

Общество с ограниченной ответственностью «МедДатаМаркет»  
125315, г. Москва, вн.тер.г. муниципальный округ Сокол,  
пр-кт Ленинградский, д. 80 к. 17, помещ. 1/1, Тел: +7 (925) 107-28-77  
ОКПО 98035622, ОГРН 1247700446399, ИНН / КПП 7743452881 / 774301001

---

**ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ «ДАТА ФАБРИКА»  
ДЛЯ ФОРМИРОВАНИЯ ВЕРИФИЦИРОВАННЫХ НАБОРОВ  
МЕДИЦИНСКИХ ДАННЫХ**

**(ПО «Дата Фабрика»)**

ОПИСАНИЕ ФУНКЦИОНАЛЬНЫХ ХАРАКТЕРИСТИК ПРОГРАММНОГО  
ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИЯ, НЕОБХОДИМАЯ ДЛЯ УСТАНОВКИ И  
ЭКСПЛУАТАЦИИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

## Оглавление

<b>Оглавление</b>	<b>2</b>
<b>Аннотация</b>	<b>3</b>
<b>Термины и определения</b>	<b>4</b>
<b>1. Общая информация о ПО</b>	<b>5</b>
1.1 Название ПО	5
1.2 Краткое описание ПО	5
1.3 Цели	5
1.4 Задачи	5
1.5 Назначение	6
1.6 Области применения	6
<b>2. Описание функциональных характеристик ПО</b>	<b>7</b>
2.1 Загрузка и хранение медицинских изображений	7
2.2 Аннотация и разметка	7
2.3 Верификация данных	7
2.4 Управление проектами и датасетами	8
2.5 API и интеграция	8
2.6 Автоматизация процессов	8
2.7 Безопасность и управление доступом	9
2.8 Поддержка многопользовательской работы	9
<b>3. Информация, необходимая для установки и эксплуатации ПО «Дата Фабрика»</b>	<b>10</b>
3.1. Системные требования	10
3.1.1 Аппаратные требования	10
3.1.2 Программные требования к АРМ пользователя	10
3.1.3 Сетевые требования	10
3.2 Эксплуатация	10
<b>4. Требования к компетенциям пользователей</b>	<b>11</b>

## **Аннотация**

Настоящий документ содержит описание функциональных характеристик программного обеспечения «Дата Фабрика», разработанного ООО «МедДатаМаркет», а также информацию, необходимую для его установки и эксплуатации. В документе представлены сведения о ключевых функциях и возможностях ПО, его системных требованиях, особенностях эксплуатации, а также требования к компетенциям пользователей и администраторов.

Раздел 1 посвящен общей информации о программном обеспечении, его названию, краткому описанию, целям и задачам, а также области применения.

Раздел 2 содержит описание функциональных характеристик ПО. Здесь детализируются процессы загрузки и хранения медицинских изображений, аннотация и разметка данных, верификация результатов, управление проектами и датасетами, использование API для интеграции с внешними системами, а также меры безопасности и поддержка многопользовательской работы.

Раздел 3 охватывает системные и программные требования для установки программного обеспечения, включая требования к аппаратной части, программным компонентам и сетевой инфраструктуре.

Раздел 4 определяет требования к компетенциям пользователей, указывая на навыки, необходимые для эффективного использования и администрирования ПО, включая опыт работы с веб-приложениями и понимание принципов безопасности данных.

Этот документ ориентирован на технических специалистов, занимающихся установкой, эксплуатацией и поддержкой ПО «Фабрика датасетов», а также на пользователей, работающих с медицинскими изображениями.

## Термины и определения

Термин	Определение
DICOM	(Digital Imaging and Communications in Medicine) стандарт для хранения и передачи медицинских изображений, таких как рентгеновские снимки, КТ или МРТ, используемый в медицинской диагностике
NRRD	(Nearly Raw Raster Data) формат данных, предназначенный для хранения томографических данных, таких как изображения компьютерной томографии
Разметка	процесс и результат нанесения графических меток (окружности, контуры, эллипсы и т.д.) на медицинские изображения для выделения патологий или других важных объектов
Верификация данных	процесс проверки корректности разметки медицинских изображений врачами-клиницистами для обеспечения надежности данных
API	(Application Programming Interface) программный интерфейс, предоставляющий набор функций для взаимодействия с другими системами и приложениями
PACS	(Picture Archiving and Communication System) система архивирования и передачи медицинских изображений, широко используемая в радиологии для управления и хранения данных
JSON	(JavaScript Object Notation) формат данных, используемый для хранения текстовых аннотаций в виде ключ-значение
PNG	(Portable Network Graphics) формат графических файлов, используемый для хранения изображений и графической разметки

## **1. Общая информация о ПО**

### **1.1 Название ПО**

Программное обеспечение «Дата Фабрика».

