Общество с ограниченной ответственностью
«Интеллектуальные системы разминирования»

Программа для ЭВМ «Dataset Studio»

ОПИСАНИЕ ФУНКЦИОНАЛЬНЫХ ХАРАКТЕРИСТИК, УСТАНОВКА И ЭКСПЛУАТАЦИЯ

Версия 1.0.0

Операционная система: РЕД ОС 8 (x86_64)

2026

1. Общие сведения

Настоящий документ содержит описание функциональных характеристик программы для ЭВМ «Dataset Studio» версии 1.0.0 (далее — ПО), а также сведения, необходимые для её установки и эксплуатации. ПО представляет собой серверное веб-приложение для подготовки размеченных наборов изображений и функционирует под управлением операционной системы РЕД ОС 8.

Параметр	Значение
Полное наименование	Программа для ЭВМ «Dataset Studio»
Версия	1.0.0 (пакет `dataset-studio-1.0.0-4.red80.x86_64.rpm`)
Правообладатель	ООО «Интеллектуальные системы разминирования» (ООО «ИСР»), ИНН 7733480311, ОГРН 1267700013570, адрес: 125222, г. Москва, вн. тер. г. муниципальный округ Митино, ул. Генерала Белобородова, д. 18, помещ. 2/1
Тип ПО	серверное веб-приложение (клиент-сервер, доступ через браузер)
Целевая платформа	РЕД ОС 8 (x86_64), стандартная редакция
Совместимость	Сертификат технологической совместимости программных продуктов «Dataset Studio 1» и «РЕД ОС 8» от 14.10.2026, ООО «ИСР» и ООО «РЕД СОФТ»; результаты испытаний зафиксированы в протоколе тестирования
Лицензия	проприетарная, лицензионное соглашение ООО «ИСР»
Язык интерфейса и документации	русский

ПО разработано работниками ООО «ИСР». Сторонние компоненты, входящие в состав ПО, распространяются по свободным лицензиям и перечислены в разделе 4.

2. Назначение и область применения

ПО предназначено для подготовки обучающих наборов данных для нейронных сетей компьютерного зрения и обеспечивает загрузку изображений, их разметку, контроль качества разметки и выгрузку набора данных в форматах COCO, YOLO и DOTA. Изображения и результаты разметки хранятся на сервере организации-пользователя; внешние облачные сервисы для работы ПО не требуются.

ПО может применяться для:

- подготовки данных для систем обнаружения и сегментации объектов;
- разметки аэрофотоснимков и снимков, полученных с беспилотных воздушных судов, в том числе с использованием повёрнутых рамок;
- организации совместной работы разметчиков с проверкой качества и регистрацией действий пользователей;
- работы в изолированных сетях без доступа к сети Интернет.

Один экземпляр ПО рассчитан на обработку наборов объёмом до нескольких десятков тысяч изображений при одновременной работе нескольких пользователей.

