



Общество с ограниченной ответственностью «Платформа Данных
Селена» 123022, г. Москва, вн.тер.г. муниципальный округ
Пресненский, ул. 2-я Звенигородская, д. 13, стр. 37 ИНН 9703195303,
КПП 770301001, ОГРН 1247700725491

Платформа Данных Селена

РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ

Версия программы 3.0

Москва
2026

Аннотация

Данное руководство пользователя описывает программный комплекс «Платформа хранения и обработки данных Селена» (в дальнейшем Селена), включая его архитектуру, варианты использования, принципы развертывания, источники и целевые объекты, а также основные концепции.

В руководстве также описывается, как использовать интерфейс панели управления для настройки, запуска, мониторинга и администрирования.

Оглавление

1 Архитектура и компоненты

- 1.1 Мастер сервер
- 1.2 Вычислительный сервер
- 1.3 Панель управления кластером (Cluster Manager)
- 1.4 Модуль загрузки данных
- 1.5 Метакаталог данных
- 1.6 Мониторинг
- 1.7 Модуль среды разработки
- 1.8 Модуль хранения данных

2 Модуль управления кластером (Cluster Manager)

- 2.1 Вход в систему
- 2.2 Установка и обновление лицензии
- 2.3 Добавление виртуальных машин
- 2.4 Добавление фронтенд (мастер) узлов
- 2.5 Добавление бэкенд (вычислительных) узлов
- 2.6 Декомиссия и Удаление узлов
- 2.7 Конфигурирование параметров Кластера
- 2.8 Перезапуск вычислительного кластера
- 2.9 Удаление вычислительного кластера

3 Управление пользователями

- 3.1 Создание группы
- 3.2 Создание пользователя
- 3.3 Указание пароля пользователя
- 3.4 Смена пароля пользователя
- 3.5 Добавление пользователя в группу
- 3.6 Удаление пользователя
- 3.7 Удаление группы
- 3.8 Создание роли
- 3.9 Удаление роли
- 3.10 Предоставление полномочий для роли
- 3.11 Отзыв полномочий для роли
- 3.12 Назначение роли для группы пользователей
- 3.13 Отзыв роли для группы пользователей

4 Работа с модулем Мониторинга

- 4.1 Мониторинг

5 Работа с модулем среды разработки

- 5.1 Описание интерфейса
- 5.2 Вход в систему
- 5.3 Смена языка интерфейса
- 5.4 Выход из системы
- 5.5 Интерфейс работы с запросами
- 5.6 Выполнение Ad-hoc запросов
- 5.7 Выполнение Ad-hoc запросов с помощью AI
- 5.8 Выполнение поиска
- 5.9 Просмотр состава таблиц
- 5.10 Просмотр и поиск сниппетов
- 5.11 Создание сниппетов
- 5.12 Изменение сниппетов
- 5.13 Удаление сниппетов
- 5.14 Просмотр блокнотов
- 5.15 Создание блокнота
- 5.16 Создание блокнота с помощью AI
- 5.17 Типы блоков в блокноте
- 5.18 Управление блоками
- 5.19 Добавление переменной в блокноте
- 5.20 Дополнительные возможности
- 5.21 Работа с ИИ-агентом
- 5.22 Интерфейс работы с ИИ-агентом
- 5.23 Выполнение запроса через ИИ-агента
- 5.24 Начало нового чата
- 5.25 Работа с историей общения

6 Подключение к системе из внешних инструментов

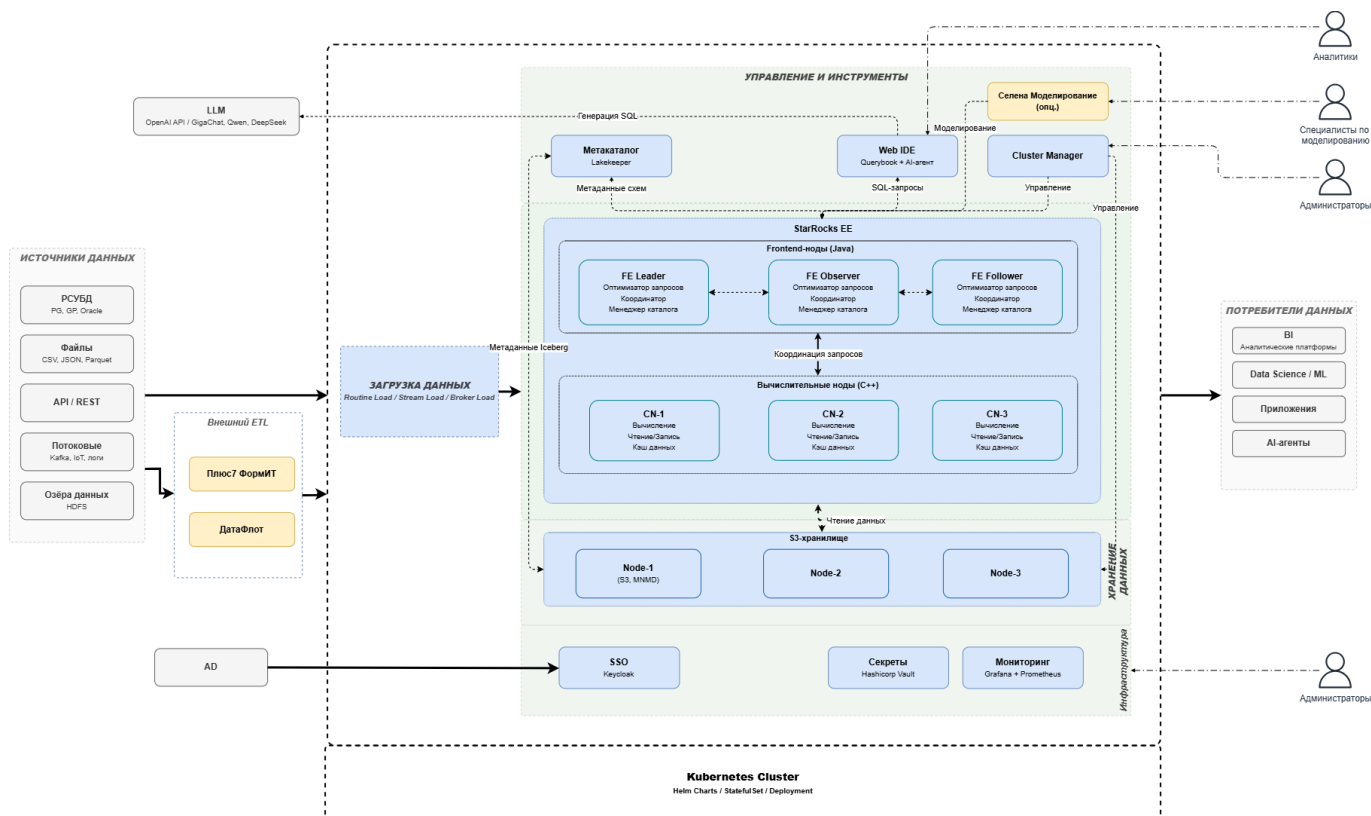
- 6.1 Подключение к платформе данных из Apache Superset
- 6.2 Подключение к платформе данных из Tableau Desktop
- 6.3 Подключение к платформе данных из DataGrip
- 6.4 Подключение к платформе данных из DBeaver
- 6.5 Подключение к платформе данных из Jupyter

