Forgot password?". At the bottom of the form is a prominent green button labeled "Log in". Below the button is a link labeled "New user? Register".

Рисунок 1

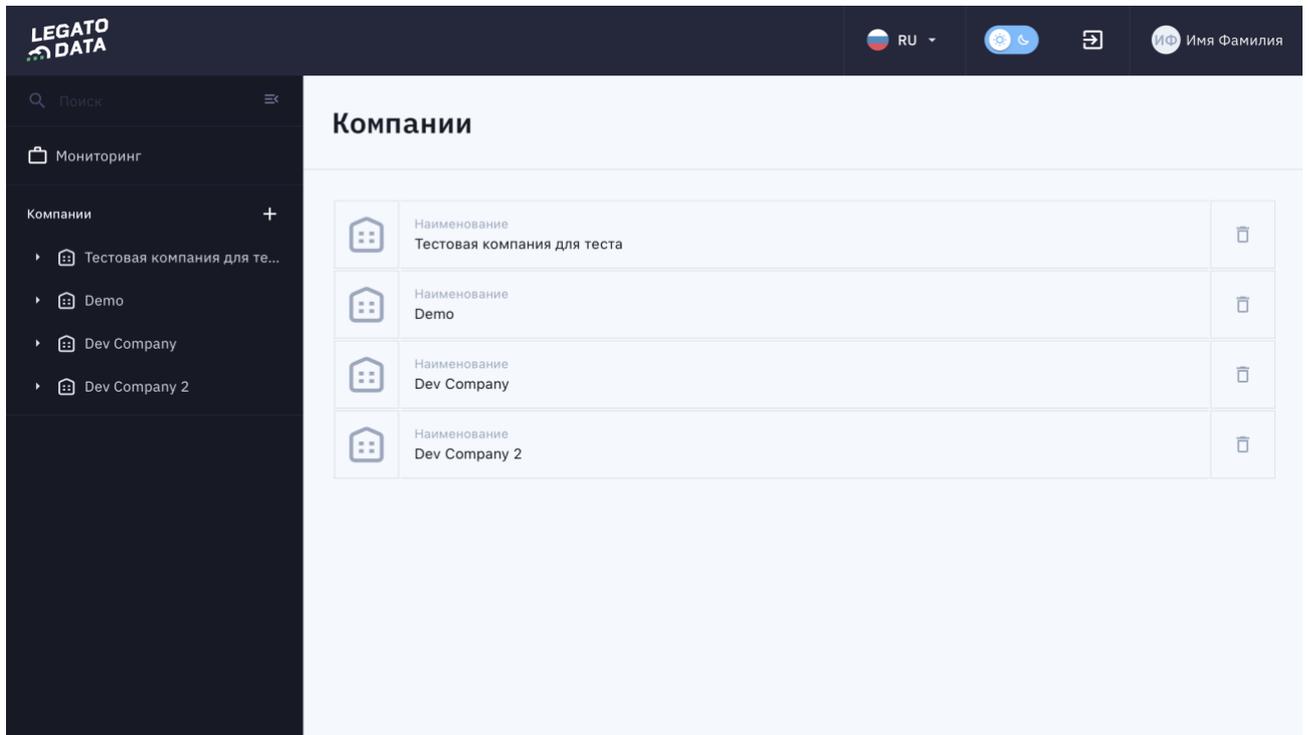


Рисунок 2

## 4. Описание операций

### 4.1 Функции управления профилем пользователя

#### 4.1.1 Функция просмотра и редактирования профиля пользователя

Для просмотра и редактирования профиля пользователя, необходимо:

1. Авторизоваться в Платформе;
2. Нажать кнопку «Профиль» (Рисунок 3);
3. Платформа предоставит пользователю окно просмотра профиля пользователя, которое включает в себя вкладки «Информация» и «Безопасность»:
  - a. Вкладка «Информация» (Рисунок 4) содержит общую информацию о пользователе:
    - i. Имя и фамилия;
    - ii. Аватар пользователя;
    - iii. Кнопка «Изменить фотографию» (используется для смены Аватара пользователя, при нажатии открывается системное модальное окно для загрузки файла);
    - iv. Email – поле для изменения контактного eMail пользователя, при изменении значения изменения автоматически сохраняются в Платформе;
    - v. Поля для заполнения первого, второго и третьего имени пользователя (с целью интернационализации). При изменении значений изменения автоматически сохраняются в Платформе;
  - b. Вкладка «Безопасность» (Рисунок 5) содержит информацию о безопасности пользователя:
    - i. Срок истечение действия пароля;
    - ii. Кнопка «Изменить пароль» (при нажатии на кнопку открывается модальное окно (Рисунок 6), содержащее поля для ввода старого пароля, нового пароля и повтора нового пароля. При успешном вводе старого пароля и одинаковых значениях нового пароля после нажатия кнопки «Сохранить» изменения сохраняются в Платформе).

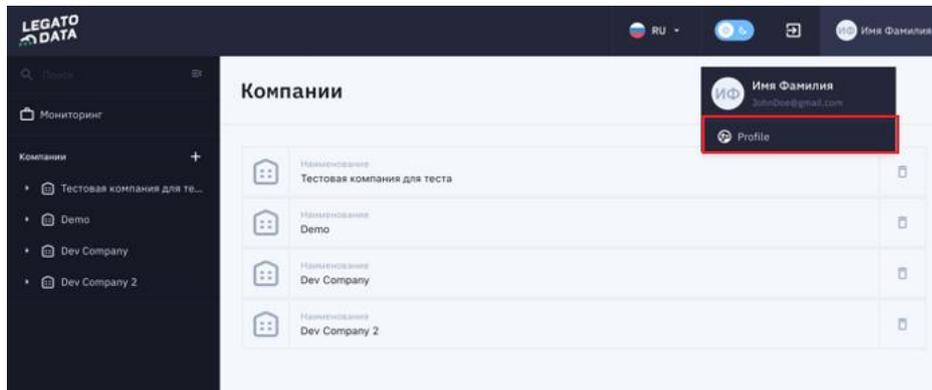


Рисунок 3

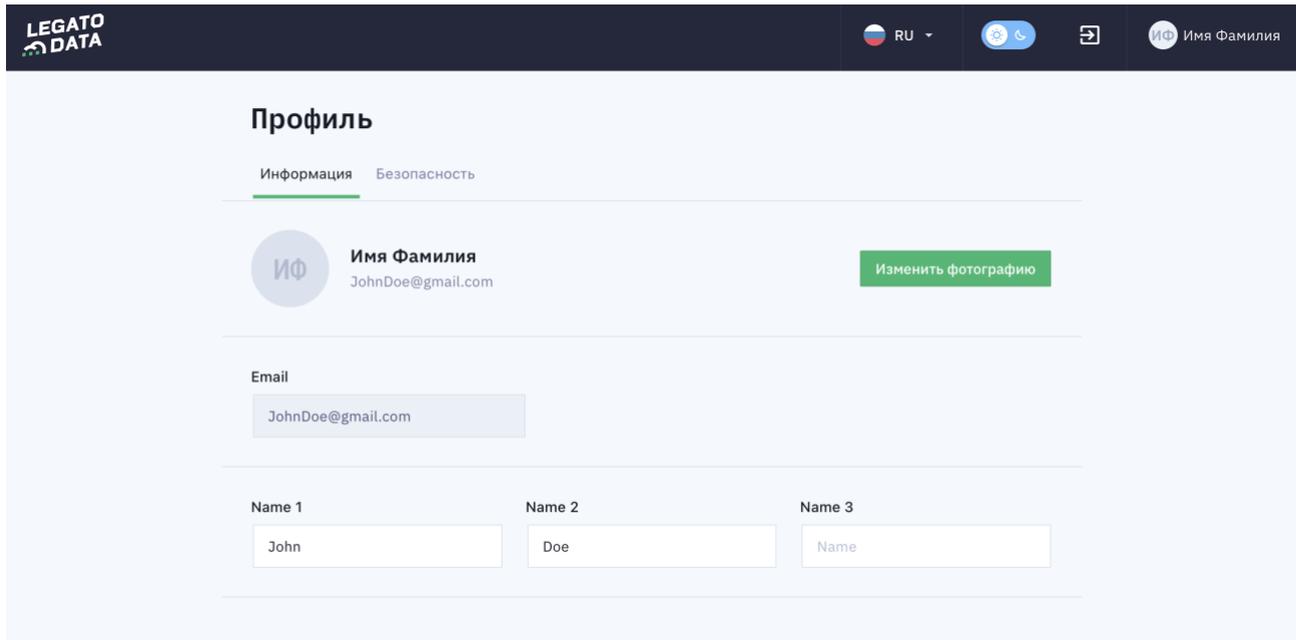


Рисунок 4

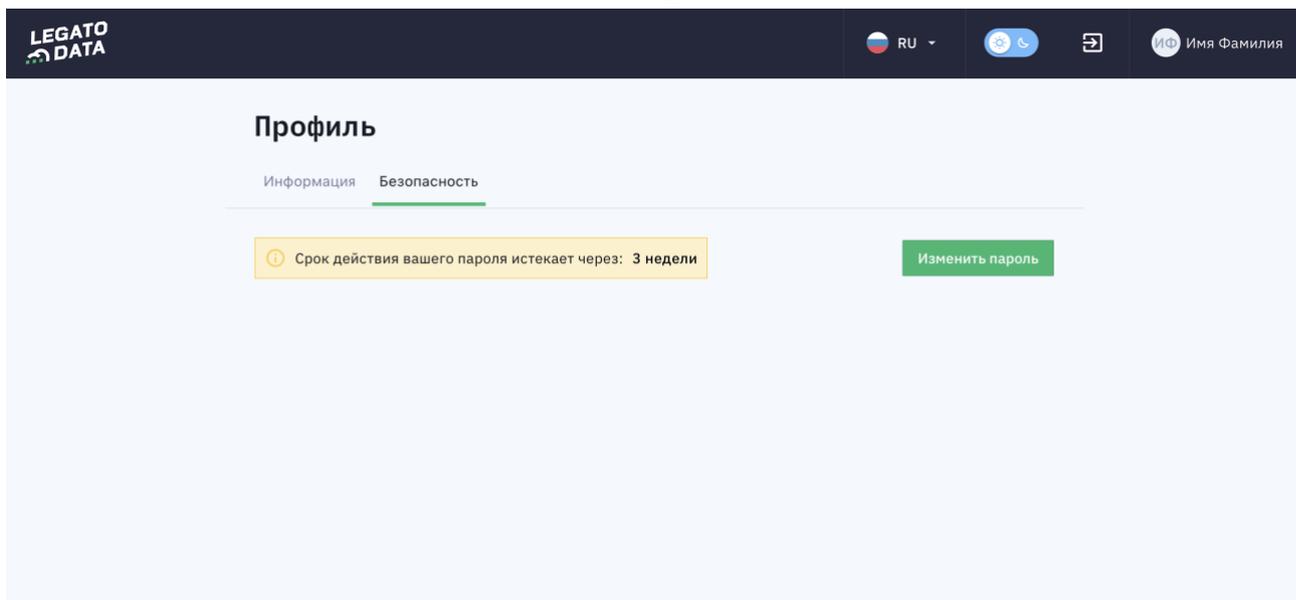
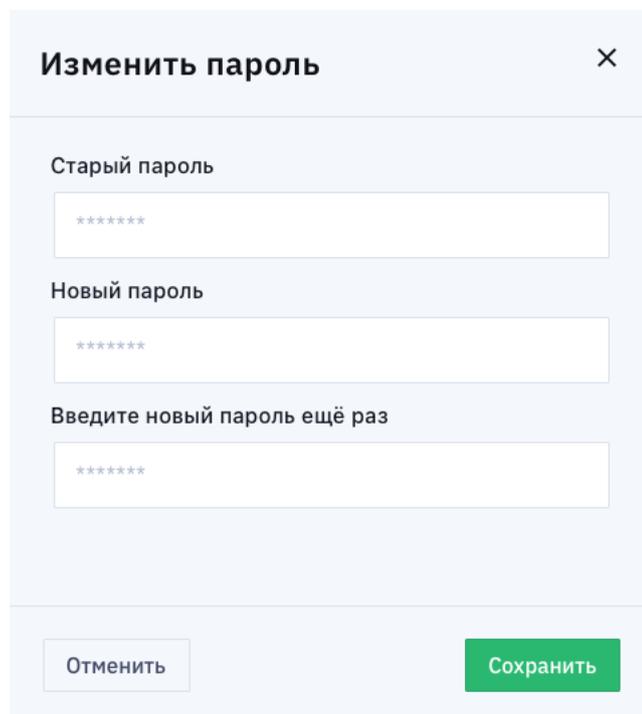


Рисунок 5



**Изменить пароль** ×

Старый пароль  
\*\*\*\*\*

Новый пароль  
\*\*\*\*\*

Введите новый пароль ещё раз  
\*\*\*\*\*

Отменить Сохранить

Рисунок 6

#### 4.1.2 Функция выхода из учетной записи

Для выхода из учетной записи необходимо нажать кнопку «Выйти» (Рисунок 7) – доступна на любом экране Платформы. После нажатия кнопки «Выйти» откроется окно входа в Платформу (Рисунок 1).