3. Функциональные характеристики

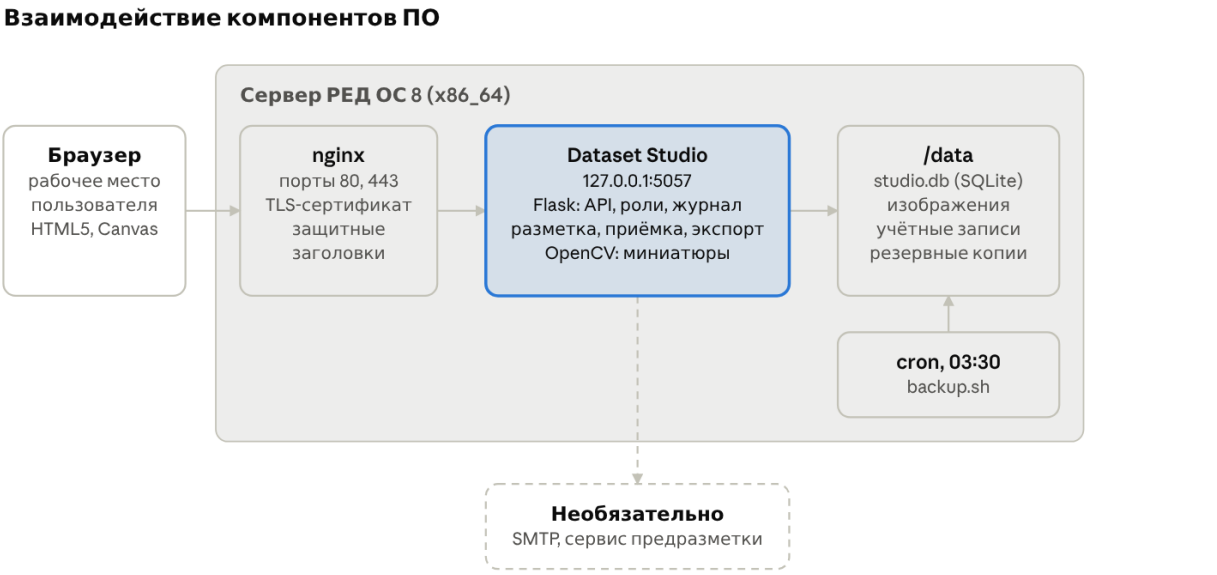
ПО обеспечивает выполнение всех этапов подготовки набора данных: загрузку изображений, разметку, приёмку результатов, анализ распределения классов и экспорт. Перечень функций приведён ниже; обозначения функций соответствуют программе испытаний на совместимость с РЕД ОС 8.

Код	Функция	Описание
F-01	Пользователи и роли	вход по адресу электронной почты и паролю, срок действия сеанса 30 дней; роли «Разметчик», «Аналитик», «Администратор»; приглашение и восстановление пароля по ссылке
F-02	Проекты и загрузка	создание проектов с собственным перечнем классов; загрузка изображений в форматах JPEG и PNG; исключение дубликатов по контрольной сумме SHA-256; разбиение загрузки на партии до 500 изображений
F-03	Разметка	разметка прямоугольными рамками, полигонами и повёрнутыми рамками (OBB); масштабирование и перемещение изображения; автоматическое сохранение; отбраковка непригодных изображений
F-04	Приёмка	проверка разметки с возможностью принять изображение, вернуть его на доработку с комментарием или отбраковать; восстановление отбракованных изображений
F-05	Импорт COCO	импорт ранее размеченного набора данных, в том числе в качестве предварительной разметки, требующей проверки
F-06	Баланс классов	распределение объектов по классам и размерам (малые, средние, крупные); предупреждение о несбалансированности классов
F-07	Экспорт	COCO JSON, YOLO (включая YOLO-seg), DOTA; ограничение числа объектов по классам и размерам с предварительным расчётом состава; формирование архива в фоновом режиме
F-08	Предразметка	автоматическая расстановка рамок внешним сервисом распознавания организации-пользователя; такие рамки включаются в набор данных только после подтверждения пользователем (дополнительная функция)
F-09	Аналитика	тепловая карта расположения объектов на кадре; отчёт по разметчикам за период
F-10	Журнал и откат	регистрация действий пользователей с указанием исполнителя и времени, отбор записей по фильтрам; отмена отдельного изменения разметки

Кроме того, ПО позволяет вести справочник классов (в том числе переименовывать и объединять классы в размеченных данных), распределять партии изображений между разметчиками и выполнять резервное копирование данных (раздел 9).

4. Архитектура и состав ПО

ПО поставляется в виде RPM-пакета, содержащего исполняемый файл серверной части со всеми необходимыми библиотеками. Исходные тексты в комплект поставки не входят. Пользователи работают с ПО через веб-браузер. Запросы принимает веб-сервер nginx и передаёт их серверной части, которая принимает соединения только по локальному адресу 127.0.0.1 (порт 5057). Схема взаимодействия компонентов приведена на рисунке 1.



Взаимодействие компонентов ПО
Сервер РЕД ОС 8 (x86_64)
Браузер
рабочее место пользователя
HTML5, Canvas
nginx
порты 80, 443
TLS-сертификат
защитные заголовки
Dataset Studio
127.0.0.1:5057
Flask: API, роли, журнал
разметка, приёмка, экспорт
OpenCV: миниатюры
/data
studio.db (SQLite)
изображения
учётные записи
резервные копии
cron, 03:30
backup.sh
Необязательно
SMTP, сервис предразметки

Рисунок 1 — Схема развёртывания ПО

Внешние соединения, показанные на схеме пунктиром, используются только после их настройки администратором.