1 Архитектура и компоненты

Платформа Данных Селена включает следующие компоненты:

1. Мастер сервер
2. Вычислительный сервер
3. Панель управления кластером (Cluster Manager)
4. Метакаatalog данных
5. Мониторинг
6. Модуль среды разработки
7. Модуль хранения данных

Мастер и вычислительные серверы составляют ядро платформы — вычислительный движок (Engine) на базе StarRocks.



Архитектура и компоненты Платформы Данных Селена

1.1 Мастер сервер

Мастер сервер отвечают за управление метаданными, управление клиентскими соединениями, планирование запросов и диспетчеризацию запросов. Каждый Мастер сервер хранит и поддерживает полную копию метаданных в своей памяти, что гарантирует неизбирательное обслуживание среди Мастер серверов. Мастер сервер может выполнять одну из ролей - лидер, последователь и наблюдатель.

Лидер считывает и записывает метаданные. Лидер обновляет метаданные, а затем синхронизирует изменения метаданных с последователями и наблюдателями. Записи данных считаются успешными только после того, как изменения метаданных синхронизированы с более чем половиной последователей.

Последователи синхронизируют и воспроизводят журналы с лидера для обновления метаданных.

1.2 Вычислительный сервер

Основной задачей Вычислительных серверов является выполнение SQL запросов. Мастер сервер разбирают каждый SQL-запрос на логический план выполнения в соответствии с семантикой запроса, а затем преобразуют логический план в физические планы выполнения, которые могут быть выполнены на Вычислительных серверах. Вычислительные сервера, хранящие данные назначения, выполняют запрос. Это устраняет необходимость в передаче и копировании данных, достигая высокой производительности запроса.

1.3 Панель управления кластером (Cluster Manager)

Cluster Manager - основной инструмент управления кластером, обеспечивает администраторам все необходимые функции по созданию и настройке кластера. Обеспечивает функции добавления и удаления нод, назначение типов и развертывание приложений на новых нодах. Помимо этого, выполняет функции безопасности, обеспечивая контроль доступа как к самим интерфейсам, используемым в платформе компонентов, так и к данным, расположенным в хранилище.

1.4 Модуль загрузки данных

Модуль загрузки данных обеспечивает бесперебойную загрузку данных из различных систем источников данных в целевое хранилище. Компонент обеспечивает возможность

извлекать данные из самых разнообразных систем включая самые популярные реляционные и No-SQL базы данных, шины данных и т.д. Кроме задач по загрузке данных, компонент реализует функции визуального формирования потоков данных, и настройки расписания запусков, а также выполнять функции отладки и мониторинга выполнения процессов работы с данными.

1.5 Метакаталог данных

Метакаталог данных содержит информацию о метаданных, хранящихся в объектном хранилище, и обеспечивает быстрое формирование запросов на данные в объектном хранилище. Каталог метаданных в архитектуре платформы Селена обеспечивает единое представление о хранящихся в объектном хранилище данных.

1.6 Мониторинг

Мониторинг в системе обеспечивается выделенным модулем, который собирает всю необходимую информацию о работе кластера, состоянии основных компонентов и вычислительных мощностях. Пользователям предоставляется целый ряд пред настроенных мониторов отображая информацию от загрузки ЦПУ на нодах мастер серверов, заканчивая производительностью и количеством выполняемых запросов на вычислительных узлах кластера.

1.7 Модуль среды разработки

Модуль среды разработки позволяет различным бизнес-пользователям быстро и просто получить данные. Для разработчиков и аналитиков обеспечивается возможность создания ad-hoc запросов и формирование блокнотов. А для бизнес-пользователей для работы с данным компонентом нет необходимости уметь писать sql-запросы. Все обращения в систему формируются используя семантику, далее они автоматически преобразуются в SQL код, который в свою очередь и выполняется, возвращая ожидаемый результат пользователю.