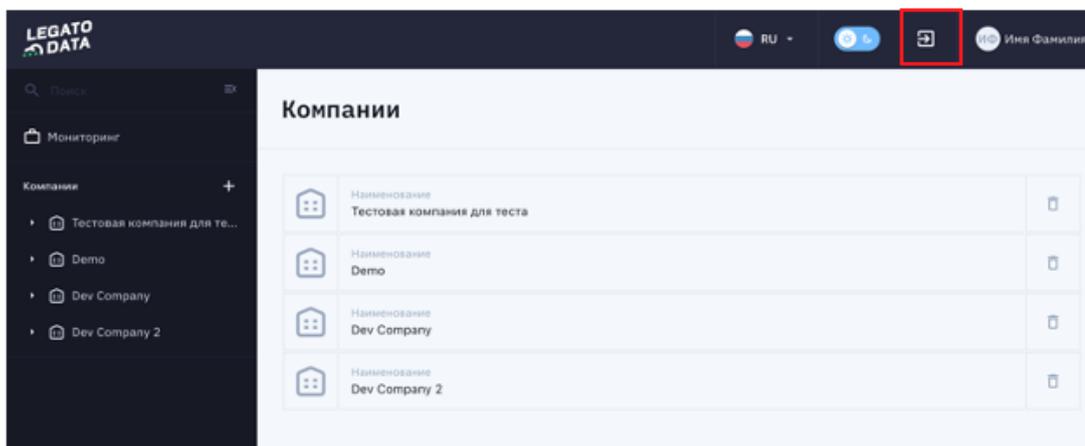


Рисунок 7

## 4.2 Функции управления отображения пользовательского интерфейса

### 4.2.1 Функция смены темы оформления пользовательского интерфейса

Платформа предлагает на выбор две темы оформления пользовательского интерфейса (светлая и темная). Для переключения темы оформления необходимо нажать кнопку «Смена темы» (Рисунок 8). После нажатия кнопки тема оформления изменяется (Рисунок 9).

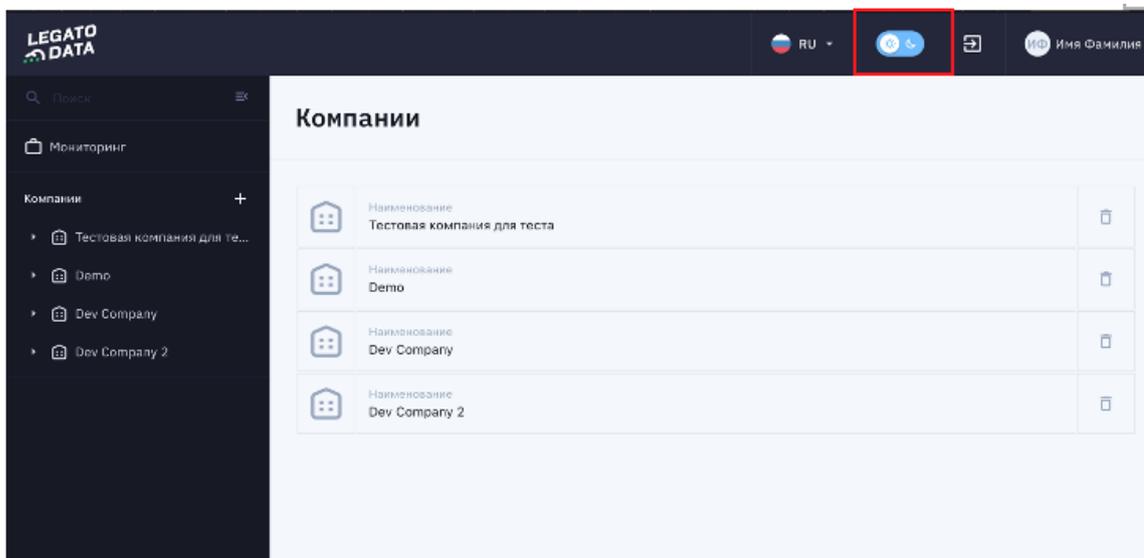


Рисунок 8

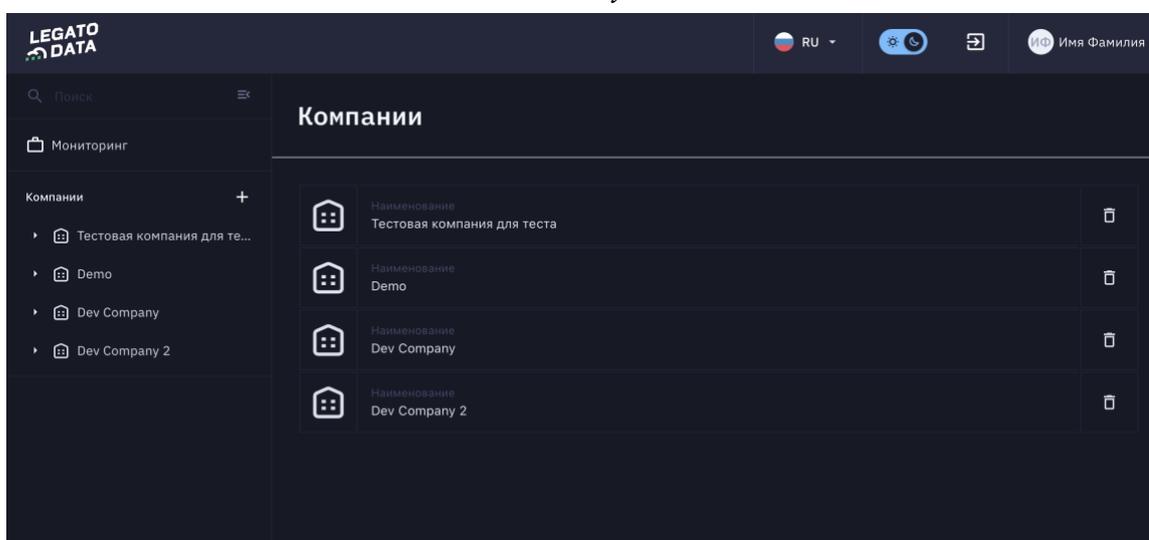


Рисунок 9

#### 4.2.2 Функция смены языка оформления пользовательского интерфейса

Платформа предлагает на выбор два языка пользовательского интерфейса (русский и английский). Для переключения языка оформления необходимо нажать кнопку «Смена языка» и выбрать язык из списка языков (Рисунок 10). После нажатия кнопки и выбора языка язык пользовательского интерфейса изменяется (Рисунок 11).

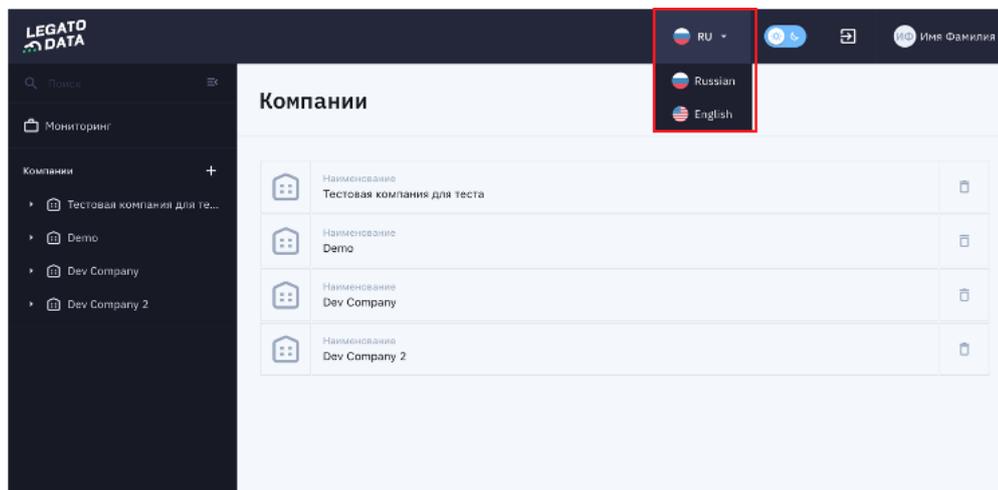


Рисунок 10

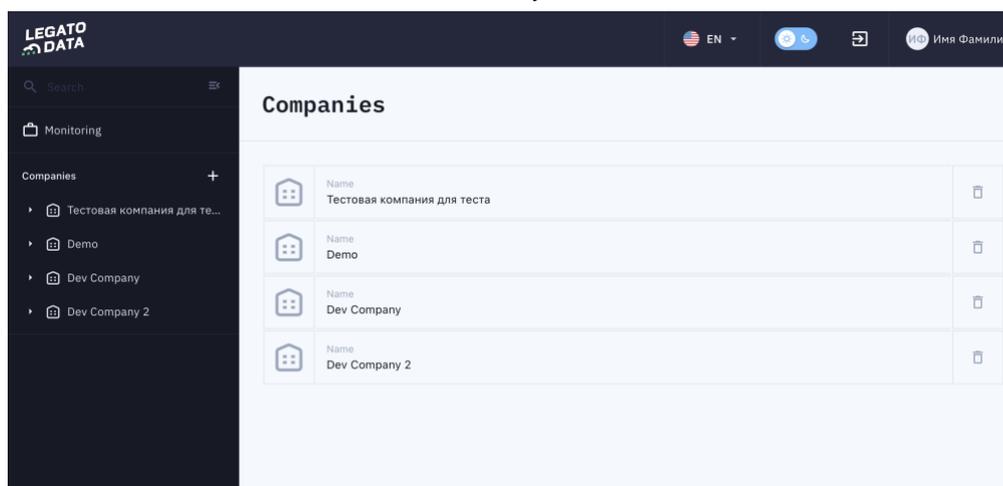


Рисунок 11

## 4.3 Функции настройки области отображения пайплайна и вспомогательных сущностей

### 4.3.1 Функция просмотра и редактирования перечня Компаний

Перечень доступных пользователю Компаний представлен пользователю сразу при авторизации в Платформе (Рисунок 2). Для удаления Компании необходимо нажать кнопку «Удалить» (Рисунок 12). После нажатия на кнопку Компания удаляется из перечня Компаний, изменения сохраняются в Платформе.

Для создания Компании в перечне Компаний необходимо нажать кнопку «Добавить» в основном (левом) меню Платформы (Рисунок 13). При нажатии на кнопку открывается модальное окно «Создание компании» (Рисунок 14), которое содержит:

1. Поле ввода наименования компании;
2. Поле ввода описания компании;

3. Кнопка «Отменить» - при нажатии кнопки модальное окно закрывается, изменения не сохраняются в Платформе;
4. Кнопка «Сохранить» - при нажатии кнопки модальное окно закрывается, Компания создается, изменения сохраняются в Платформе.

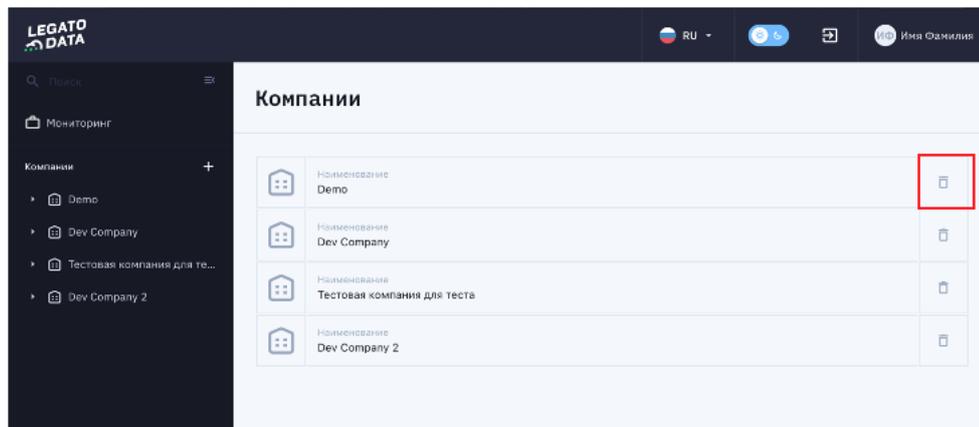


Рисунок 12

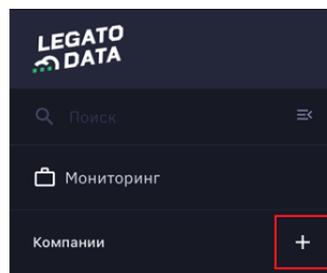


Рисунок 13

**Создание компании** [X]

Наименование

Описание  
 0/255

Рисунок 14

#### 4.3.2 Функция просмотра и редактирования общей информации для Компании

Для просмотра и редактирования общей информации о Компании необходимо перейти в Карточку Компании (Рисунок 15), для этого необходимо выбрать Компанию в перечне Компаний (Рисунок 2).

Для просмотра и редактирования общей информации о Компании, необходимо перейти во вкладку «Информация» Карточки Компании (Рисунок 16). Вкладка содержит в себе:

1. Поле для просмотра/редактирования наименования компании;
2. Поле для просмотра даты регистрации Компании в Платформе;
3. Поле для просмотра/редактирования описания Компании;
4. Кнопка «Редактировать». При нажатии на кнопку вкладка переходит в режим редактирования (Рисунок 17), кнопка «Редактировать» пропадает, появляются кнопки «Отменить» (для отмены изменений в результате редактирования) и «Сохранить» (для сохранения изменений в Системе), поля вкладки становятся доступными для редактирования.