Размещение компонентов ПО на сервере приведено в таблице ниже.

Компонент	Расположение
Исполняемый файл и скрипты резервного копирования	`/opt/dataset-studio/bin/`
Данные: изображения, база `studio.db`, учётные записи, резервные копии	`/data/` (владелец — служебный пользователь `datasetstudio`)
Параметры службы	`/etc/dataset-studio/env`
Служба systemd	`dataset-studio.service`
Конфигурация nginx	`/etc/nginx/conf.d/dataset-`

Компонент	Расположение
	`studio.conf`
Расписание резервного копирования	`/etc/cron.d/ds-backup` (ежедневно в 03:30)

В состав ПО входят следующие сторонние компоненты, распространяемые по свободным лицензиям, не предусматривающим лицензионных отчислений.

Компонент	Роль	Лицензия
Python (CPython)	среда исполнения, встроена в исполняемый файл	PSF
Flask 3.1, Werkzeug, Jinja2, MarkupSafe, itsdangerous, click, blinker	веб-фреймворк, HTTP API	BSD-3-Clause, MIT
waitress	встроенный HTTP-сервер приложения	ZPL-2.1
OpenCV (opencv-python-headless 4.12), NumPy 2.2	обработка изображений, миниатюры	Apache-2.0, BSD-3-Clause
SQLite	хранение состояния проектов и разметки (режим WAL)	Public domain
nginx, rsync, cronie, systemd	из штатного репозитория РЕД ОС 8	BSD-2-Clause, GPL, LGPL

Пользовательский интерфейс разработан на HTML5, CSS3 и JavaScript без использования сторонних фреймворков. Все файлы интерфейса загружаются с сервера ПО; обращения к внешним сетям доставки контента, сервисам статистики и телеметрии не выполняются.

5. Системные требования

Для функционирования ПО требуется сервер под управлением РЕД ОС 8 и рабочие места пользователей с веб-браузером. Подключение к сети Интернет не требуется.

Параметр	Минимум	Рекомендуется
Операционная система	РЕД ОС 8 x86_64, стандартная редакция	РЕД ОС 8 x86_64 с установленными обновлениями
Процессор	2 ядра	4 ядра
Оперативная память	4 ГБ	8 ГБ
Диск для данных (`/data`)	в соответствии с объёмом изображений и резервных копий	отдельный том ext4 или xfs
Пакеты ОС	nginx, rsync, cronie, policycoreutils-python-utils, python3, curl	устанавливаются автоматически из штатного репозитория РЕД ОС
Сеть	входящий TCP 80	TCP 443 с TLS-сертификатом
Рабочее место	Яндекс Браузер, Google Chrome или Microsoft Edge актуальной версии	разрешение экрана не менее 1920×1080, манипулятор «мышь» с колесом прокрутки

6. Установка

Рекомендуемый способ установки — из RPM-пакета. При установке пакета автоматически устанавливаются зависимости, создаётся системная служба и выполняется запуск ПО. Приведённые ниже команды выполняются пользователем с правами администратора.

6.1. Установка RPM-пакета

1. Скопируйте на сервер файл пакета и файл контрольной суммы `.sha256`.

2. Проверьте целостность: `sha256sum -c dataset-studio-1.0.0-4.red80.x86_64.rpm.sha256`. Результат проверки должен быть «OK».
3. Установите пакет: `sudo dnf install ./dataset-studio-1.0.0-4.red80.x86_64.rpm`. В графическом окружении РЕД ОС пакет можно установить двойным щелчком по файлу.
4. По завершении установки выводится сообщение `Dataset Studio 1.0.0: http://<адрес сервера>/`.
5. Проверьте работу: `systemctl status dataset-studio` (состояние active) и `curl http://127.0.0.1/api/health` (ответ `{"ok": true, ...}` с номером версии ПО).