1.8 Модуль хранения данных

В качестве компонента, отвечающего за непосредственное хранение данных, могут выступать два решения для хранения: объектное хранилище S3 и HDFS. Платформа хранения и обработки данных Селена может поставляться вместе с S3 на базе Minio, но при этом может быть использовано любое клиентское S3 Compatible решение, а также HDFS для реализации хранения данных. Платформа обеспечивает возможность

записывать и хранить данные в различных форматах, в зависимости от использования внешнего метакаталлога данные будут храниться в открытом формате, либо в нативном.

2 Модуль управления кластером (Cluster Manager)

2.1 Вход в систему

1. Откройте URL Cluster Manager'a.
2. Нажмите «войти через Keycloak».
3. Введите логин и пароль.
4. Нажмите «Sign In».

2.2 Установка и обновление лицензии

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите кнопку «Лицензия».
3. Нажмите «Загрузить новую лицензию».
4. В открывшемся окне выберите лицензионный файл.
5. Нажмите «Открыть».

2.3 Добавление виртуальных машин

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите «Виртуальные машины».
3. Нажмите «Добавить VM».
4. Укажите:
 - Имя VM
 - Описание
 - Hostname или IP
 - Порт хоста
 - Имя пользователя.
5. Выберите «Тип ноды»: FE/BE/CN.
6. Укажите (при необходимости) пути для:
 - Корневой каталог данных

- Каталог метаданных FE
- Каталог хранения (указывается по одному на строку).

7. Доступ (Тип аутентификации):

- SSH-пароль — укажите пароль пользователя. Дополнительно задайте Отдельный пароль `sudo` при необходимости.
- SSH-ключ — укажите текст ключа.

8. Нажмите «Создать».

Для проверки корректности подключения нажмите «Проверить SSH».

2.4 Добавление фронтенд (мастер) узлов

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите «Статус и операции».
3. В разделе FE-узлы нажмите «Изменить кол-во нод».
4. Увеличьте количество нод в соответствующем поле.
5. Нажмите «Изменить кол-во нод».

2.5 Добавление бэкенд (вычислительных) узлов

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите «Статус и операции».
3. В разделе CN-узлы или BE-узлы нажмите «Изменить кол-во нод».
4. Увеличьте количество нод в соответствующем поле.
5. Нажмите «Изменить кол-во нод».

2.6 Декомиссия и Удаление узлов

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите «Виртуальные машины».
3. В строке узла нажмите кнопку «Декомиссия».
4. Подтвердите свой выбор.
5. Дождитесь завершения процесса.
6. После выполнения декомиссии станет доступна кнопка «удалить».

7. В строке узла нажмите кнопку «Удалить».
8. Дождитесь завершения процесса.

2.7 Конфигурирование параметров Кластера

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите «Конфигурация».
3. В блоке JSON-спецификация Engine внесите необходимые параметры.
4. Нажмите «Применить конфигурацию».

2.8 Перезапуск вычислительного кластера

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите «Статус и операции».
3. Напротив строки кластера нажмите «Действие».
4. Нажмите «Перезапустить кластер» из выпадающего списка.
5. Подтвердите действие.
6. Дождитесь выполнения операции.

2.9 Удаление вычислительного кластера

1. Войдите в Cluster Manager.
2. В навигационной панели нажмите «Статус и операции».
3. Напротив строки кластера нажмите «Действие».
4. Нажмите «Удалить кластер» из выпадающего списка.
5. Подтвердите действие.
6. Дождитесь выполнения операции.

3 Управление пользователями

3.1 Создание группы

1. Откройте консоль Selene SSO (Keycloak).
2. Обязательно выберите realm — Selena.
3. В навигационной панели нажмите «Groups».
4. Нажмите «Create group».
5. В поле «Name» введите имя группы. Не указывайте в имени символ «/», его Keycloak подставит автоматически.
6. При необходимости укажите описание группы в поле «Description».
7. Нажмите «Create».

При необходимости создайте подгруппы (например, `/cm/admins` — родительская группа + дочерняя). Для этого:

1. Создайте родительскую группу (например `cm`).
2. Выберите созданную группу.
3. Перейдите на вкладку «Child groups».
4. Нажмите «Create group».
5. В поле «Name» введите имя подгруппы (например `admins`). Не указывайте в имени символ «/», его Keycloak подставит автоматически.
6. При необходимости, укажите описание группы в поле «Description».
7. Нажмите «Create».

Полный путь подгруппы из примера станет `/cm/admins`.

3.2 Создание пользователя

1. Откройте консоль Selene SSO (Keycloak).
2. Обязательно выберите realm — Selena.
3. В навигационной панели нажмите «Users».
4. Нажмите «Add user».

5. Укажите:

- Username — идентификатор пользователя
- Email — почтовый адрес пользователя
- First name — имя пользователя
- Last name — фамилия пользователя.

6. Выберите «Email verified», если не требуется подтверждение почтового адреса.

7. Нажмите «Join Groups», если необходимо сразу прикрепить пользователя к группе.

8. Нажмите «Create».

3.3 Указание пароля пользователя

1. Откройте консоль Selene SSO (Keycloak).

2. Обязательно выберите realm — Selena.

3. В навигационной панели нажмите «Users».

4. Выберите пользователя.

5. На странице пользователя откройте вкладку «Credentials».

6. Нажмите «Set password».

7. Введите одинаковый пароль в «Password» и Password «confirmation».

8. Выберите:

- «Temporary On» для смены пароля при первом входе.
- «Temporary Off», если пароль постоянный.

9. Нажмите «Save», и далее «Set password».

3.4 Смена пароля пользователя

1. Откройте консоль Selene SSO (Keycloak).