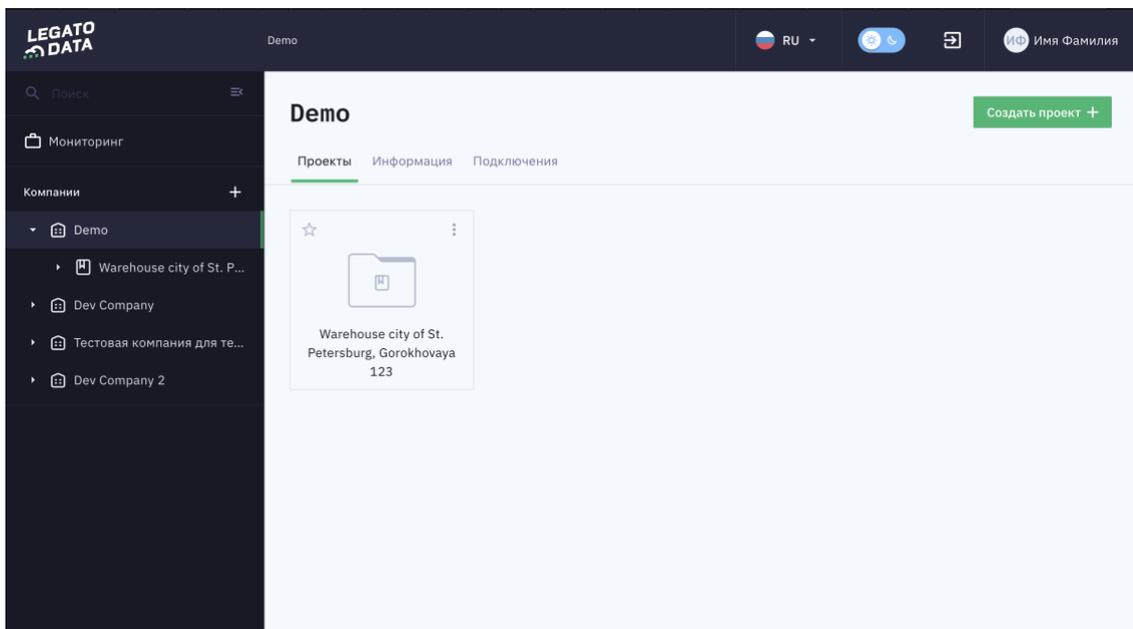


Рисунок 15

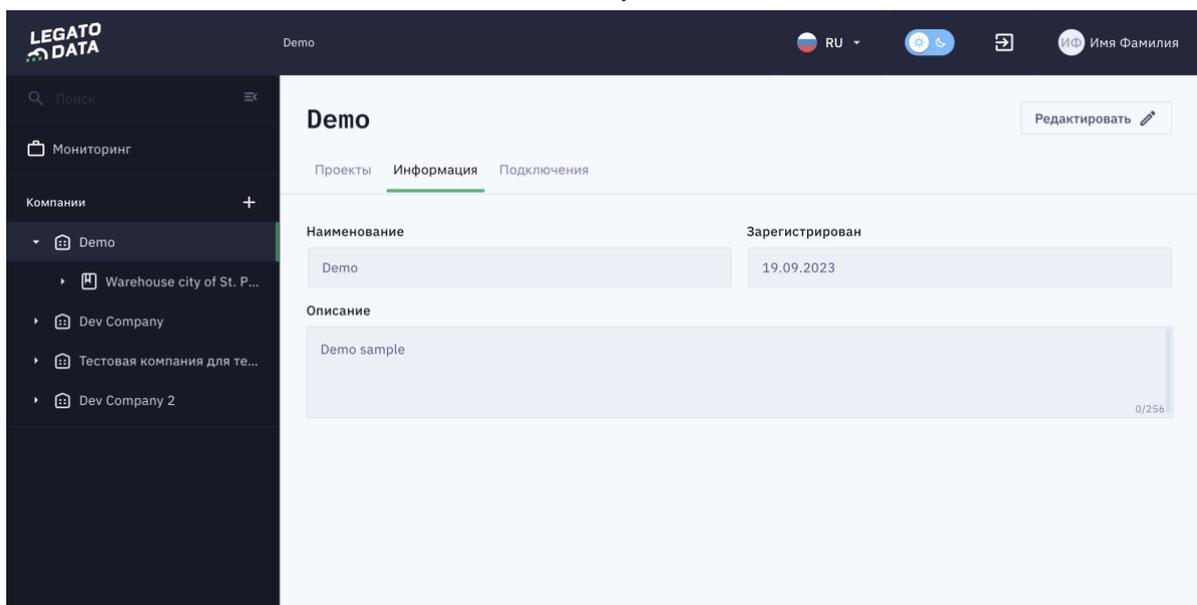


Рисунок 16

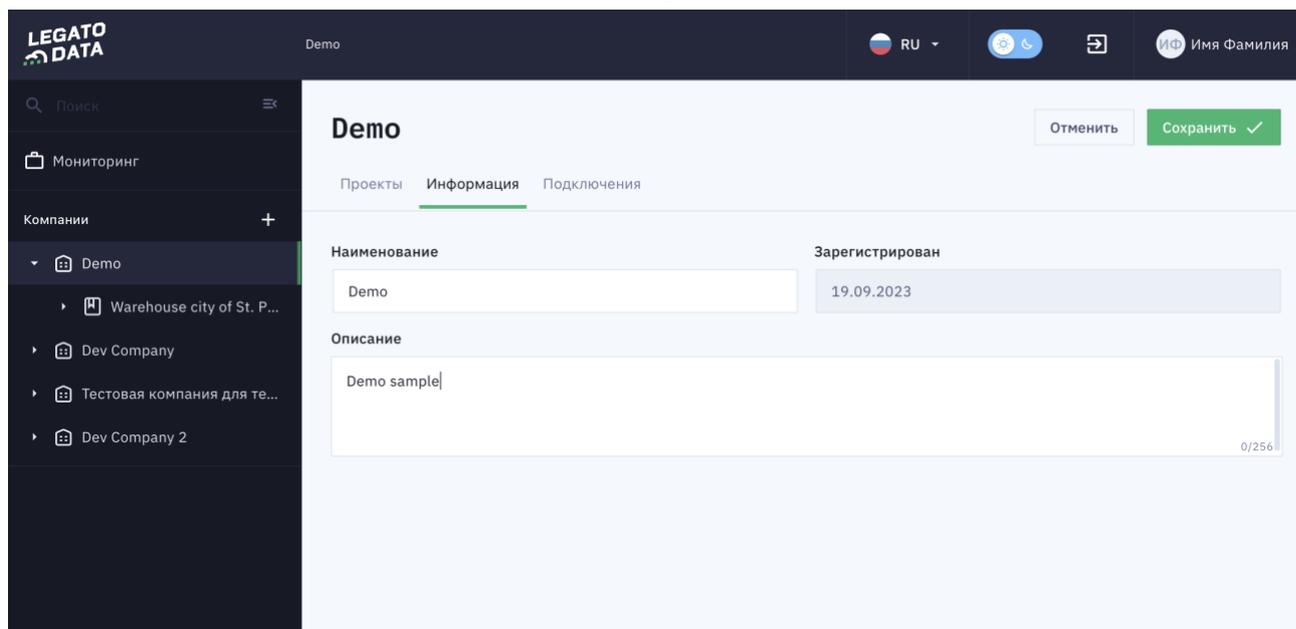


Рисунок 17

### 4.3.3 Функция просмотра и редактирования Подключений в рамках Компании

Для просмотра и редактирования общей информации о Компании необходимо перейти в Карточку Компании (Рисунок 15), для этого необходимо выбрать Компанию в перечне Компаний (Рисунок 2). Далее необходимо выбрать вкладку «Подключения» (Рисунок 18). Вкладка содержит следующие компоненты:

1. Кнопка «Добавить подключение» (описано ниже в этом же разделе);
2. Выпадающий список выбора типа сортировки Подключений в перечне Подключений по Компании (возможны сортировка по наименованию Подключения или по дате последнего обновления Подключения);
3. Иконка типа подключения;
4. Наименование Подключения;
5. Наименование технологии (типа подключения – назначается при создании);
6. Владелец Подключения (назначается автоматически при создании);
7. Дата последнего обновления Подключения (назначается автоматически при редактировании Подключения);
8. Кнопка «Удалить» (при нажатии на кнопку подключение удаляется из Платформы, исключение составляют «системные Подключения» - Spark и Kafka, их нельзя редактировать и удалять);
9. Кнопка «Редактировать» (при нажатии открывается модальное окно изменения Подключения (описано ниже в этом же разделе).

При нажатии на кнопку «Добавить подключение» открывается модальное окно «Создание подключения» (Рисунок 19). Модальное окно содержит:

1. Наименование подключения – поле для редактирования наименования Подключения;
2. Тип подключения – поле для выбора Типа подключения. В результате выбора типа Подключения ниже отображаются поля в зависимости от выбранного типа Подключения (пример отображен на Рисунок 20);
3. Кнопка «Тест» - при нажатии на кнопку «Тест» производится проверка Подключения, в случае успешной проверки становится активной кнопка «Сохранить»;
4. Кнопка «Сохранить» - при нажатии на кнопку «Сохранить» информация о Подключении сохраняется в Платформе, модальное окно закрывается;
5. Кнопка «Отменить» - при нажатии на кнопку «Отменить» информация о Подключении не сохраняется в Платформе, модальное окно закрывается.

При нажатии на кнопку «Редактировать подключение» открывается модальное окно «Редактирование подключение», аналогичное модальному окну «Создание подключения» (Рисунок 19) по содержанию и функционалу.

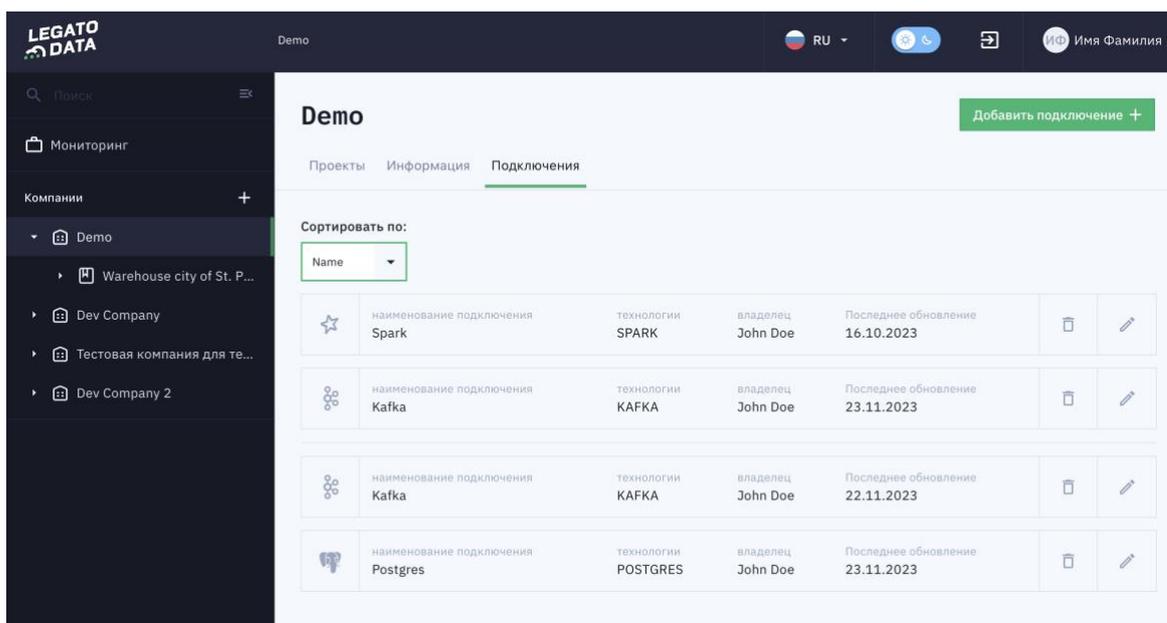


Рисунок 18

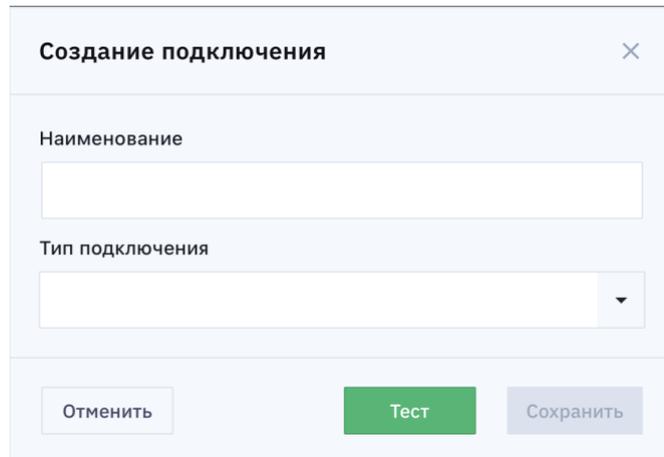


Рисунок 19

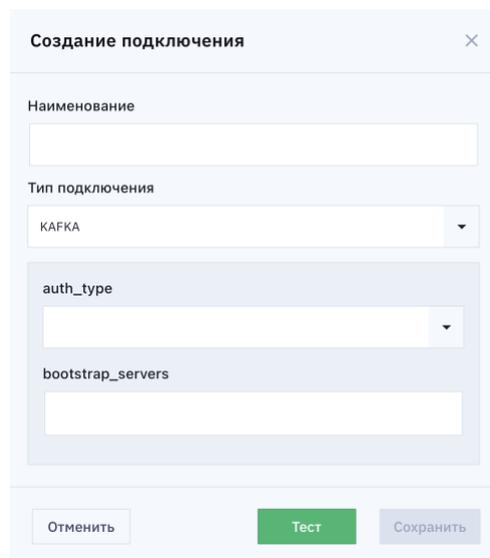


Рисунок 20

#### 4.3.4 Функция просмотра и редактирования перечня Проектов в рамках Компании

Для просмотра и редактирования перечня Проектов в рамках Компании необходимо перейти в Карточку Компании (Рисунок 15), для этого необходимо выбрать Компанию в перечне Компаний (Рисунок 2). На вкладке «Проекты» (Рисунок 15) будет отображаться перечень Проектов. Каждый проект отображается в виде системной папки с наименованием и кнопкой «Действия» с выпадающим списком возможных действий по проекту (Рисунок 21):

1. Редактировать (описано [здесь](#));
2. Обновить (при нажатии обновляется наименование Проекта, если оно до этого было изменено другим пользователем Платформы);
3. Удалить (при нажатии Проект удаляется из Платформы со всеми входящими в него Пайплайнами).