В ходе установки пакет выполняет следующие действия:

- устанавливает nginx, rsync, cronie и policycoreutils-python-utils из штатного репозитория РЕД ОС;
- создаёт служебного пользователя `datasetstudio` без права входа и каталог `/data`;
- включает параметр SELinux `httpd_can_network_connect`, разрешающий веб-серверу обращаться к серверной части ПО;
- открывает в firewalld службы http и https, если межсетевой экран включён;
- включает автозапуск служб dataset-studio, nginx и crond.

6.2. Установка из архива (альтернативный способ)

1. Проверьте контрольную сумму архива `dataset-studio-1.0.0-redos8-x86_64.tar.gz` аналогично п. 6.1.
2. Распакуйте архив командой `tar -xzf dataset-studio-1.0.0-redos8-x86_64.tar.gz` и перейдите в созданный каталог.
3. Запустите установщик: `sudo ./deploy/redos/install.sh`. Результат установки аналогичен установке из RPM-пакета.

6.3. Обновление и удаление

- Обновление выполняется командой `sudo dnf install ./dataset-studio-<новая версия>.rpm`. Данные в `/data` и настройки в `/etc/dataset-studio/env` при этом сохраняются, служба перезапускается автоматически.
- Удаление выполняется командой `sudo dnf remove dataset-studio`. Данные и служебная учётная запись при этом сохраняются; для их удаления выполните `sudo rm -rf /data && sudo userdel datasetstudio`.
- При установке из архива удаление выполняется командой `sudo ./deploy/redos/uninstall.sh`; при указании ключа `--purge` удаляются также данные.

7. Первичная настройка

После установки следует создать учётную запись администратора и, если требуется, настроить защищённое соединение (HTTPS).

1. Откройте в браузере `http://<адрес сервера>/`. При первом обращении отображается форма создания учётной записи администратора, в которой указываются адрес электронной почты и пароль.
2. В разделе «Команда» добавьте пользователей и назначьте роли. Если отправка почты не настроена, администратор передаёт сотруднику ссылку-приглашение самостоятельно.

3. При необходимости укажите в настройках SMTP-сервер для автоматической отправки приглашений и ссылок восстановления пароля. Пароль SMTP хранится в зашифрованном виде.
4. Для работы по HTTPS установите сертификат организации по шаблону блока 443 в `/etc/nginx/conf.d/dataset-studio.conf` или выпустите его через certbot: `certbot --nginx -d <домен>`. После изменения: `sudo nginx -t && sudo systemctl reload nginx`.

Параметры службы задаются в `/etc/dataset-studio/env`. Для применения изменений выполните `sudo systemctl restart dataset-studio`.

Параметр	По умолчанию	Назначение
`DATASET_STUDIO_DATA`	`/data`	каталог данных
`DATASET_STUDIO_DB`	`1`	хранение разметки в базе SQLite
`DATASET_STUDIO_PORT`	`5057`	локальный порт приложения (должен совпадать с `upstream` в nginx)
`DATASET_STUDIO_THREADS`	`8`	число потоков обработки запросов
`DS_APP_USER`	`datasetstudio`	пользователь, от которого идёт резервное копирование

8. Эксплуатация

Работа с данными в ПО ведётся в рамках проектов и включает загрузку изображений, разметку, приёмку, формирование набора данных и экспорт. Доступные пользователю функции определяются его ролью.

Роль	Полномочия
Разметчик	разметка изображений и отправка их на проверку
Аналитик	полномочия разметчика, приёмка, баланс классов, аналитика, ведение списка классов
Администратор	все функции: проекты, загрузка, назначения, экспорт, пользователи, настройки, журнал действий

Типовой порядок работы:

1. Администратор создаёт проект и задаёт список классов объектов.
2. Администратор загружает изображения в разделе «Загрузка». Дубликаты исключаются автоматически, загруженные изображения объединяются в партию. Ранее размеченный набор данных загружается через раздел «Импорт COCO».
3. Администратор назначает партию разметчику на доске «Разметка».
4. Разметчик рисует рамки (`B`), полигоны (`P`) или повёрнутые рамки (`O`). Непригодное изображение отбраковывается клавишей `X`. Разметка сохраняется автоматически.
5. Разметчик нажимает кнопку «На проверку», после чего размеченные изображения передаются на приёмку.
6. Аналитик в разделе «Приёмка» принимает изображение, возвращает его на доработку с комментарием или отбраковывает. Принятые изображения включаются в набор данных.
7. Аналитик оценивает распределение классов и тепловую карту расположения объектов.
8. Администратор в разделе «Экспорт» задаёт лимиты по классам и размерам, выбирает формат COCO, YOLO или DOTA и скачивает собранный архив.

Основные сочетания клавиш редактора разметки приведены в таблице ниже.

Действие	Клавиши
Масштаб	`Ctrl` + колесо мыши
Перемещение изображения	`Space` + перетаскивание
Удалить выделенный объект	`Del`
Соседнее изображение	`←` `→`
Подтвердить предразметку	`Enter`
Отбраковать изображение	`X`

После обновления ПО пользователям следует перезагрузить страницу с очисткой кэша браузера (`Ctrl+F5`).

9. Администрирование и обслуживание

Управление службой ПО выполняется штатными средствами systemd. Резервное копирование выполняется автоматически ежедневно в 03:30.

Задача	Команда
Состояние службы	`systemctl status dataset-studio`
Перезапуск	`sudo systemctl restart dataset-studio`
Журнал приложения	`journalctl -u dataset-studio -f`
Журналы веб-сервера	`/var/log/nginx/access.log`, `/var/log/nginx/error.log`
Проверка работоспособности	`curl http://127.0.0.1/api/health`
Резервная копия вручную	`sudo /opt/dataset-studio/bin/backup.sh backup`
Список резервных копий	`sudo /opt/dataset-studio/bin/backup.sh list`
Восстановление	`sudo /opt/dataset-studio/bin/backup.sh restore <архив>`

Автоматическое резервное копирование включает:

- метаданные: архив, содержащий согласованную копию базы данных `studio.db` (включая всю разметку), учётные записи и настройки; срок хранения 14 дней;
- изображения: инкрементальные копии, создаваемые утилитой rsync (неизменённые файлы повторно места не занимают); хранятся 7 последних копий.

Копии сохраняются в `/data/backups` на том же диске и защищают от ошибочного удаления данных, но не от отказа носителя. Рекомендуется регулярно копировать каталог `/data/backups` на внешнее хранилище средствами организации.

Ключ шифрования секретов `/data/auth/.key` по умолчанию в архив не включается. Для переноса ПО на другой сервер создайте резервную копию с ключом `--with-key`; такой архив относится к конфиденциальной информации.

10. Информационная безопасность

ПО функционирует с минимально необходимыми правами и не обращается к внешним сервисам без соответствующей настройки.

- Служба выполняется от имени непривилегированной учётной записи `datasetstudio` с изоляцией systemd (`ProtectSystem`, `ProtectHome`, `PrivateTmp`, `NoNewPrivileges`); запись разрешена только в `/data`.
- Серверная часть принимает соединения только по адресу 127.0.0.1; из внешней сети доступен только веб-сервер nginx. SELinux работает в режиме Enforcing.

- Пароли пользователей хранятся в виде хэшей PBKDF2; пароль SMTP зашифрован локальным ключом.
- Доступ к функциям разграничивается по ролям; действия пользователей регистрируются в журнале.
- Веб-сервер передаёт защитные HTTP-заголовки (`X-Frame-Options`, `X-Content-Type-Options`, `Referrer-Policy`) и скрывает версию сервера.
- Исходящие соединения возможны только с SMTP-сервером и сервисом предварительной разметки, если они настроены администратором. Передача телеметрии и сетевая проверка лицензии не выполняются.

11. Контакты

По вопросам установки и эксплуатации ПО следует обращаться в службу технической поддержки ООО «ИСР»: электронная почта info@imr.center, телефон [заполнить], сайт imr.center. Порядок обращения и сроки рассмотрения обращений приведены в документе «Описание процессов, обеспечивающих поддержание жизненного цикла ПО».