2. Обязательно выберите realm — Selena.

3. В навигационной панели нажмите «Users».

4. Выберите пользователя.

5. На странице пользователя откройте вкладку «Credentials».

6. Нажмите «Reset password».

7. Введите одинаковый пароль в «Password» и «New password confirmation».
8. Выберите:
 - «Temporary On» для смены пароля при первом входе.
 - «Temporary Off», если пароль постоянный.
9. Нажмите «Save», и далее «Set password».

3.5 Добавление пользователя в группу

1. Откройте консоль Selene SSO (Keycloak).
2. Обязательно выберите realm — Selena.
3. В навигационной панели нажмите «Users».
4. Выберите пользователя.
5. Перейдите на вкладку «Groups».
6. Нажмите «Join Group».
7. Выберите необходимые группы.
8. Нажмите «Join».

3.6 Удаление пользователя

1. Откройте консоль Selene SSO (Keycloak).
2. Обязательно выберите realm — Selena.
3. В навигационной панели нажмите «Users».
4. Выберите пользователя.
5. Нажмите «Delete user».
6. Подтвердите действие нажав «Delete».

3.7 Удаление группы

1. Откройте консоль Selene SSO (Keycloak).
2. Обязательно выберите realm — Selena.
3. В навигационной панели нажмите «Groups».
4. Выберите группу (:).
5. Нажмите напротив неё кнопку меню.

6. Нажмите «Delete».
7. Подтвердите действие нажав «Delete».

3.8 Создание роли

1. Войдите в Cluster Manager.
2. В навигационном меню нажмите «Роли и привилегии».
3. Нажмите «Создать роль».
4. Укажите название.
5. Нажмите «Создать».

3.9 Удаление роли

1. Войдите в Cluster Manager.
2. В навигационном меню нажмите «Роли и привилегии».
3. Выберите роль.
4. Нажмите «Удалить».
5. Подтвердите выбор.

3.10 Предоставление полномочий для роли

1. Войдите в Cluster Manager.
2. В навигационном меню нажмите «Роли и привилегии».
3. Выберите необходимую роль.
4. Перейдите на вкладку «Привилегии».
5. Нажмите «Выдать привилегию».
6. В открывшемся окне, задайте:
 - Тип объекта (например: Каталог, Таблица, База данных, Функция и т.д.)
 - Область (в зависимости от выбранного типа, это может быть конкретная таблица или все таблицы в БД)
 - Привилегии для объекта.
7. Нажмите «Выдать».

3.11 Отзыв полномочий для роли

1. Войдите в Cluster Manager.
2. В навигационном меню нажмите «Роли и привилегии».
3. Выберите необходимую роль.
4. Перейдите на вкладку «Привилегии».
5. Напротив необходимой привилегии нажмите «Отозвать».
6. Подтвердите выбор.

3.12 Назначение роли для группы пользователей

1. Войдите в Cluster Manager.
2. В навигационном меню нажмите «Группы и SQL роли».
3. Напротив необходимой группы нажмите «Управление ролями».
4. В блоке «Выдать роль» выберите необходимую роль из выпадающего списка.
5. Нажмите «Выдать».

3.13 Отзыв роли для группы пользователей

1. Войдите в Cluster Manager.
2. В навигационном меню нажмите «Группы и SQL роли».
3. Напротив необходимой группы нажмите «Управление ролями».
4. В блоке «Текущие роли» напротив необходимой роли нажмите «Отозвать».

4 Работа с модулем Мониторинга

4.1 Мониторинг

Для контроля за состоянием кластера:

1. Войдите в Cluster Manager.
2. В навигационном меню нажмите «Grafana».
3. В открывшемся окне появится Dashboard.
4. Метрики сгруппированы по типам:
 - Overview — общее состояние кластера, ключевые показатели производительности и доступности.
 - Cluster Overview — сводная информация по узлам кластера, их статусам и нагрузке.
 - Query Statistic — статистика выполнения запросов, включая время выполнения, количество запросов и ошибки.
 - Resource Group — использование ресурсов по группам, распределение нагрузки между ними.
 - Jobs — состояние фоновых и плановых заданий, время их выполнения и статусы.
 - Transaction — показатели работы транзакций: количество, длительность, ошибки.
 - FE JVM — метрики работы JVM фронтенд-узлов (память, сборка мусора, потоки).
 - BE — общие показатели работы backend-узлов (нагрузка CPU, состояние служб).
 - BE Memory — использование оперативной памяти backend-узлами.
 - BE Tasks — статистика выполнения задач на backend-узлах.
 - BE Load — показатели загрузки backend-узлов, распределение вычислительных ресурсов.

Раскрывая каждую группу, можно просмотреть соответствующие графики и показатели.

5 Работа с модулем среды разработки

5.1 Описание интерфейса

Интерфейс среды разработки реализован в виде веб-приложения и содержит следующие элементы:

- Левая панель инструментов:
 - Консоль — выполнение Ad-hoc запросов.
 - Поиск — форма поиска по различным объектам.
 - Расписание — список блокнотов, запланированных на запуск по расписанию.
 - ИИ агент — окно работы с искусственным интеллектом.
 - Язык — выбор языка интерфейса.
 - Настройки — конфигурация параметров и выход из системы.
- Правая панель инструментов:
 - Блокноты — создание и управление блокнотами.
 - Таблицы — просмотр объектов, доступных в хранилище данных.
 - Снимпеты — создание и сохранение пользовательских фрагментов кода.
 - Вып. запросы — история выполненных запросов и их результаты.
- Центральная область:
 - Область редактора — ввод и редактирование запросов.
 - Область результатов — отображение результатов выполнения запросов в табличной форме.