Так же вкладка «Проекты» содержит кнопку «Создать проект». При нажатии на кнопку открывается модальное окно «Создание проекта» (Рисунок 22). Модальное окно содержит:

1. Поле ввода наименования Проекта;
2. Поле ввода описания Проекта;
3. Кнопка «Сохранить» (доступна при введенном наименовании Проекта) – при нажатии Проект сохраняется в Платформе, модальное окно закрывается;
4. Кнопка «Отменить» – при нажатии Проект не сохраняется в Платформе, модальное окно закрывается;

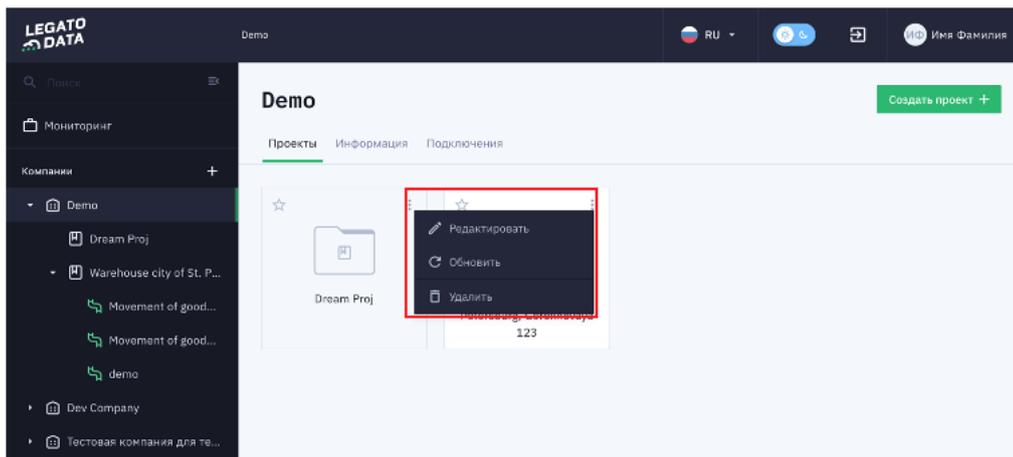


Рисунок 21

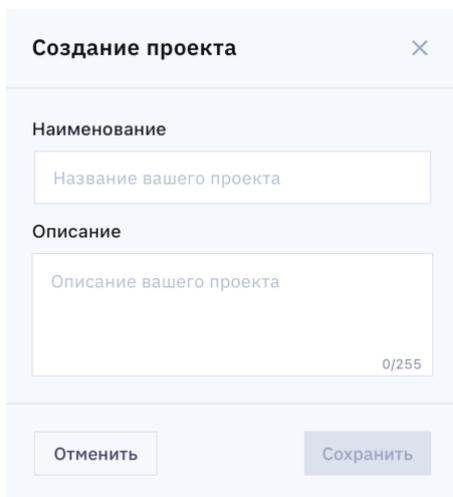


Рисунок 22

#### 4.3.5 Функция просмотра и редактирование информации о Проекте

Для просмотра и редактирования информации о Проекте необходимо выполнить одно из следующих действий:

1. Перейти в Карточку Проекта, выбрав Проект в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта, выбрав Проект во вкладке «Проекты» Карточки Компании;

3. Перейти в Карточку Проекта, выбрав «Редактировать» в кнопке действий по Проекту во вкладке «Проекты» Карточки Компании.

В результате любого из этих действий Платформа предоставляет Карточку Проекта (Рисунок 23). Для просмотра и редактирования общей информации о Проекте необходимо перейти во вкладку «Информация» на Карточке Проекта (Рисунок 24).

Вкладка содержит в себе:

1. Поле для просмотра/редактирования наименования Проекта;
2. Поле для просмотра даты регистрации Проекта в Платформе;
3. Поле для просмотра/редактирования описания Проекта;
4. Кнопка «Редактировать». При нажатии на кнопку вкладка переходит в режим редактирования (Рисунок 25), кнопка «Редактировать» пропадает, появляются кнопки «Отменить» (для отмены изменений в результате редактирования) и «Сохранить» (для сохранения изменений в Платформе), поля вкладки становятся доступными для редактирования.

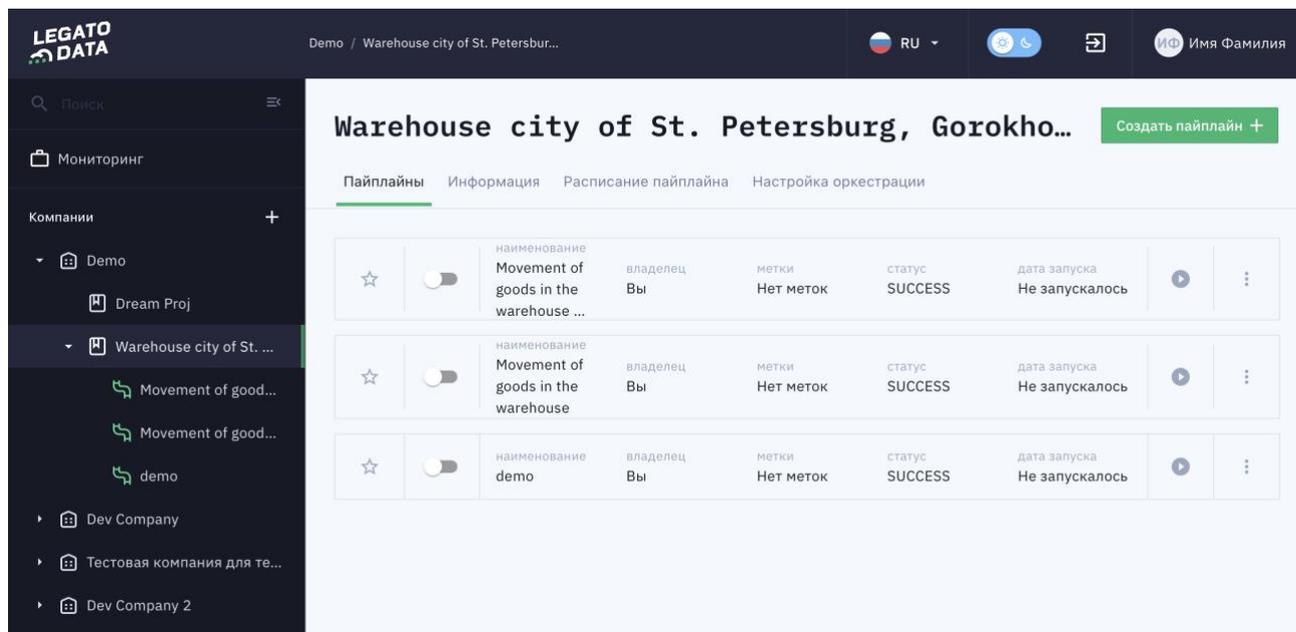


Рисунок 23

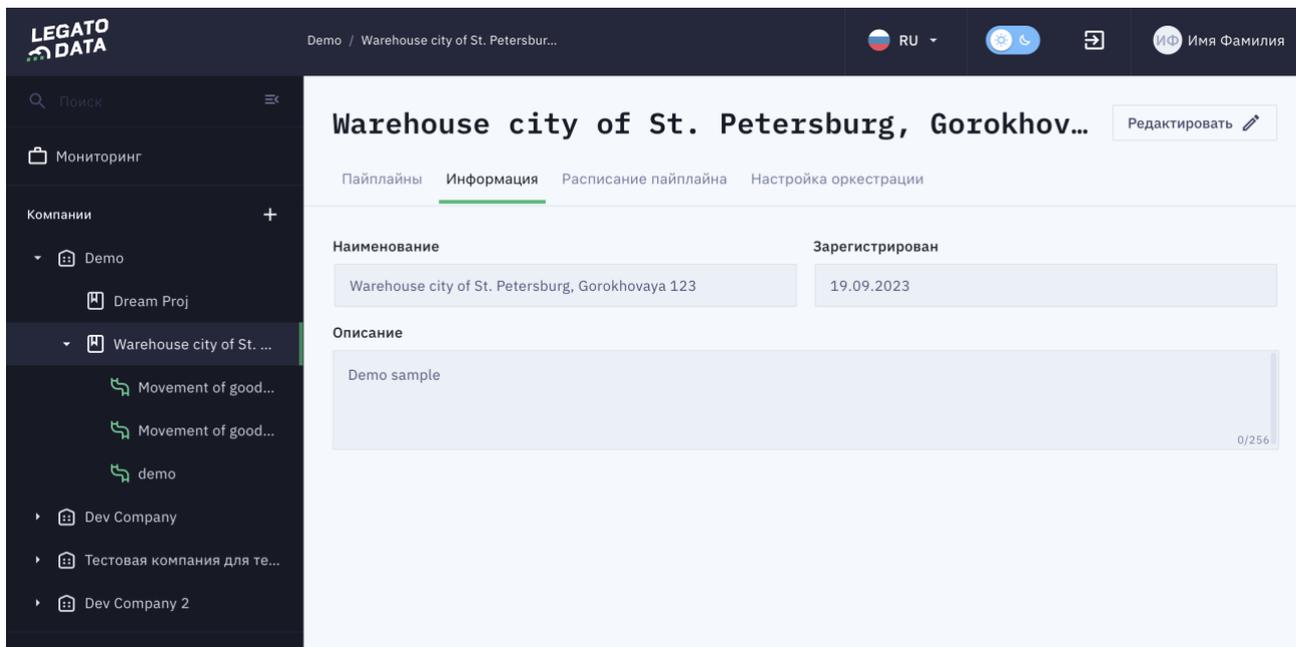


Рисунок 24

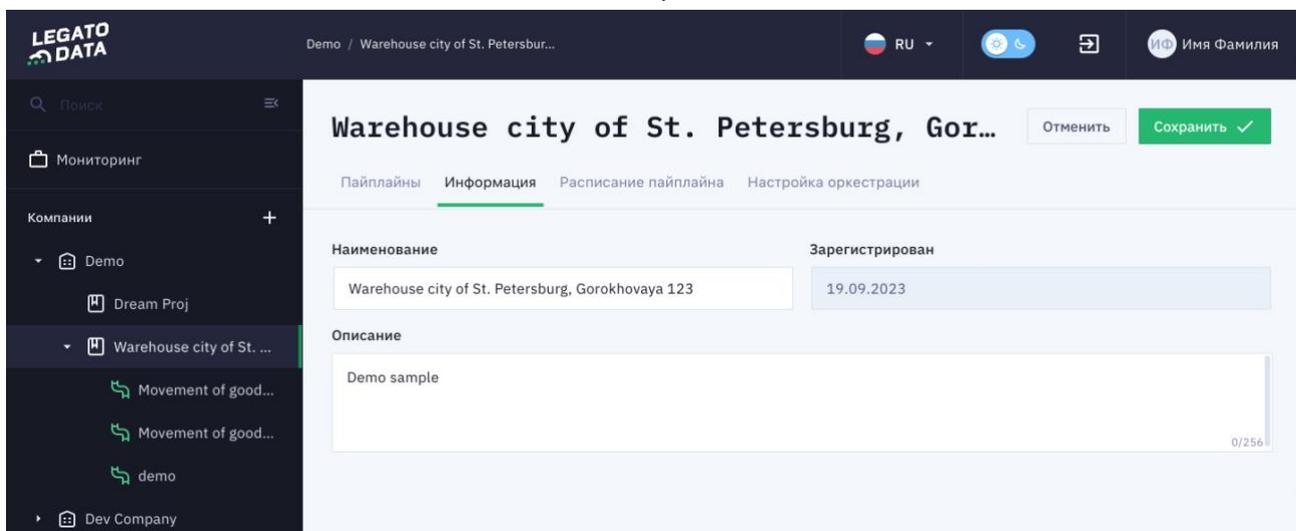


Рисунок 25

#### 4.3.6 Функция настройки расписания Пайплайнов по-умолчанию в рамках Проекта

Для просмотра информации о Проекте необходимо выполнить одно из следующих действий:

1. Перейти в Карточку Проекта, выбрав Проект в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта, выбрав Проект во вкладке «Проекты» Карточки Компании;
3. Перейти в Карточку Проекта, выбрав «Редактировать» в кнопке действий по Проекту во вкладке «Проекты» Карточки Компании.

В результате любого из этих действий Платформа предоставляет Карточку Проекта (Рисунок 23). Для просмотра и редактирования расписания Пайплайна в рамках Проекта необходимо перейти во вкладку «Расписание пайплайна» (Рисунок 26). Данная вкладка по-

умолчанию открывается в режиме «Расписание пайплайна задается условиями». Так же есть режим «Расписание пайплайна задается скриптом» (Рисунок 27).

При выборе режима «Расписание пайплайна задается условиями» отображается перечень условий (если он уже есть по данному Пайплайну) и кнопка «Добавить условие». При нажатии на данную кнопку добавляется новое условие (при этом данное условие по умолчанию выключено до момента полной настройки данного условия. При этом в левой части каждого условия есть возможность включения/выключения условия. В правой части каждого условия в перечне условий располагается кнопка «Удалить условие». При нажатии на данную кнопку должно открываться модальное окно с текстом подтверждения удаления, при подтверждении пользователем условие удаляется.

Описание настроек представлено в Таблица 5:

№п/п	Наименование	Описание
1	Номер условия	Автоинкрементный номер условия для отображения. Отображается в формате «Условие {Номер условия}». Не редактируемое поле
2	Номер периода	Обязательное для ввода поле. По-умолчанию у нового условия проставляется значение «1». Целое число от 1 до 500
3	Тип периода	Обязательное для ввода поле. Выпадающий список. Возможные значения:  1. День 2. Неделя 3. Месяц

№п/п	Наименование	Описание
		В зависимости от выбора данного поля отображаются дальнейшие поля настройки условия
4	Тип периода «День»	
4.1	Время запуска пайплайна	Обязательное для заполнения поля. По умолчанию, значение «00:00».. При заполненном значении является конечным полем настройки условия
5	Тип периода «Неделя»	
5.1	День недели	Обязательное для заполнения поле из predetermined значений:  <ol style="list-style-type: none"> <li>1. пн</li> <li>2. вт</li> <li>3. ср</li> <li>4. чт</li> <li>5. пт</li> <li>6. сб</li> <li>7. вс</li> </ol> Возможен множественный выбор.
5.2	Время запуска пайплайна	Обязательное для заполнения поля. Доступно при выборе хотя бы одного дня недели. По умолчанию, значение «00:00». При заполненном значении является конечным полем настройки условия
6	Тип периода «Месяц»	

№п/п	Наименование	Описание
6.1	Какого числа/в какой день	Выбор из вариантов «Какого числа» и «В какой день». По умолчанию, используется значение «Какого числа».
6.1.1	Тип периода «Месяц»/«Какого числа»	
6.1.1.1	Число	Число месяца для запуска пайплайна, выбор целого числа от 1 до 31. По умолчанию, значение = «1»
6.1.1.2	Время запуска пайплайна	Обязательное для заполнения поля. По умолчанию, значение «00:00». При заполненном значении является конечным полем настройки условия
6.1.2	Тип периода «Месяц»/«В какой день»	
6.1.2.1	Номер периода месяца	Обязательное для заполнения поле. Выпадающий список со значениями:  <ol style="list-style-type: none"> <li>1. 1 (по-умолчанию)</li> <li>2. 2</li> <li>3. 3</li> <li>4. 4</li> <li>5. 5</li> <li>6. Первый</li> <li>7. Последний</li> </ol>
6.1.2.2	Период месяца	Обязательное для заполнения поле.

