**Документация, содержащая информацию, необходимую для эксплуатации экземпляра программного обеспечения**

**Модуль «Модерация контента»**

**г. Москва 2025**

[1. Введение](#_heading=h.p8b57rqzzzfa) 3

[2. Архитектура системы](#_heading=h.l9r5tvv8enzd) 3

[3. Процесс установки модуля](#_heading=h.t57onvrs8us0) 3

[4. Запуск системы](#_heading=h.3hkpy1w24u56) 4

[5. Функциональность](#_heading=h.kjsvihqutcqd) 4

[6. Обновление и поддержка](#_heading=h.e54fyd3ay0r) 4

##

## **1. Введение**

Данный документ содержит инструкции по эксплуатации, настройке и обновлению модуля «Модерация контента». Модуль предназначен для автоматизированной разметки видеоконтента (фильмы, сериалы) в онлайн-кинотеатре для модерации и обеспечения безопасной и здоровой онлайн-среды для пользователей.

# **2. Архитектура системы**

Модуль «Модерация контента» включает в себя следующие компоненты:

* Серверная часть: Обрабатывает запросы, анализирует контент и формирует отчеты.
* База данных ClickHouse: Хранит информацию о модерации и метаданные контента.
* Интерфейс командной строки: Позволяет запускать модули анализа и формировать отчеты.
* Подключение к внешним сервисам (Кинопоиск, IMDB): Получает данные о видеоконтенте.
* Система логирования: Записывает процесс работы и фиксирует возможные ошибки.

# **3. Процесс установки модуля**

Требования к окружению

Перед установкой необходимо убедиться, что в системе доступны следующие компоненты:

* Операционная система: Ubuntu (рекомендуемая версия 20.04 и выше).
* Python: Версия 3.10 или выше.
* Библиотеки Python: Pandas, Selenium, Requests, OpenPyXL, Loguru и другие (полный список приведен в разделе 5).
* ClickHouse: Должен быть развернут и доступен для записи данных.
* Доступ к интернету: Необходим для взаимодействия с API Кинопоиск и IMDB.

Установка модуля

1. Клонировать репозиторий с исходным кодом:
git clone https://mlgit.cyrm.ru/ml-team/mediasurf.git
2. Перейти в папку с проектом:
cd mediasurf
3. Установить зависимости:
pip install -r requirements.txt
4. Проверить корректность установки:
python app.py --help
Если команда вывела справочную информацию, установка прошла успешно.

# **4. Запуск системы**

Запуск модуля осуществляется через командную строку.

* Подготовка входных данных

Перед запуском необходимо подготовить входной файл XLSX с перечнем контента.

* Запуск модуля
1. Указать путь к входному файлу в параметрах запуска:
python app.py --input data.xlsx --output results.xlsx
2. Дождаться завершения обработки.
3. Проверить выходной файл results.xlsx с разметкой контента.

# **5. Функциональность**

Модуль «Модерация контента» выполняет следующие задачи:

* Автоматическая классификация видеоконтента на основе предустановленных критериев (наличие сцен курения, алкоголя, наркотиков, ЛГБТ).
* Получение данных о контенте из внешних сервисов (Кинопоиск, IMDB).
* Формирование отчетов с цветовой разметкой (зеленый – безопасно, желтый – возможно, красный – нарушение).
* Ведение логов работы модуля.
* Возможность интеграции с BI-системами для визуализации данных.

# **6. Обновление и поддержка**

Обновление модуля выполняется стандартными командами:

git pull origin main

pip install -r requirements.txt --upgrade

В случае обнаружения ошибок или необходимости доработки функционала, следует обратиться в службу технической поддержки.