5.2 Вход в систему

Процесс аутентификации пользователя в веб-интерфейсе подсистемы:

1. Откройте указанный URL-адрес в браузере.
2. Нажмите кнопку «Login».
3. Введите логин и пароль.
4. Нажмите кнопку «Sign In».

5.3 Смена языка интерфейса

Изменение языка отображения элементов интерфейса:

1. Войдите в подсистему среды разработки.
2. В левой панели инструментов нажмите кнопку «Язык».
3. Выберите Русский или English.

5.4 Выход из системы

Завершение сеанса работы пользователя:

1. Войдите в подсистему среды разработки.
2. В левой панели инструментов нажмите кнопку «Настройки».
3. Нажмите кнопку «Выход».

5.5 Интерфейс работы с запросами

Интерфейс состоит из следующих элементов:

- Область редактирования — ввод SQL-запросов или текстовых инструкций для ИИ.
- Область результатов — отображение данных, полученных в результате выполнения запроса.
- Дополнительные действия  :
 - Форматировать — переформатирование запроса в удобный для чтения вид.
 - Скрыть запрос — удаление текста запроса из области редактирования.
 - Выполнение — запуск выполнения запроса.

В области результатов доступны:

1. Скрыть колонки — фильтрация отображения колонок результирующей таблицы.
2. Full Screen — разворачивание результирующей таблицы на весь экран.
3. Просмотр результатов нескольких запросов — переключение между результатами через выпадающее меню кнопки *Запрос 1 из N*.

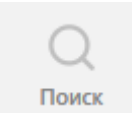
5.6 Выполнение Ad-hoc запросов

1. Войдите в подсистему среды разработки.
2. В левой панели инструментов нажмите кнопку «Консоль».
3. В панели редактора введите необходимый SQL-запрос в поле запроса.
4. Нажмите кнопку «Выполнить».
5. В области результатов отобразится результирующая таблица.

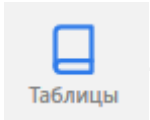
5.7 Выполнение Ad-hoc запросов с помощью AI

1. Войдите в подсистему среды разработки.
2. В левой панели инструментов нажмите кнопку «Консоль».
3. В панели редактора введите текстовый запрос в поле работы с AI.
4. Нажмите Enter.
5. Подтвердите правильность созданного запроса, нажав Разрешить.
6. Нажмите Выполнить.
7. В области результатов отобразится результирующая таблица.

5.8 Выполнение поиска

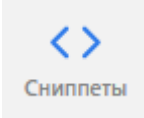
1. Войдите в подсистему среды разработки.
2. В левой панели инструментов нажмите кнопку «Поиск» .
3. В открывшемся окне выберите тип объекта: запрос, блокнот, таблицы, список.
4. Введите интересующую информацию в поле поиска.
5. При необходимости используйте фильтры в правой части окна для уточнения поиска.

5.9 Просмотр состава таблиц

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Таблицы» .
3. В открывшейся оснастке разверните каталог.

4. Разверните необходимую базу данных.
5. Нажмите на нужную таблицу.
6. Нажмите «Посмотреть таблицу».
7. В открывшейся оснастке переключайтесь по вкладкам «Обзор», «Колонки», «Примеры строк», «Исходный запрос», «Предупреждения», — для просмотра информации.
8. Для работы с таблицей доступны действия:
 - Добавить в избранное — нажмите кнопку «Добавить в избранное».
 - Добавить теги — нажмите .
 - Добавить себя как владельца — нажмите  напротив поля «Владельцы».
 - Редактировать описание — нажмите «Редактировать», введите текст и нажмите «Сохранить».
 - Добавить в список — нажмите  и выберите нужный.

5.10 Просмотр и поиск сниппетов

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Сниппеты» .
3. В открывшейся оснастке выберите:
 - Публичный — доступен всем пользователям.
 - Приватный — доступен только вам.
4. Для поиска введите имя сниппета в поле поиска.

5.11 Создание сниппетов

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Сниппеты».
3. В открывшейся оснастке нажмите кнопку .
4. Введите название, описание, запрос, выберите область видимости (Приватный или Публичный).

5. Нажмите «Сохранить».

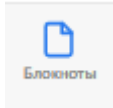
5.12 Изменение сниппетов

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Сниппеты».
3. В открывшейся оснастке выберите сниппет.
4. Внесите изменения в запрос или описание.
5. Нажмите «Обновить».

5.13 Удаление сниппетов

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Сниппеты».
3. В открывшейся оснастке выберите сниппет.
4. Нажмите «Удалить».

5.14 Просмотр блокнотов

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Блокноты»  .
3. В открывшейся оснастке разверните группированные списки:
 - Недавние — последние открытые блокноты.
 - Избранное — блокноты, добавленные в избранное.
 - Списки — пользовательские или общие списки блокнотов.
 - Мои блокноты — блокноты, созданные пользователем.
4. Выберите необходимый блокнот.

5.15 Создание блокнота

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Блокноты».
3. В открывшейся оснастке нажмите кнопку  .

4. Выберите команду «Создать блокнот».
5. Введите название блокнота.

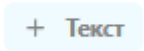
5.16 Создание блокнота с помощью AI

1. Войдите в подсистему среды разработки.
2. В правой панели инструментов нажмите кнопку «Блокноты».
3. В открывшейся оснастке нажмите кнопку  .
4. Выберите команду «Создать блокнот с помощью AI».
5. Выберите каталог.
6. Выберите базу данных.
7. Введите текстовый запрос.
8. Нажмите «Создать».