### **1.2 Краткое описание ПО**

ПО представляет собой специализированное программное обеспечение для создания, аннотации и верификации медицинских датасетов, используемых в области лучевой диагностики и разработки моделей искусственного интеллекта.

### **1.3 Цели**

ПО предназначено для создания высококачественных верифицированных наборов медицинских изображений, а также их обработки, разметки и аннотации. ПО позволяет врачам-клиницистам загружать DICOM и NRRD изображения, проводить сбор интерпретированных данных, верифицировать и комплектовать датасеты с документацией для использования в разработке и тестировании искусственных интеллектуальных моделей в области лучевой диагностики.

### **1.4 Задачи**

- Создание верифицированных медицинских датасетов: обеспечение функционала для загрузки, аннотирования и разметки медицинских изображений (в форматах DICOM и NRRD), с целью точного выявления патологий и диагностически значимых признаков.
- Верификация данных: обеспечение точной проверки и подтверждения корректности разметки и аннотаций, выполненных врачами-клиницистами, для создания качественных и надежных наборов данных.
- Поддержка разработки ИИ-моделей: обеспечение возможности использования сформированных датасетов для обучения нейронных сетей, используемых в медицинской диагностике.
- Автоматизация процессов обработки данных: сокращение времени и ресурсов, затрачиваемых на разметку и верификацию данных, за счет предоставления удобных и эффективных инструментов для пользователей.

### **1.5 Назначение**

ПО предназначено для создания верифицированных наборов медицинских изображений, пригодных для использования в разработке и тестировании моделей искусственного интеллекта в области медицинской диагностики. ПО также используется для разметки и аннотирования изображений специалистами, что обеспечивает создание качественных датасетов, необходимых для научных исследований и разработки ИИ-моделей, способных повысить точность медицинской диагностики.

### **1.6 Области применения**

- Медицина и лучевая диагностика: использование для разметки, аннотации и создания наборов данных для ИИ-моделей, применяемых для диагностики заболеваний на основе медицинских изображений.
- Разработка ИИ-систем в медицине: использование в процессе обучения и тестирования нейронных сетей, применяемых для анализа медицинских изображений.
- Образование и подготовка специалистов: применение в учебных и образовательных учреждениях для повышения квалификации медицинских специалистов в области работы с медицинскими данными и ИИ-технологиями.

## 2. Описание функциональных характеристик ПО

Функциональные характеристики ПО включают следующие ключевые компоненты и возможности, которые поддерживают процесс создания, аннотации и верификации медицинских изображений, в первую очередь для обучения и тестирования моделей искусственного интеллекта.

### 2.1 Загрузка и хранение медицинских изображений

- Поддерживаемые форматы: ПО поддерживает загрузку и обработку изображений в форматах DICOM (стандарт для медицинских изображений) и NRRD (формат для работы с томографическими данными). Это позволяет работать с изображениями, полученными с помощью таких методов диагностики, как компьютерная томография (КТ) и маммография (МГ).
- Хранение данных: изображения и связанные данные (включая разметку) хранятся в объектном хранилище, основанном на технологии MinIO, что обеспечивает высокую производительность и гибкость хранения.

### 2.2 Аннотация и разметка

- Инструменты разметки: ПО предоставляет графические инструменты для разметки, такие как окружности, контуры, эллипсы и кисти, для точной аннотации изображений. Эти инструменты позволяют врачам-клиницистам выделять патологические участки или другие диагностически значимые признаки.
- Типы разметки: ПО поддерживает как графическую, так и текстовую аннотацию изображений.
- Формат хранения разметки: разметка сохраняется в формате JSON для неграфических признаков и в формате PNG для графических меток.

### 2.3 Верификация данных

- Первичная валидация: ПО позволяет проводить первичную проверку загруженных изображений для оценки их соответствия требованиям к разметке и аннотации. Это помогает выявлять ошибки до начала детальной обработки данных.