№п/п	Наименование	Описание
		<p>Выпадающий список со значениями:</p> <ol style="list-style-type: none"> <li>1. пн</li> <li>2. вт</li> <li>3. ср</li> <li>4. чт</li> <li>5. пт</li> <li>6. сб</li> <li>7. вс</li> <li>8. день (по-умолчанию)</li> <li>9. будний</li> <li>10. выходной</li> </ol>
6.1.2.3	Время запуска пайплайна	<p>Обязательное для заполнения поля. По-умолчанию, значение «00:00». При заполненном значении является конечным полем настройки условия</p>

Таблица 5

При выборе режима «Расписание пайплайна задается скриптом» отображается область ввода скрипта, кнопка «Сохранить». Введенное пользователем значение скрипта записывается в случае нажатия кнопки «Сохранить».

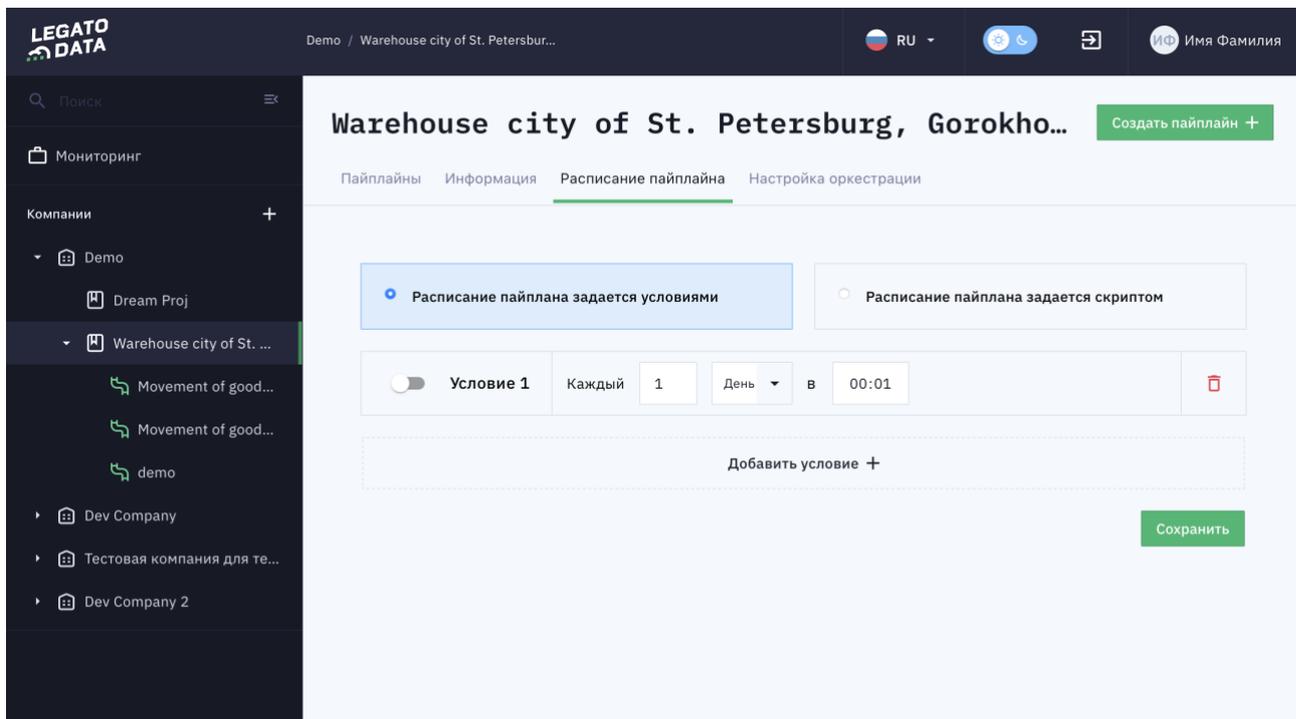


Рисунок 26

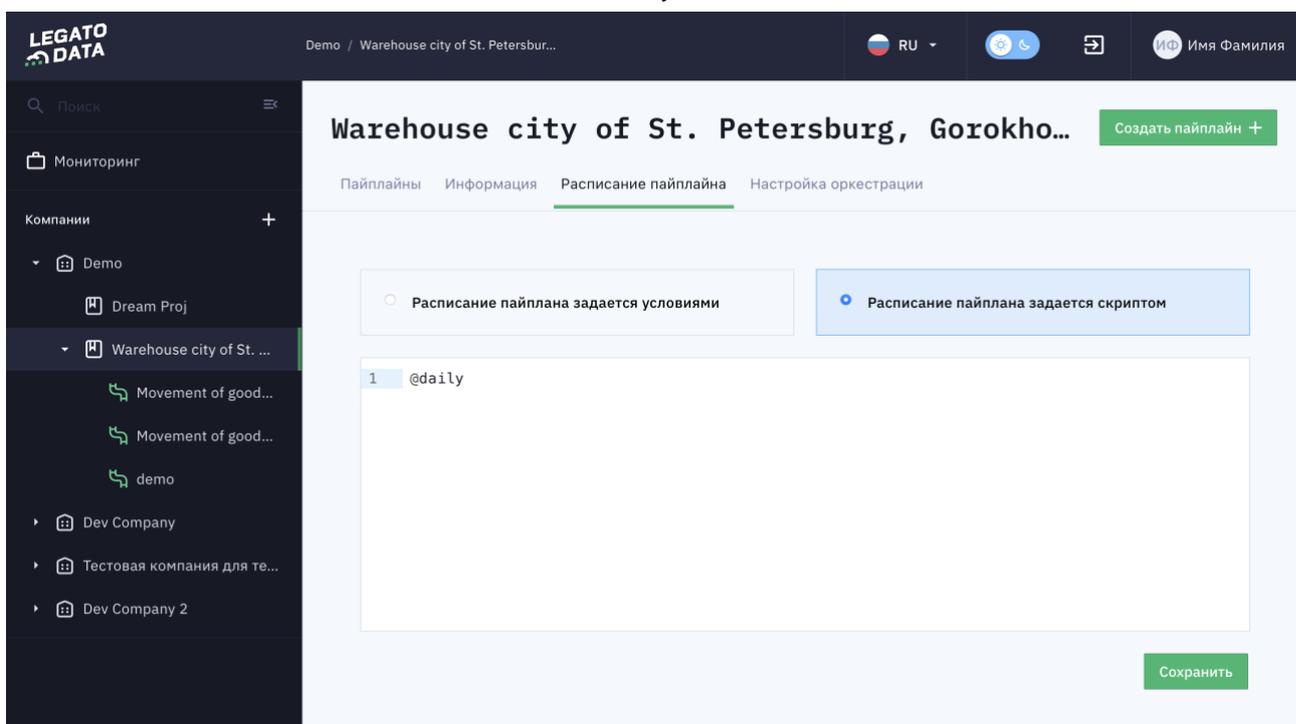


Рисунок 27

#### 4.3.7 Функция настройки оркестрации Пайплайнов по-умолчанию в рамках Проекта

Для просмотра информации о Проекте необходимо выполнить одно из следующих действий:

1. Перейти в Карточку Проекта, выбрав Проект в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта, выбрав Проект во вкладке «Проекты» Карточки Компании;

3. Перейти в Карточку Проекта, выбрав «Редактировать» в кнопке действий по Проекту во вкладке «Проекты» Карточки Компании.

В результате любого из этих действий Платформа предоставляет Карточку Проекта (Рисунок 23). Для просмотра и редактирования настроек оркестрации Пайплайна в рамках Проекта необходимо перейти во вкладку «Настройка оркестрации» Карточки Проекта (Рисунок 28). Вкладка содержит в себе следующие поля для просмотра и редактирования (редактирование доступно в режиме редактирования (Рисунок 29) после нажатия на кнопку «Редактировать» - в случае нажатия далее кнопки «Сохранить», изменения сохраняются в Платформе):

1. Приоритет задачи;
2. Дата старта данных;
3. Дата финиша данных;
4. Максимальное количество параллельных процессов в рамках задачи;
5. Максимальное количество параллельных запусков одной задачи;
6. Количество попыток;
7. Пауза между попытками.

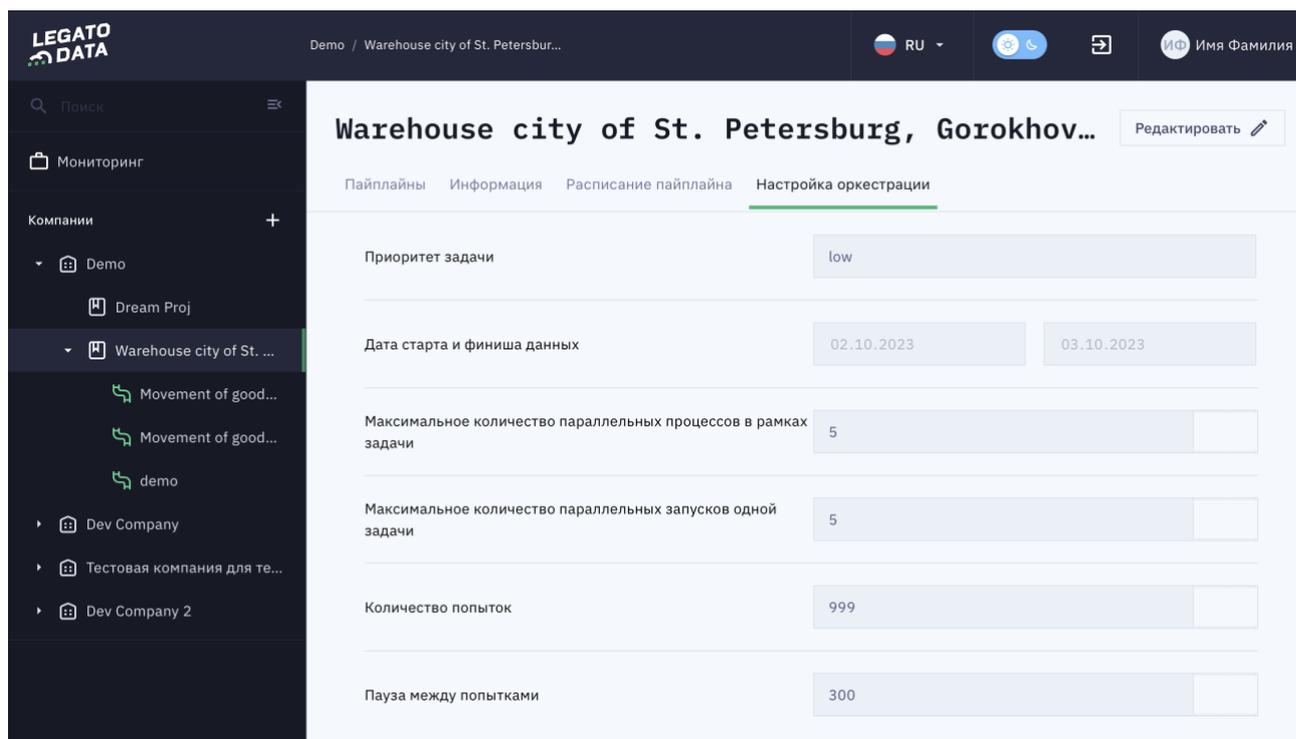


Рисунок 28

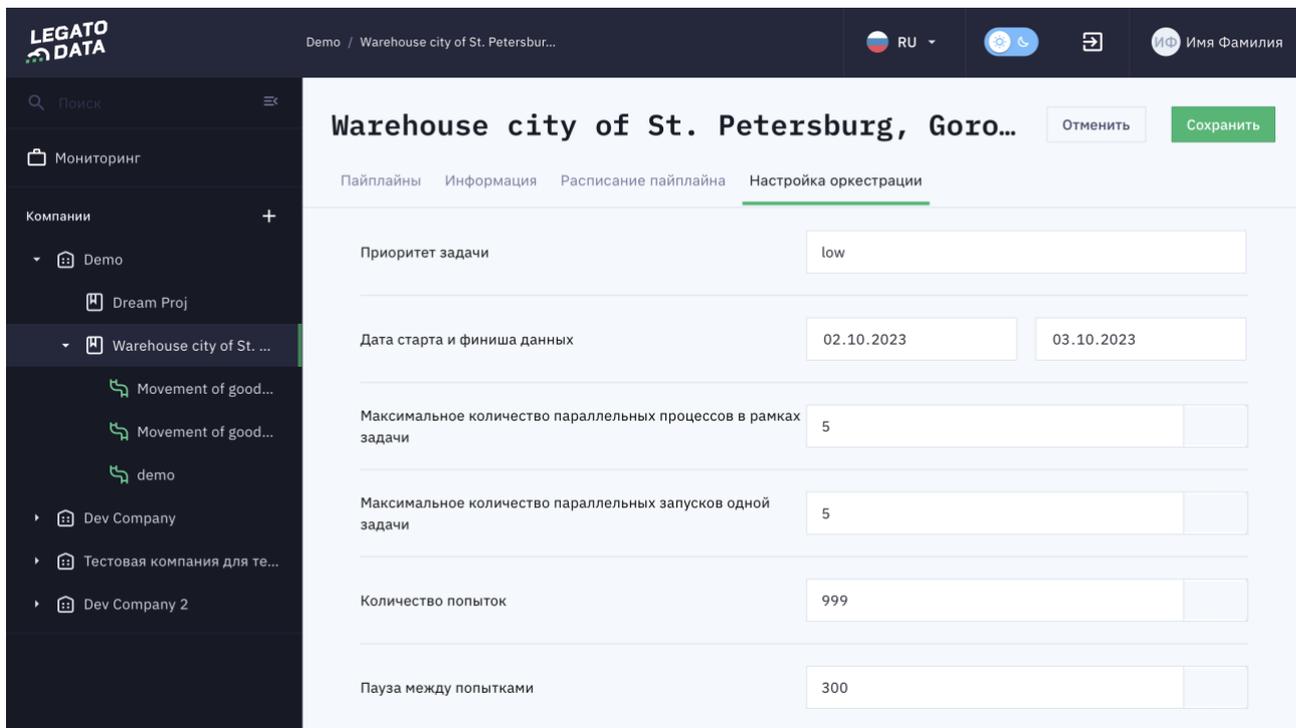


Рисунок 29

#### 4.3.8 Функция просмотра и редактирования перечня Пайплайнов в рамках Проекта

Для просмотра информации о Проекте необходимо выполнить одно из следующих действий:

1. Перейти в Карточку Проекта, выбрав Проект в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта, выбрав Проект во вкладке «Проекты» Карточки Компании;
3. Перейти в Карточку Проекта, выбрав «Редактировать» в кнопке действий по Проекту во вкладке «Проекты» Карточки Компании.