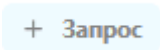
5.17 Типы блоков в блокноте

В блокнотах поддерживается использование трёх типов блоков.

Текстовый блок

1. Нажмите кнопку «Текст»  .
2. Введите текстовое описание.

Блок SQL-запроса

1. Нажмите кнопку «Запрос»  .
2. Введите название запроса.
3. В блок редактирования введите SQL-запрос.

Использование AI при работе с запросами:

4. В поле работы с AI введите текстовый запрос.
5. Нажмите Enter.
6. Проверьте корректность сгенерированного SQL-запроса.

7. Нажмите «Разрешить».

8. При необходимости измените текстовый запрос и повторно подтвердите.

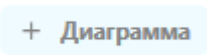
Выполнение запроса:

9. Нажмите кнопку «Выполнить»  .

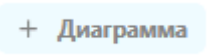
Дополнительные действия через меню запроса:

- Форматировать запрос — привести к читаемому виду.
- Объяснить запрос — получить описание запроса с помощью AI.
- Скрыть запрос — оставить только результаты.
- Отобразить шаблон — применить шаблон.

Блок диаграммы

1. Нажмите кнопку «Диаграмма»  .

2. Если диаграмма создаётся сразу после запроса, он будет выбран источником данных по умолчанию.

3. Для изменения источника нажмите «Настройки диаграммы» ( → вкладка «Данные»).

4. Выберите:

- Текущий запрос в блокноте.
- Запрос в блокноте (с указанием номера).
- Из выполненного запроса (с указанием названия и номера).

5. Настройте отображение во вкладках «Диаграмма» и «Визуализация».

6. Нажмите «Сохранить».

5.18 Управление блоками

1. Удалить блок — нажмите кнопку «Удалить»  .

2. Комментарий — нажмите кнопку «Комментарий»  , введите текст и нажмите «Comment».

5.19 Добавление переменной в блокноте

1. Нажмите + Новая переменная.
2. Укажите имя, тип и значение.
3. Сохраните изменения.

5.20 Дополнительные возможности

1. Запустить все  Запустить все — выполнение всех блоков в блокноте.
2. Добавить в избранное — значок звёздочки.
3. Доступ — выбрать Приватный или Публичный режим, при приватном можно назначить других пользователей.
4. Меню дополнительных действий  :
 - Клонировать — копия блокнота.
 - Экспортировать в PDF — сохранить в файл PDF.
 - Списки — просмотр списка, в котором используется блокнот.
 - Шаблон — сохранить или применить шаблон.
 - Планировщик — задать расписание выполнения.
 - Удалить — удалить блокнот.

5.21 Работа с ИИ-агентом

ИИ-агент предоставляет возможность интерактивного взаимодействия с моделью LLM через интеграцию по протоколу MCP. Подсистема позволяет формировать запросы на естественном языке, получать автоматически сгенерированные SQL-запросы, выполнять их и просматривать результаты.

5.22 Интерфейс работы с ИИ-агентом

Экран работы с ИИ-агентом состоит из следующих элементов:

- Область ввода запроса — текстовое поле для ввода вопроса или задачи.
- Кнопка «Выполнить» — отправка запроса ИИ-агенту.
- Левая область — чат с ИИ-агентом (история запросов и ответов).

- Правая область — результаты выполнения запросов (таблицы, графики, текстовые описания).
- Кнопка «История» — просмотр предыдущих диалогов с ИИ-агентом.
- Кнопка «Новый чат» — очистка текущей истории и начало нового диалога.

5.23 Выполнение запроса через ИИ-агента

1. Войдите в подсистему среды разработки.
2. В левой панели инструментов нажмите кнопку «ИИ-агент»  .
3. В приветственном окне в области ввода запроса введите вопрос или задачу на естественном языке.
4. Нажмите кнопку «Выполнить».
5. Дождитесь появления ответа в левой области чата.
6. Если ИИ-агент сгенерировал SQL-запрос:
 - Проверьте корректность предложенного запроса.
 - При необходимости, внесите изменения.
 - Подтвердите выполнение.
7. Результаты выполнения запроса будут отображены в правой области экрана.

5.24 Начало нового чата

1. В интерфейсе ИИ-агента нажмите кнопку «Новый чат»  .
2. Текущая история общения будет очищена.
3. Введите новый запрос в область ввода и нажмите «Выполнить».

5.25 Работа с историей общения

1. В интерфейсе ИИ-агента нажмите кнопку «История»  .
2. В открывшемся списке выберите интересующий диалог.
3. Просмотрите предыдущие запросы и ответы.

