**Документация, содержащая описание функциональных характеристик модуля программного обеспечения**

**Модуль «Сегментатор данных»**

**г. Москва 2025**

**Содержание**

[1. Введение 3](#_heading=h.65bfswssbp75)

[2. Назначение модуля 3](#_heading=h.3sut7fe6effz)

[3. Функциональные возможности 3](#_heading=h.16mt3crutyvo)

[4. Архитектура и принцип работы 4](#_heading=h.4ls06p6ko6i3)

[5. Роли и полномочия 7](#_heading=h.sfy1cgcbm2en)

[6. Используемые языки программирования и фреймворки 7](#_heading=h.c68imox5zlzs)

# **1. Введение**

Модуль «Сегментатор данных» предназначен для автоматизированного сбора, обработки и анализа данных пользователей, с возможностью их сегментации по различным критериям. Он предоставляет интерфейс для гибкой настройки сегментов и их выгрузки в различных форматах.

# **2. Назначение модуля**

Модуль разработан для использования в компаниях, работающих с большими объемами данных, включая:

* Электронную коммерцию (анализ эффективности маркетинговых кампаний);
* Рекламные агентства (оптимизация рекламных кампаний);
* Разработчиков мобильных приложений (анализ поведения пользователей);
* Компании-разработчики ПО (оптимизация работы сервисов);
* Онлайн-сервисы (повышение удовлетворенности клиентов).

Аутентификация проходит через сервисы Яндекс ID; Авторизация на основе ролевой модели ( две роли: Администратор, Пользователь); Управление сегментами (создание, редактирование, удаление). Выгрузка данных сегментов в форматах CSV и XLSX. Интеграция с различными источниками данных (PostgreSQL, ClickHouse). Автоматизированное формирование аналитических отчетов.

# **3. Функциональные возможности**

Модуль «Сегментатор данных» выполняет следующие функции:

* Авторизация на основе ролевой модели (Администратор, Пользователь).

Управление сегментами, включая:

* Создание новых сегментов.
* Редактирование существующих сегментов.
* Удаление сегментов.
* Переименование и копирование сегментов.

Настройка сегментов:

* Настройка фильтров по атрибутам.
* Группировка фильтров.
* Выгрузка сегментов в различных форматах (CSV, XLSX).
* Просмотр и анализ ранее сформированных сегментов.
* Интеграция с источниками данных, такими как PostgreSQL и ClickHouse.
* Автоматизированное формирование аналитических отчетов.
* Поиск и фильтрация сегментов по заданным критериям.

Работа с пользователями (для Администратора):

* Поиск пользователей.
* Создание и удаление пользователей.
* Назначение ролей и управление доступами.

#

# **4. Архитектура и принцип работы**

Принцип работы

1. Пользователь загружает данные из различных источников.
2. Происходит обработка данных и разбиение их на сегменты по заданным критериям.
3. Пользователь получает сформированные сегменты и может выгрузить их в различных форматах.
4. Сегменты доступны для анализа и интеграции с другими сервисами.

Структура данных

Модель данных PostgreSQL:

| Поле | Тип | Описание |
| --- | --- | --- |
| id | integer | Уникальный идентификатор |
| name | varchar(50) | Имя сегмента |
| rule | jsonb | Правило выгрузки |
| updated\_at | timestamp(6) | Дата обновления |

Модель данных ClickHouse:

| Поле | Тип | Описание |
| --- | --- | --- |
| user\_id | String | Идентификатор пользователя |
| date | Date | Дата события |
| platform\_main | String | Платформа пользователя |
| cohort | Date | Когорта |

Модуль состоит из следующих компонентов:

Клиентская часть (Web, Mobile, другие устройства);

Сервер визуализации и обработки метрик, включающий:

* + Frontend app (отображение данных);
	+ Backend app (обработка данных);
	+ Базы данных PostgreSQL и Redis;

Сервер сбора метрик, включающий:

* + S3 БД (хранение данных);
	+ Главная БД ClickHouse (агрегация данных);
	+ Модуль сбора данных;
	+ Модуль обработки данных;
	+ Dagster оркестратор (управление потоками данных);

Сервисы сбора аналитических данных (интеграция с внешними источниками).



*Рис. 1 Общая схема сервиса*

Потоки данных, включая протоколы обмена и сценарии

Ниже представлены основные сценарии в виде диаграмм последовательности.



*Рис. 2 Диаграмма последовательности: Создание пользователя/изменение.*

**

*Рис. 3 Диаграмма последовательности: Выбор сегмента/создание/изменение.*

**

*Рис. 4 Диаграмма последовательности: Выгрузка сегментов.*

# **5. Роли и полномочия**

Администратор

* Управление пользователями (создание, удаление, назначение ролей);
* Контроль доступа к данным и функциям системы;
* Настройка параметров выгрузки и отчетности.

Пользователь

* Просмотр и управление сегментами данных;
* Выгрузка данных;
* Анализ сегментов через веб-интерфейс.

# **6. Используемые языки программирования и фреймворки**

База данных: PostgreSQL, ClickHouse

Язык программирования: Python

Фреймворки: Pandas, FastAPI, GraphQL

Инфраструктура: Yandex Cloud, Docker, Kubernetes