В результате любого из этих действий Платформа предоставляет Карточку Проекта (Рисунок 23). Для просмотра и редактирования перечня Пайплайнов в рамках Проекта необходимо перейти во вкладку «Пайплайны» Карточки Проекта (Рисунок 23) – открывается по-умолчанию при переходе в Карточку Проекта. Данная вкладка содержит перечень Пайплайнов внутри Проекта и кнопки быстрых действий с Пайплайнами:

1. Наименование пайплайна;
2. Владелец пайплайна;
3. Метки (теги) пайплайна;
4. Статус последнего запуска;
5. Дата последнего запуска;
6. Кнопка мгновенного запуска (при нажатии Пайплайн запускается);
7. Перечень быстрых действий по Пайплайну:
  - а) Редактировать (при нажатии осуществляется переход в Карточку Пайплайна);

- b) Обновить (при нажатии обновляется информация о Пайплайне на странице);
  - c) Удалить (при нажатии останавливается выполнение Пайплайна, в случае, если он в данный момент выполняется, Пайплайн удаляется из Платформы).
8. Кнопка «Создать пайплайн» (описано [здесь](#)).

#### **4.3.9    Функция создания Пайплайна**

Для просмотра информации о Проекте необходимо выполнить одно из следующих действий:

1. Перейти в Карточку Проекта, выбрав Проект в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта, выбрав Проект во вкладке «Проекты» Карточки Компании;
3. Перейти в Карточку Проекта, выбрав «Редактировать» в кнопке действий по Проекту во вкладке «Проекты» Карточки Компании.

В результате любого из этих действий Платформа предоставляет Карточку Проекта (Рисунок 23). Для создания Пайплайна в рамках Проекта необходимо перейти во вкладку «Пайплайны» Карточки Проекта (Рисунок 23) – открывается по-умолчанию при переходе в Карточку Проекта и нажать кнопку «Создать пайплайн». При нажатии на данную кнопку открывается модальное окно «Создание пайплайна» (Рисунок 30). Модальное окно содержит:

1. Поле ввода наименования Пайплайна;
2. Поле ввода описания Пайплайна;
3. Кнопка «Сохранить» (доступна при введенном наименовании Пайплайна) – при нажатии Пайплайн сохраняется в Платформе, модальное окно закрывается;
4. Кнопка «Отменить» – при нажатии Пайплайн не сохраняется в Платформе, модальное окно закрывается;

Создание пайплайна

Наименование

Название вашего пайплайна

Описание

Описание вашего пайплайна

0/255

Отменить

Сохранить

Рисунок 30

#### 4.3.10 Функция просмотра и редактирования информации о Пайплайне

Для просмотра информации о Проекте необходимо выполнить одно из следующих действий:

1. Перейти в Карточку Проекта, выбрав Проект в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта, выбрав Проект во вкладке «Проекты» Карточки Компании;
3. Перейти в Карточку Проекта, выбрав «Редактировать» в кнопке действий по Проекту во вкладке «Проекты» Карточки Компании.

В результате любого из этих действий Платформа предоставляет Карточку Проекта (Рисунок 23). Для создания Пайплайна в рамках Проекта необходимо перейти во вкладку «Пайплайны» Карточки Проекта (Рисунок 23) – открывается по-умолчанию при переходе в Карточку Проекта и нажать кнопку «Редактировать пайплайн». При нажатии на данную кнопку открывается модальное окно «Редактирование пайплайна» (аналогично модальному окну «Создание пайплайна» - Рисунок 30). Модальное окно содержит:

1. Поле ввода наименования Пайплайна;
2. Поле ввода описания Пайплайна;
3. Кнопка «Сохранить» (доступна при введенном наименовании Пайплайна) – при нажатии Пайплайн сохраняется в Платформе, модальное окно закрывается;
4. Кнопка «Отменить» – при нажатии Пайплайн не сохраняется в Платформе, модальное окно закрывается;

## **4.4 Функции запуска и настроек запуска Пайплайна**

### **4.4.1 Функция настройки оркестрации Пайплайна**

Для перехода в Карточку Пайплайна пользователю необходимо произвести одно из следующих действий:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Для настройки оркестрации Пайплайна пользователю необходимо перейти во вкладку «Настройка оркестрации» в Карточке Пайплайна. По-умолчанию, настройки повторяют настройки оркестрации в Карточке Проекта, к которому относится Пайплайн (описано [здесь](#)). Механика изменений настроек оркестрации аналогично изменению настроек оркестрации в Карточке Проекта, к которому относится Пайплайн (описано [здесь](#)).

### **4.4.2 Функция настройки расписания Пайплайна**

Для перехода в Карточку Пайплайна пользователю необходимо произвести одно из следующих действий:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Для настройки расписания Пайплайна пользователю необходимо перейти во вкладку «Настройка расписания» в Карточке Пайплайна. По-умолчанию, настройки повторяют настройки расписания в Карточке Проекта, к которому относится Пайплайн (описано [здесь](#)). Механика изменений настроек расписания Пайплайна аналогично изменению настроек расписания Пайплайна в Карточке Проекта, к которому относится Пайплайн (описано [здесь](#)).

### **4.4.3 Функция постановки Пайплайна на исполнение**

Для перехода в Карточку Пайплайна пользователю необходимо произвести одно из следующих действий:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

По-умолчанию открывается вкладка «Конструктор пайплайна» (Рисунок 31), при этом конструктор Пайплайнов содержит кнопки «Поставить в расписание» и «Старт» (Рисунок 32). При нажатии на кнопку «Поставить в расписание», пайплайн ставится в

расписание и начнет исполняться по расписанию, настроенному ранее (описано [здесь](#) и [здесь](#)), при нажатии на кнопку «Старт», Пайплайн начнет выполняться немедленно, в случае, если он еще не запущен на текущий момент. Если Пайплайн запущен, Платформа выдаст ошибку «Пайплайн запущен».

*При постановке Пайплайна на выполнение по расписанию существует ограничение на остановку выполнения на 30 секунд. В обратную сторону работает так же.*

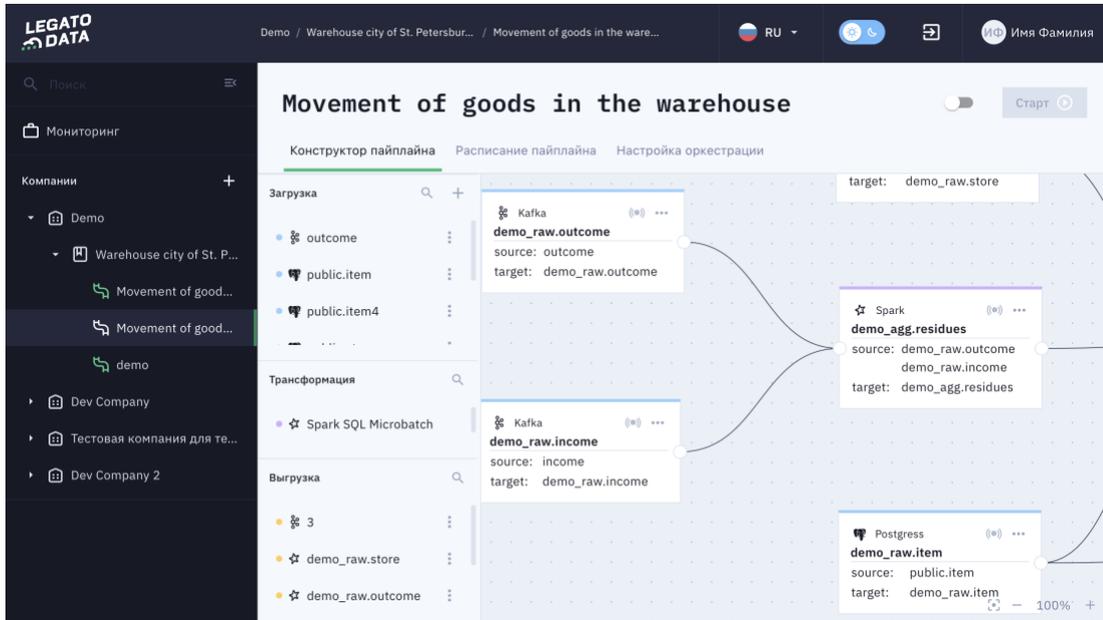


Рисунок 31

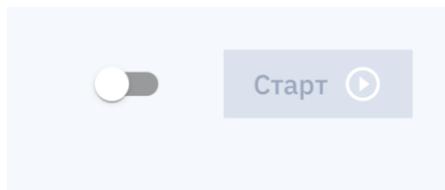


Рисунок 32

## 4.5 Функции работы с задачами загрузки/выгрузки данных

### 4.5.1 Функция создания набора данных для задач типа «Загрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

3. Выбрать Пайплайн в основном (левом) меню Платформы;
4. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна.

Для создания набора данных (датасета) типа «Загрузка» необходимо нажать кнопку «Добавить набор данных типа «Загрузка» (Рисунок 33). После этого откроется модальное окно «Создание набора данных» (Рисунок 34). Данное модальное окно состоит из двух вкладок (шагов) – «Основные» (Рисунок 34). и «Схема» (Рисунок 35).

Вкладка «Основные» необходима для указания необходимого источника данных (например, таблица в БД Postgres или топик в Kafka). В нижней части вкладки «Основные» должны располагаться кнопки «Отменить», «Тест», «Далее»:

1. При нажатии на кнопку «Отменить» пользователем, модальное окно должно закрываться, изменения в системе не должны сохраняться. Кнопка доступна всегда;
2. При нажатии на кнопку «Тест» система должна осуществлять попытку получения мета-информации о датасете (например, атрибутный состав таблицы Postgres или атрибутный состав топика Kafka). Кнопка доступна, если заполнены все обязательные поля (как из статических, так и динамических полей);
3. При нажатии на кнопку «Далее» (доступна только при успешном прохождении тестирования подключения и получения мета информации о датасете) осуществляется переход во вкладку «Схема» (вкладка «Схема» так же не доступна до успешного тестирования).

Вкладка «Схема» необходима для предоставления информации о схеме датасета в источнике (например, схема таблица БД Postgres или топик Kafka, откуда далее будет забираться информация). Вкладка так же содержит кнопки «Отменить», «Сохранить»:

1. При нажатии на кнопку «Отменить» пользователем, модальное окно должно закрываться, изменения в системе не должны сохраняться. Кнопка доступна всегда;
2. Кнопка «Назад», при нажатии на которую осуществляется переход на вкладку «Основные»;
3. При нажатии на кнопку «Сохранить», модальное окно закрывается, датасет создается в системе.

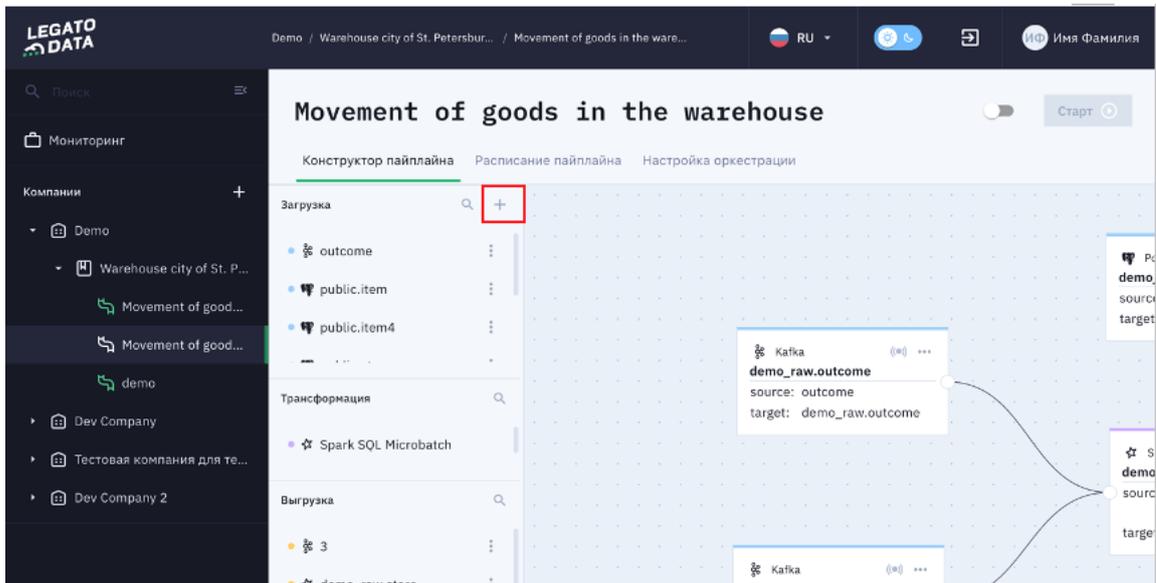


Рисунок 33

The screenshot shows a dialog box titled "Создание набора данных" (Create Data Set) with a close button (X) in the top right corner. The dialog has two tabs: "Основные" (Main) and "Схема" (Schema). The "Основные" tab is selected. It contains the following fields:

- "Наименование" (Name): A text input field.
- "Тип подключения" (Connection Type): A dropdown menu.
- "Подключение" (Connection): A dropdown menu.