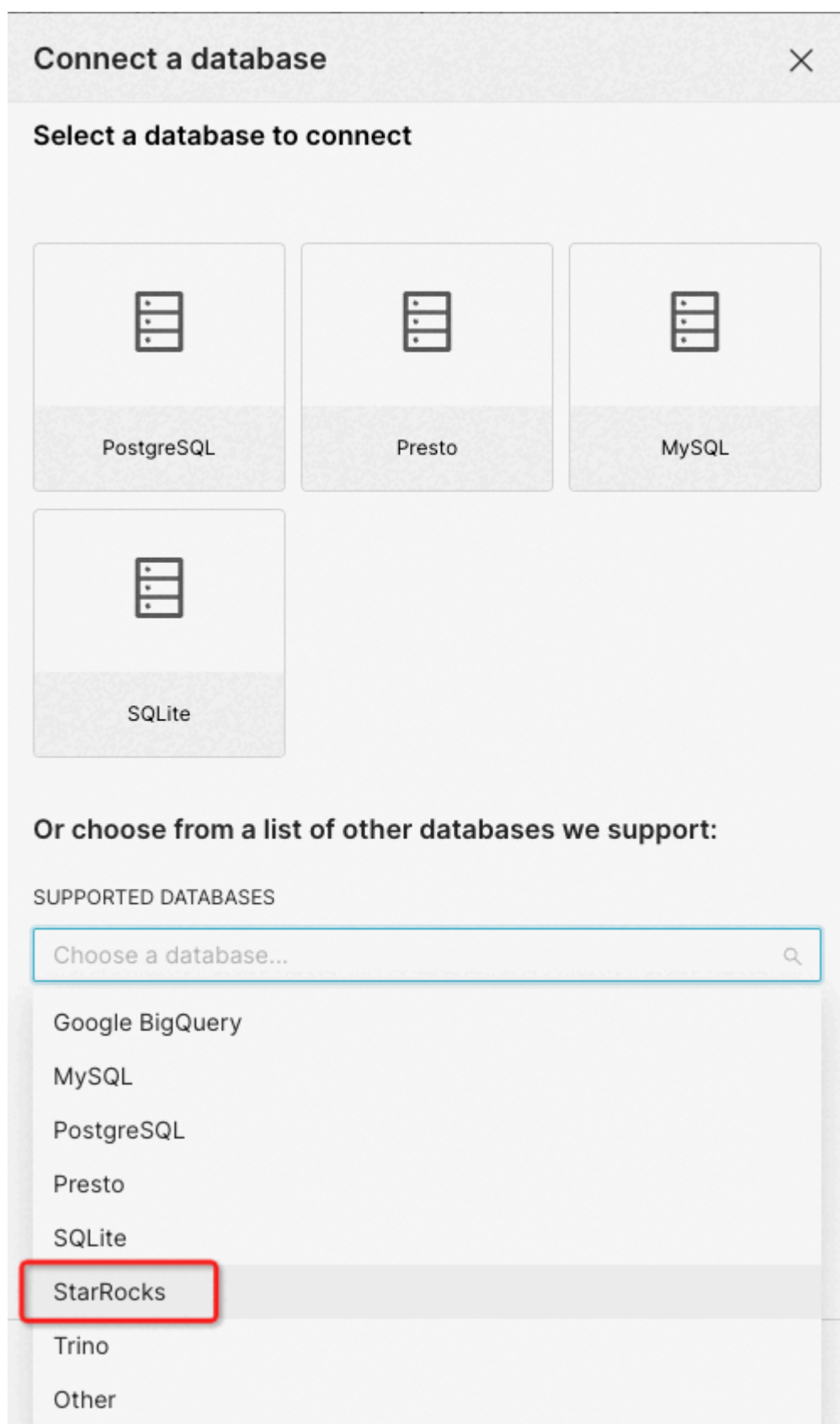
6 Подключение к системе из внешних инструментов

6.1 Подключение к платформе данных из Apache Superset

1. На сервере Apache Superset установите Python client StarRocks, выполнив:

```
pip install starrocks
```

2. В Apache Superset перейдите в раздел подключения БД в качестве источника данных.
3. В разделе поддерживаемых Баз Данных (supported databases) из выпадающего списка выберите StarRocks.



Apache Superset — выбор БД

4. На Этапе настройки подключения укажите:
- Отображаемое имя в поле «Display Name».
 - Путь подключения в поле «SQLAlchemy URI»:

```
starrocks://<User>:<Password>@<Host>:<Port>/<Catalog>.<Database>
```

- User — имя пользователя технической УЗ.
- Password — пароль от УЗ.
- Host — hostname или ip адрес мастер сервера.
- Port — порт мастер сервера.
- Catalog — имя целевого каталога, поддерживаются внутренние и внешние каталоги.
- Database — имя целевой База Данных, поддерживаются внутренние и внешние базы данных.

Connect a database

×

STEP 2 OF 2

Enter Primary Credentials

Need help? Learn how to connect your database [here](#).

BASIC

ADVANCED

DISPLAY NAME *

sr_hive_catalog

Pick a name to help you identify this database.

SQLALCHEMY URI *

starrocks://root:xxx@192.168.7.221:9030/hive_catalog.hive_test

Refer to the for more information on how to structure your URI.

TEST CONNECTION

i

Additional fields may be required
Select databases require additional fields to be completed in the Advanced tab to successfully connect the database.
Learn what requirements your databases has [here](#).

BACK

CONNECT

Apache Superset — настройки подключения

6.2 Подключение к платформе данных из Tableau Desktop

Tableau Desktop поддерживает запросы и визуализацию как внутренних, так и внешних данных в Селене.

1. Создайте базу данных в Tableau.
2. Выберите Other Databases (JDBC) в качестве источника данных.
3. Для Dialect выберите MySQL.
4. Для URL введите URL в формате MySQL URI, как показано ниже:

```
jdbc:mysql://<Host>:<Port>/<Catalog>.<Databases>
```

Параметры в URL описаны следующим образом:

- Host: hostname или IP-адрес хоста мастер сервера вашего кластера.
- Port: порт хоста мастер сервера, например, 9030.
- Catalog: целевой каталог в вашем кластере. Поддерживаются как внутренние, так и внешние каталоги.
- Database: целевая база данных в вашем кластере. Поддерживаются как внутренние, так и внешние базы данных.

5. Настройте имя пользователя и пароль.

6.3 Подключение к платформе данных из DataGrip

DataGrip поддерживает запросы как внутренних, так и внешних данных.

Создайте источник данных в DataGrip. Обратите внимание, что в качестве источника данных необходимо выбрать MySQL.