- Верификация разметки: ПО включает этап проверки разметки для подтверждения ее корректности и точности. Этот процесс позволяет исключить дефекты и ошибки в аннотации перед финальным использованием датасета.

#### **2.4 Управление проектами и датасетами**

- Создание и управление датасетами: пользователи могут создавать новые проекты и датасеты, а также наполнять их медицинскими изображениями. ПО позволяет определять для датасетов клинические задачи и отслеживать прогресс их выполнения. Каждое исследование или набор данных проходит через различные этапы: загружено, в процессе аннотации, верифицировано и готово к использованию.
- Просмотр изображений: ПО предоставляет интерфейс для просмотра изображений и разметок в браузере, что позволяет удобно анализировать изображения и взаимодействовать с результатами аннотации.

#### **2.5 API и интеграция**

- Программный интерфейс (API): ПО предоставляет набор API для взаимодействия с другими системами и программами, поддерживая расширяемость и интеграцию с другими приложениями. Примером может служить интеграция с ПО MedMarker для дополнительных функций аннотации.
- Поддержка микросервисной архитектуры: ПО реализовано в виде набора микросервисов, что обеспечивает модульность, гибкость в развертывании и масштабируемость системы. Это включает контейнеры для взаимодействия с хранилищами данных, API и выполнения фоновых задач.

#### **2.6 Автоматизация процессов**

- Пакетная обработка: Пользователи могут загружать изображения пакетами, что позволяет одновременно обрабатывать несколько файлов и ускорять рабочий процесс.
- Фоновые задачи: ПО поддерживает выполнение фоновых задач, таких как преобразование медицинских изображений в нужные форматы (например, перевод из DICOM в PNG). Это снижает нагрузку на пользователей и ускоряет обработку данных.

## **2.7 Безопасность и управление доступом**

- Авторизация и аутентификация: ПО использует ключи доступа для контроля доступа к данным, обеспечивая безопасное хранение и обработку медицинских изображений и аннотаций.
- Управление пользователями: администраторы системы могут создавать и управлять учетными записями пользователей, а также выдавать и отзывать ключи доступа.

## **2.8 Поддержка многопользовательской работы**

- ПО поддерживает одновременную работу до 15 пользователей без существенного снижения производительности. Пользователи могут одновременно загружать изображения, выполнять аннотацию и верификацию данных.

### **3. Информация, необходимая для установки и эксплуатации ПО «Дата Фабрика»**

ПО «Дата Фабрика» является серверным приложением и не требует установки или выполнения специальных процедур на рабочих станциях пользователей. Доступ к функциональности программы осуществляется через сетевое взаимодействие. Отсутствует необходимость в дополнительных действиях со стороны пользователей для работы с ПО.

Для успешной эксплуатации ПО необходимо соблюдение определенных требований к оборудованию, программному обеспечению и сетевой инфраструктуре, а также выполнение последовательности действий по запуску системы. Ниже представлены ключевые аспекты, которые следует учитывать:

#### **3.1. Системные требования**

##### **3.1.1 Аппаратные требования**

Пользовательские рабочие станции: Процессор не ниже Intel Core i7, минимум 8 ГБ оперативной памяти, операционная система Windows 10 Pro или выше.

##### **3.1.2 Программные требования к АРМ пользователя**

Веб-браузеры: Google Chrome версии 121 и выше, Mozilla Firefox версии 122 и выше, Microsoft Edge версии 121 и выше.

##### **3.1.3 Сетевые требования**

Пропускная способность сети: Рекомендуется наличие сети с пропускной способностью не менее 1 Гбит/с для обеспечения быстрого доступа к данным и минимизации задержек при обмене данными.

#### **3.2 Эксплуатация**

Доступ к функциональности ПО осуществляется через веб-интерфейс с использованием современных браузеров. Пользователи входят в систему по предоставленным Разработчиком учетным данным.

Пользователи получают доступ к возможностям ПО через интуитивно понятный интерфейс.

#### **4. Требования к компетенциям пользователей**

Пользователи должны обладать базовыми навыками работы с веб-приложениями и основами безопасности данных (использование паролей). Знание работы с системами PACS и медицинскими информационными системами будет преимуществом.

Администраторы системы должны иметь опыт работы с серверными операционными системами, навыки управления контейнеризацией и понимание баз данных для эффективного управления и настройки ПО.