Below the fields, there is a note: "Обратите внимание что, после заполнения данных, необходимо провести тестирование." (Pay attention that, after filling in the data, it is necessary to perform testing). At the bottom of the dialog, there are three buttons: "Отменить" (Cancel), "Тест" (Test), and "Далее" (Next).

Рисунок 34

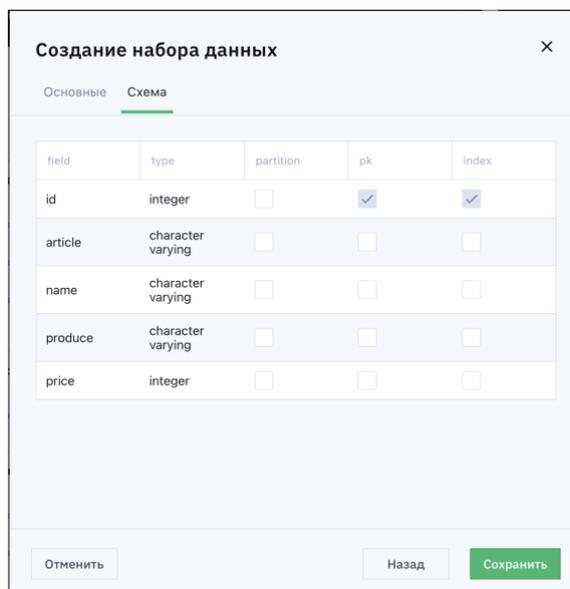


Рисунок 35

#### 4.5.2 Функция редактирования набора данных для задач типа «Загрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. В перечне существующих наборов данных типа «Загрузка» необходимо нажать на кнопку «Быстрые действия» у нужного набора данных и выбрать там действие «Редактировать» (Рисунок 36). После этого откроется модальное окно «Редактирование набора данных», аналогичное модальному окну «Создание набора данных» по отображению и функционалу (Рисунок 34) – описано [здесь](#).

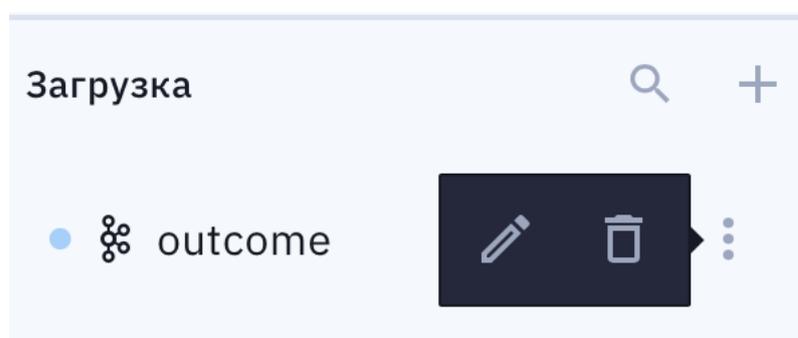


Рисунок 36

### 4.5.3 Функция удаления набора данных для задач типа «Загрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. В перечне существующих наборов данных типа «Загрузка» необходимо нажать на кнопку «Быстрые действия» у нужного набора данных и выбрать там действие «Удалить» (Рисунок 36). После нажатия на кнопку «Удалить», Платформа удалит выбранный набор данных, если на его основе не создано ни одной задачи. В случае, если существует задача, основанная на данном наборе данных, Платформа отобразит ошибку «Набор данных используется».

### 4.5.4 Функция создания задачи типа «Загрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. Для создания задачи загрузки данных необходимо, чтобы уже был создан хотя бы один набор данных типа «Загрузка» (процесс создания описан [здесь](#)). В случае, если пользователь собирается создать задачу загрузки данных на основе существующего набора данных, необходимо перетянуть набор данных на рабочую область конструктора пайплайна, после чего откроется окно «Extract Task» (Рисунок 37). Целью работы с данным модальным окном является создание задачи импорта данных, для чего необходимо выбрать тип задачи (поточковая или пакетная), название задачи, создание target-датасета в одном из внутренних хранилищ (source-датасет автоматически назначается тот, который пользователь перенес на рабочую область конструктора пайплайна на предыдущем пункте), настройка типа выполнения задачи (полная или интервальная загрузка):

1. Создание потоковой задачи импорта данных:
  - a) Пользователь переводит флаг «Stream task» в положение «true» (от данного значения зависят отображаемые поля ниже);
  - b) Пользователь вводит наименование задачи;

- c) Пользователь вводит наименование топика Kafka (куда будут записываться данные при потоковом импорте данных);
  - d) Пользователь нажимает кнопку «Далее» или переходит во вкладку «Схема»;
  - e) Пользователю отображается схема данных из датасета-назначения;
  - f) Пользователь нажимает кнопку «Сохранить»;
  - g) Модальное окно закрывается, задача отображается на рабочей области конструктора пайплайна в виде ярлыка задачи на рабочей области конструктора Пайплайна (далее-ноды);
2. Создание пакетной задачи импорта данных:
- a) Пользователь переводит флаг «Stream task» в положение «false» (от данного значения зависят отображаемые поля ниже);
  - b) Пользователь вводит наименование задачи;
  - c) Пользователь вводит наименования БД и таблицы внутреннего хранилища Spark (куда будут записываться данные при пакетном импорте данных);
  - d) Пользователь выбирает тип перегрузки (полная или интервальная):
    - i) При выборе полного типа перегрузки:
      - (1) Пользователь правит значение размера пакета - «fetch size» Значение по умолчанию = 100;
    - ii) При выборе интервального типа перегрузки:
      - (1) Пользователь правит значение размера пакета - «fetch size» Значение по умолчанию = 100;
      - (2) Пользователь заполняет значение «interval» - время, за которое должна происходить загрузка (в днях);
      - (3) Пользователь заполняет значение «start\_date» - дата начала интервала;
      - (4) Пользователь заполняет значение «created\_at\_column» - колонка данных в исходной таблице с типом date/datetime, на основе которой отбираются данные за указанный интервал;
  - e) Пользователь нажимает кнопку «Далее» или пользователь переходит во вкладку «Схема» Открывается вкладка «Схема», отображается схема таблицы хранилища Spark, куда будут записываться импортированные данные;
  - f) Пользователь выбирает поля, необходимые для импорта/записи данных;
  - g) Пользователь заполняет описания полей;
  - h) Пользователь выбирает поля, по которым необходимо партиционировать данные;
  - i) Пользователь нажимет кнопку «Сохранить»;

- j) Модальное окно закрывается, задача отображается на рабочей области конструктора пайплайна в виде ноды.

The screenshot shows a modal window titled "Extract task" with a close button in the top right corner. Below the title bar, there are two tabs: "Основные" (Active) and "Схема". The main content area contains a "Stream task" toggle switch, a text input field for "Наименование набора данных источника" (Source data set name) with the value "public.item4", a "Task name" input field, a "Table" input field, and a "Database" input field. At the bottom, there are two buttons: "Отменить" (Cancel) and "Далее" (Next).

Рисунок 37

#### 4.5.5 Функция редактирования задачи типа «Загрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. После чего необходимо нажать кнопку быстрых действий на ноде выбранной задачи на рабочей области Конструктора Пайплайна и нажать кнопку «Редактировать» (Рисунок 38). После выполненных действий откроется модальное окно «Edit task» (Рисунок 39). Модальное окно должно быть при открытии заполнено значениями, назначенными в Системе ранее.

Модальное окно состоит из двух вкладок (они же «шаги» модального окна):

1. Основные;
2. Схема.

Вкладка «Основные» содержит кнопки «Далее» и «Отменить». При нажатии на кнопку «Отменить» пользователем, модальное окно закрывается, изменения не сохраняются в системе.

При нажатии на кнопку «Далее» производится проверка на заполнение обязательных полей модального окна, осуществляется переход на вкладку «Схема»

Вкладка «Схема» содержит кнопки «Сохранить» и «Отменить». При нажатии на кнопку «Отменить» пользователем, модальное окно закрывается, изменения не сохраняются в системе. При нажатии на кнопку «Сохранить» производится проверка на заполнение обязательных полей модального окна (обоих шагов), осуществляется сохранение изменений в Системе.

В остальном функционал редактирования задачи схож с функцией создания задачи (описано [здесь](#)).

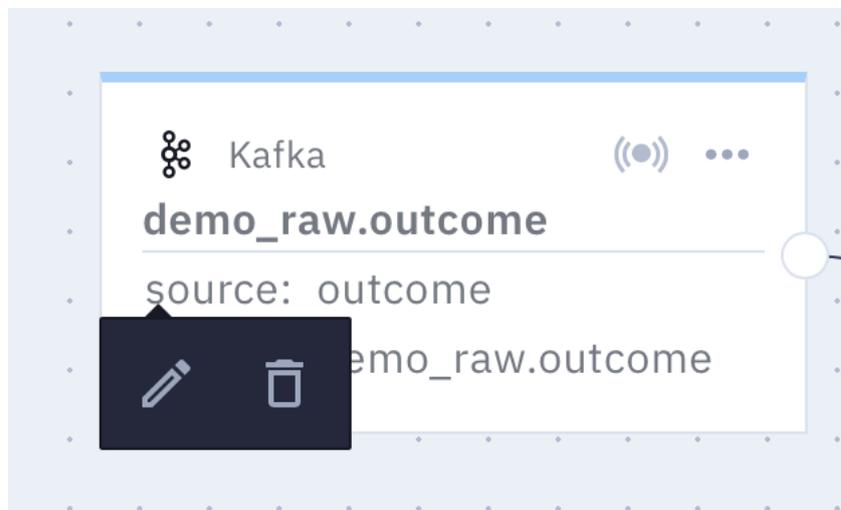


Рисунок 38

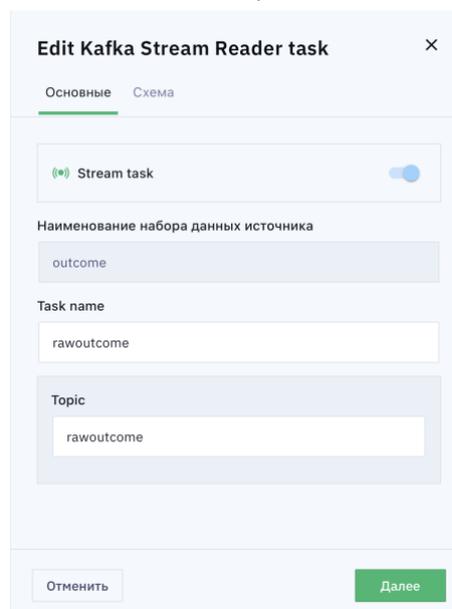


Рисунок 39

#### 4.5.6 Функция удаления задачи типа «Загрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. После чего необходимо нажать кнопку быстрых действий на ноде выбранной задачи на рабочей области Конструктора Пайплайна и нажать кнопку «Удалить» (Рисунок 38). После выполненных действий задача удалится из Платформы, при этом Платформа выдаст информационное сообщение «Задача удалена».

#### 4.5.7 Функция редактирования набора данных типа «Выгрузка»

В случае создания задачи загрузки, трансформации или выгрузки данных пользователь создает датасет - назначение в рамках создания задачи. В случае создания датасета через кнопку, описываемую в рамках текущей статьи, создается набор данных - источник для будущих задач загрузки или выгрузки данных. Механизмы создания подобных задач описаны [здесь](#), [здесь](#) и [здесь](#). Механизм редактирования набора данных типа «Выгрузка» аналогичен механизму редактирования набора данных типа «Загрузка» (описан [здесь](#)).

#### 4.5.8 Функция создания задачи типа «Выгрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. Для создания задачи выгрузки данных необходимо, чтобы уже был создан хотя бы один набор данных типа «Выгрузка» (процесс создания описан [здесь](#)). В случае, если пользователь собирается создать задачу выгрузки данных на основе существующего набора данных, необходимо перетянуть набор данных на рабочую область конструктора пайплайна, после чего откроется окно «Выгрузка данных» (Рисунок 40). Целью работы с данным модальным окном является создание задачи выгрузки, для чего необходимо выбрать тип задачи (потокковая или пакетная), название задачи, создание набора данных-назначения в

одно из возможных внешних или внутренних хранилищ данных (при этом возможно использование как существующего набора данных, так и нового):

1. Создание пакетной задачи экспорта данных (не подразумевается выбор типа задачи, тип задачи определяется набором данных - источником):
  - a) Пользователь выбирает тип подключения;
  - b) Пользователь выбирает соединение;
  - c) Пользователь вводит наименование задачи;
  - d) Пользователь создает/выбирает набор данных-назначения:
    - i) Новый:
      - (1) Пользователь нажимает кнопку «Новый»;
      - (2) Пользователь в ручном режиме вводит опции места записи данных (например, схема и таблица датасета-назначения);
      - (3) Пользователь заполняет значение поля «Размер пакета» («fetch size»);
    - ii) Существующий:
      - (1) Пользователь нажимает кнопку «Существующий»;
      - (2) Пользователь выбирает опции места записи данных (например схема и таблица датасета-назначения из области видимости пользователя);
      - (3) Пользователь заполняет значение поля «Размер пакета» («fetch size»);
  - e) Пользователь нажимает кнопку «Тест» Выполняется попытка получения мета-информации о схеме выбранной на предыдущем этапе таблицы БД;
  - f) Пользователь нажимает кнопку «Далее» или переходит во вкладку «Схема» Открывается вкладка «Схема», отображается информация о схеме датасета-назначения;
  - g) Пользователь нажимает «Сохранить»;
  - h) Модальное окно закрывается, задача отображается на рабочей области конструктора пайплайна в виде ноды. В случае создания нового датасета он отображается в перечне плагинов в разделах «Загрузка» и «Выгрузка»
2. Создание потоковой задачи экспорта данных (не подразумевается выбор типа задачи, тип задачи определяется набором данных - источником):
  - a) Пользователь выбирает тип подключения;
  - b) Пользователь выбирает соединение;
  - c) Пользователь вводит наименование задачи;
  - d) Пользователь создает/выбирает набор данных - назначение:
    - i) Новый:
      - (1) Пользователь нажимает кнопку «Новый»;

- (2) Пользователь в ручном режиме вводит опции места записи данных (например, схема и таблица датасета-назначения);
- ii) Существующий:
- (1) Пользователь нажимает кнопку «Существующий»;
- (2) Пользователь выбирает опции места записи данных (например схема и таблица датасета-назначения из области видимости пользователя);
- e) Пользователь нажимает кнопку «Тест» Выполняется попытка получения мета информации о схеме выбранной на предыдущем этапе таблицы БД;
- f) Пользователь нажимает кнопку «Далее» или переходит во вкладку «Схема» Открывается вкладка «Схема», отображается информация о схеме набора данных - назначения;
- g) Пользователь нажимает «Сохранить»;
- h) Модальное окно закрывается, задача отображается на рабочей области конструктора пайплайна в виде ноды. В случае создания нового набора данных он отображается в перечне плагинов в разделах «Загрузка» и «Выгрузка»

**Выгрузка данных** ×

Основные **Схема**

Наименование набора данных источника

rawitem

Task name Тип подключения

Подключение

Обратите внимание что, после заполнения данных, необходимо провести тестирование.