Параметры, которые вам необходимо настроить, описаны ниже:

- Хост: hostname или IP-адрес хоста мастер сервера вашего кластера.
- Порт: порт хоста мастер сервера, например, 9030.
- Аутентификация: метод аутентификации, который вы хотите использовать. Выберите Имя пользователя и пароль.

- Пользователь: имя пользователя, которое используется для входа в ваш кластер, например, admin.
- Пароль: пароль, который используется для входа в ваш кластер.
- База данных: источник данных, к которому вы хотите получить доступ в вашем кластере. Значение этого параметра имеет формат `<catalog_name>`.
`<database_name>`.
- catalog_name: имя целевого каталога в вашем кластере. Поддерживаются как внутренние, так и внешние каталоги.
- database_name: имя целевой базы данных в вашем кластере. Поддерживаются как внутренние, так и внешние базы данных.

6.4 Подключение к платформе данных из DBeaver

DBeaver — это клиентское программное обеспечение SQL и инструмент администрирования баз данных.

Чтобы подключиться к базе данных, выполните следующие действия:

1. Запустите DBeaver.
2. Щелкните значок плюса (+) в верхнем левом углу окна DBeaver или выберите «База данных» → «Новое подключение к базе данных» в строке меню, чтобы получить доступ к помощнику.
3. Выберите драйвер MySQL.
4. На этапе выбора базы данных вам будет представлен список доступных драйверов. Нажмите «Analytical» на левой панели, чтобы быстро найти драйвер MySQL. Затем дважды щелкните значок «MySQL».
5. Настройте подключение к базе данных.
6. На этапе «Параметры подключения» перейдите на вкладку «Главное» и настройте следующие основные параметры подключения:
 - Хост сервера: hostname или IP-адрес хоста мастер сервера вашего кластера.
 - Порт: порт хоста мастер сервера, например, 9030.
 - База данных: целевая база данных в вашем кластере. Поддерживаются как внутренние, так и внешние БД.

- Имя пользователя: имя пользователя, которое используется для входа в ваш кластер, например, admin.
 - Пароль: пароль, который используется для входа в ваш кластер.
7. Вы также можете просматривать и редактировать свойства драйвера MySQL на вкладке «Свойства драйвера», если это необходимо. Чтобы изменить определенное свойство, щелкните строку в столбце «Значение» для этого свойства.
8. Протестируйте подключение к базе данных.
9. Нажмите «Проверить подключение», чтобы проверить точность параметров подключения. Появится диалоговое окно с информацией о драйвере MySQL. Нажмите «ОК» в диалоговом окне, чтобы подтвердить информацию. После успешной настройки параметров подключения нажмите «Готово», чтобы завершить процесс.
10. После установки подключения вы можете просмотреть его в дереве подключений к базе данных слева, и DBeaver сможет эффективно подключиться к базе данных.

6.5 Подключение к платформе данных из Jupyter

Вы можете использовать JupySQL поверх Jupyter для выполнения запросов к Selena.

После загрузки данных в кластер вы можете запрашивать и визуализировать их с помощью построения графиков SQL.

Перед началом работы вам необходимо установить локально следующее программное обеспечение:

- JupySQL: `pip install jupysql`
- Jupyterlab: `pip install jupyterlab`
- SKlearn Evaluation: `pip install sklearn-evaluation`
- Python
- pymysql: `pip install pymysql`

После выполнения вышеуказанных требований вы можете открыть Jupyter lab, просто вызвав `jupyterlab` — это откроет интерфейс блокнота. Если Jupyter lab уже запущен в блокноте, вы можете просто запустить ячейку ниже, чтобы получить зависимости.

```
%pip install --quiet jupysql sklearn-evaluation pymysql
```

Для использования обновленных пакетов может потребоваться перезапуск ядра.

```
import pandas as pd
from sklearn_evaluation import plot
# Import JupyterSQL Jupyter extension to create SQL cells.
%load_ext sql
%config SqlMagic.autocommit=False
```

Вам нужно будет настроить строку подключения в соответствии с типом экземпляра, к которому вы пытаетесь подключиться (url, пользователь и пароль). В примере ниже используется локальный экземпляр.

Подключение к Selena через JupyterSQL

В этом примере используется экземпляр Docker, который отражает данные в строке подключения.

Пользователь `root` используется для подключения к локальному экземпляру Selena, создание базы данных и проверки того, что данные действительно могут быть прочитаны и записаны в таблицу.

```
%sql mysql+pymysql://root:@localhost:9030
```

Создать JupyterSQL БД:

```
%sql CREATE DATABASE jupysql;

%sql USE jupysql;
```

Создать таблицу:

```
%%sql
```

```
CREATE TABLE tbl(c1 int, c2 int) distributed by hash(c1) properties  
(\\"replication_num\\" = \\"1\\");
```

```
INSERT INTO tbl VALUES (1, 10), (2, 20), (3, 30);
```

```
SELECT * FROM tbl;
```

Сохранение и загрузка запросов

Теперь, после создания базы данных, вы можете записать в неё некоторые образцы данных и запросить их.

JupySQL позволяет разбивать запросы на несколько ячеек, упрощая процесс создания больших запросов.

Вы можете писать сложные запросы, сохранять их и выполнять при необходимости, аналогично CTE в SQL.

```
%%sql --save initialize-table --no-execute
```

```
CREATE TABLE tbl(c1 int, c2 int) distributed by hash(c1) properties  
("replication_num" = "1");
```

```
INSERT INTO tbl VALUES (1, 1), (2, 2), (3, 3);
```

```
SELECT * FROM tbl;
```

Обратить внимание, что используется `--with;`, это извлечет ранее сохраненные запросы и добавит их (используя CTE). Затем мы сохраняем запрос в `track_fav`.