Существующий  Новый

Отменить Тест Далее

Рисунок 40

#### 4.5.9 Функция редактирования задачи типа «Выгрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;

2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. После чего необходимо нажать кнопку быстрых действий на ноде выбранной задачи на рабочей области Конструктора Пайплайна и нажать кнопку «Редактировать» (Рисунок 38). После выполненных действий откроется модальное окно «Edit task» (Рисунок 41). Модальное окно должно быть при открытии заполнено значениями, назначенными в Системе ранее.

Модальное окно состоит из двух вкладок (они же «шаги» модального окна):

1. Основные;
2. Схема.

Вкладка «Основные» содержит кнопки «Далее», «Отменить» и «Тест»:

1. При нажатии на кнопку «Отменить» пользователем, модальное окно закрывается, изменения не сохраняются в системе.
2. При нажатии на кнопку «Далее» производится на заполнение обязательных полей модального окна, осуществляется переход на вкладку «Схема».
3. При нажатии на кнопку «Тест» производится проверка подключения;

Вкладка «Схема» содержит кнопки «Сохранить» и «Отменить». При нажатии на кнопку «Отменить» пользователем, модальное окно закрывается, изменения не сохраняются в системе. При нажатии на кнопку «Сохранить» производится проверка на заполнение обязательных полей модального окна (обоих шагов), осуществляется сохранение изменений в Системе. В остальном функционал редактирования задачи схож с функцией создания задачи (описано [здесь](#)).

Рисунок 41

#### 4.5.10 Функция удаления задачи типа «Выгрузка»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. После чего необходимо нажать кнопку быстрых действий на ноде выбранной задачи на рабочей области Конструктора Пайплайна и нажать кнопку «Удалить» (Рисунок 38). После выполненных действий задача удалится из Платформы, при этом Платформа выдаст информационное сообщение «Задача удалена».

#### 4.6 Функции работы с задачами трансформации данных

##### 4.6.1 Функция создания задачи типа «Трансформация»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. Для создания задачи трансформации данных необходимо перенести шаблон трансформации данных (Рисунок 43) на рабочую область Конструктора Пайплайнов (шаблон выбирается в зависимости от необходимого типа задач). В результате Платформа откроет модальное окно создания задачи трансформации данных «Transform task» (Рисунок 43).

1. Создание задачи пакетного преобразования:

- a) Система открывает модальное окно создания/редактирования задачи трансформации данных;
- b) Пользователь вводит наименование задачи;
- c) Пользователь вводит место, куда записывать данные результата задачи трансформации (набор данных-назначение). Для пакетной обработки указываются «Наименование БД» и «Наименование таблицы» внутреннего хранилища Spark;
- d) Пользователь вводит код SQL-запроса для обработки данных;
- e) Пользователь заполняет значение поля “partition\_column\_write“ - колонка/колонки, по которым должно производиться партиционирование, формат ввода - без кавычек, в случае массива колонок - через запятую;

2. Создание задачи потокового преобразования:

- a) Система открывает модальное окно создания/редактирования задачи трансформации данных;
- b) Пользователь вводит наименование задачи;
- c) Пользователь вводит место, куда записывать данные результата задачи трансформации (датасет-назначение). Для потоковой обработки указываются «Топик Kafka» внутренней Kafka LegatoData для хранения данных, требующих пакетного ETL;
- d) Пользователь вводит код SQL-запроса для обработки данных;

3. Пользователь нажимает кнопку «Сохранить»;

4. Модальное окно закрывается, задача отображается на рабочей области конструктора пайплайна в виде ноды. В случае создания нового набора данных он отображается в перечне плагинов в разделе «Выгрузка».

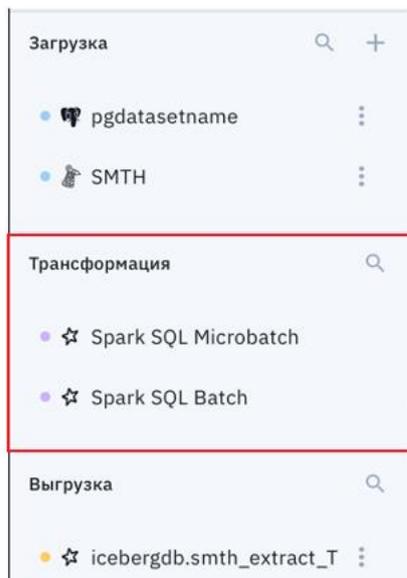


Рисунок 42

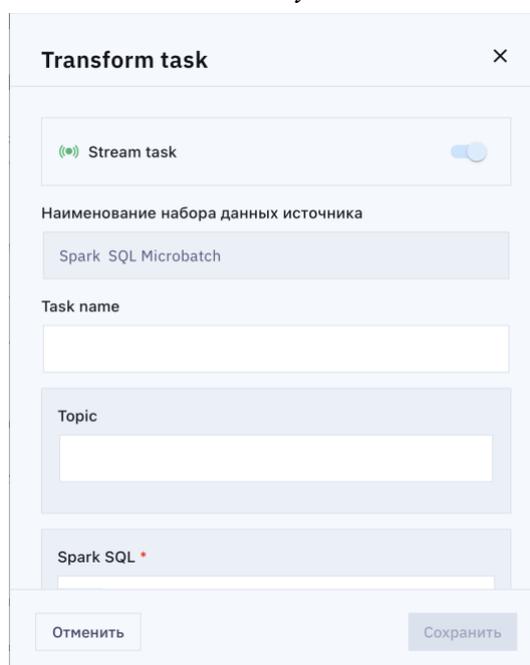


Рисунок 43

#### 4.6.2 Функция редактирования задачи типа «Трансформация»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. После чего необходимо нажать кнопку быстрых действий на ноде выбранной задачи на рабочей области Конструктора Пайплайна и нажать кнопку «Редактировать»

(Рисунок 38). После выполненных действий откроется модальное окно «Edit task» (Рисунок 44). Модальное окно должно быть при открытии заполнено значениями, назначенными в Системе ранее.

Модальное окно состоит из одной вкладки.

Модальное окно содержит кнопки «Отменить» и «Сохранить»:

1. При нажатии на кнопку «Отменить» пользователем, модальное окно закрывается, изменения не сохраняются в Системе.
2. При нажатии на кнопку «Сохранить» пользователем, модальное окно закрывается, изменения (если они есть), сохраняются в Системе.

В остальном функционал редактирования задачи схож с функцией создания задачи (описано [здесь](#)).



Рисунок 44

#### 4.6.3 Функция удаления задачи типа «Трансформация»

Для перехода в Конструктор Пайплайна пользователю необходимо произвести одно из следующих действий для перехода в Карточку Пайплайна:

1. Выбрать Пайплайн в основном (левом) меню Платформы;
2. Перейти в Карточку Проекта и выбрать вкладку «Пайплайны», после чего нажать на строку с необходимым Пайплайном.

Далее необходимо перейти во вкладку «Конструктор пайплайна» Карточки Пайплайна. После чего необходимо нажать кнопку быстрых действий на ноде выбранной задачи на рабочей области Конструктора Пайплайна и нажать кнопку «Удалить» (Рисунок 38). После выполненных действий задача удалится из Платформы, при этом Платформа выдаст информационное сообщение «Задача удалена».

## 4.7 Функции мониторинга

### 4.7.1 Мониторинг выполнения задач

Для мониторинга выполнения задач необходимо выбрать пункт основного (левого) меню Платформы «Мониторинг» (Рисунок 45). В результате откроется окно мониторинга выполняемых задач (Рисунок 46).

В верхней части формы мониторинга задач располагаются строка поиска и область фильтрации. Строка поиска работает по частичному совпадению с отображаемыми полями таблицы мониторинга. В случае, если введено значение поля из области вспомогательной таблицы (описано ниже), отображается строка таблицы перечня таблицы мониторинга задач (основной). Область фильтрации содержит три выпадающих списка - Компания, Проект, Пайплайн. В случае, если пользователю доступно только одно значение из выпадающего списка, по-умолчанию выбрано оно. В случае, если доступно несколько значений - состояние выпадающего списка = «Не выбрано», при этом в таблице отображаются все задачи. Каждый выпадающий список имеет возможность множественного выбора. Значения внутри списка сортируются в алфавитном порядке.

Описание таблиц перечня задач Spark представлено в Таблица 6:

№п/п	Наименование	Описание
1	Наименование	Наименование задачи
2	Компания	Наименование Компании, к которой относится Пайплайн
3	Пайплайн	Наименование Пайплайна, который инициирует задачу
4	Владелец	Владелец задачи
5	Статус	Статус задачи. Отображается с цветовой индикацией. Возможные варианты: 1. Running 2. Stopped 3. Success 4. Failed
6	Дата запуска	Дата последнего запуска задачи (без времени)
7	Тип	Тип процесса
8	Кнопка остановки задачи	Доступна только в статусе «Running», в остальных статусах «Disabled». при нажатии останавливается процесс

№п/п	Наименование	Описание
9	Шеврон раскрытия вспомогательной таблицы с описанием процессов	При нажатии вызывает отображение вспомогательной таблицы
10	Поля вспомогательной таблицы	
10.1	Версия	Версия процесса
10.2	Идентификатор задачи	Идентификатор процесса
10.3	Тип	Тип процесса
10.4	Началась	Дата и время запуска процесса
10.5	Закончилась	Дата и время окончания процесса
10.6	Продолжительность	Продолжительность процесса. Отображается в часах, минутах и секундах в формате hh:mm:ss
10.7	Последнее обновление	Дата и время последнего изменения процесса
10.8	Кнопка загрузки логов задачи	При нажатии на кнопку скачивается файл логов процесса
10.9	Кнопка перехода в окно подробной информации о задаче	При нажатии на кнопку осуществляется переход в окно подробной информации о задаче Spark. Окно должно содержать кнопку «Назад» (при нажатии на данную кнопку открывается форма перечня задач Spark, при этом она отображается в том же состоянии, из которого был осуществлен переход в окно подробной информации о задаче Spark - строка поиска/фильтры должны быть заполнены значениями, который вводил пользователь ранее, должны быть развернуты те же строки таблицы) и окно подробной информации о процессе Spark (встроенный плагин, интерфейс Spark)

Таблица 6

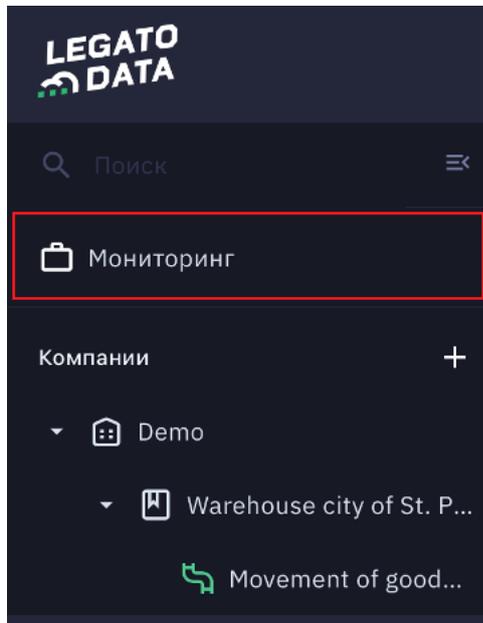


Рисунок 45

Наименование	Компания	Пайплайн	владелец	Статус	дата запуска	Тип
icebergdb.kkkwspw	Dev Company 2	asd	Dev Company 2	Completed	-	batch
Версия	Идентификатор задачи	Тип	Началась	Закончилась	Продолжительность	Последнее обновление
3.4.1	spark-f339598765ed4f69aa2f10a4df0e755	SparkApplication	29.11.2023 10.04	29.11.2023 10.05	1m 4s	29.11.2023 13.05
3.4.1	spark-a466dfc8cd9544259c3960b8cb9c967d	SparkApplication	29.11.2023 10.00	29.11.2023 10.01	51s	29.11.2023 13.01
3.4.1	spark-b60ee1fe37104799841c6300338716c5	SparkApplication	29.11.2023 09.53	29.11.2023 09.54	1m 8s	29.11.2023 12.54
3.4.1	spark-5d67fb1bdd854e8a80e2fea1c74ee79	SparkApplication	29.11.2023 09.52	29.11.2023 09.53	50s	29.11.2023 12.53
3.4.1	spark-cdfdea4ab78e4f8086bd9b68e28d1103	SparkApplication	29.11.2023 09.51	29.11.2023 09.52	1m 5s	29.11.2023 12.52
3.4.1	spark-757fb33c80345e5a5a1f1f32583c332	SparkApplication	29.11.2023 09.51	29.11.2023 09.52	1m 9s	29.11.2023 12.52
3.4.1	spark-e18d5efec6a0415ea1e5f40f88b6c90	SparkApplication	29.11.2023 09.51	29.11.2023 09.52	45s	29.11.2023 12.52
demo_raw.outcome	Demo	Movement of goods in the warehouse	Demo	Completed	-	batch
demo_raw.income	Demo	Movement of goods in the warehouse	Demo	Completed	-	batch

Рисунок 46