

**Программа для ЭВМ**  
**«Рукат - системное ядро автоматизированного переводчика»**  
**(ПрЭВМ «Рукат»)**

Документация, содержащая описание функциональных характеристик программного обеспечения и информацию, необходимую для установки и эксплуатации программного обеспечения

На 32 листах

## **Аннотация**

Настоящий документ распространяется на Программу для ЭВМ «Рукат - системное ядро автоматизированного переводчика» (далее – Система, ПрЭВМ «Рукат»), программное обеспечение).

Данный документ содержит описание функциональных характеристик программного обеспечения и информацию, необходимую для установки и эксплуатации программного обеспечения.

В разделе «Термины и сокращения» указаны основные аббревиатуры, понятия и их описание.

В разделе «Общие сведения о программном продукте» указаны основное положение, возможности системы, принципы и подходы к созданию системы.

В разделе «Функциональное описание системы» приведены сведения об эксплуатационном назначении системы, функциональных возможностях системы, общей архитектуре системы, функциональном описании модулей системы.

В разделе «Информация, необходимая для установки и эксплуатации программного обеспечения» приведены общие сведения необходимые для установки программного обеспечения, также требования к персональному компьютеру пользователя, требования к серверному обеспечению, требования к браузерам.

## **ВНИМАНИЕ!**

Сведения о структуре, принципе организации Системы, составе аппаратных и программных средств, а также другие данные, которые имеют отношение к функционированию Системы, являются строго конфиденциальными, не подлежат разглашению.

## Содержание

1.	Термины и сокращения.....	4
2.	Общие сведения о программном продукте. ....	6
2.1	Основное положение .....	6
2.2	Принципы и подходы к созданию системы.....	7
3.	Функциональное описание системы.....	10
3.1	Эксплуатационное назначение Системы.....	10
3.2	Функциональные возможности Системы .....	10
3.3	Общая архитектура системы .....	11
3.3.1	Структура хранения данных .....	14
3.3.2	Описание микросервисов.....	14
3.3.3	Микросервисы ядра.....	16
3.4	Модули ПрЭВМ «Рукат» .....	19
3.4.1	Функция модуля «Проекты» .....	19
3.4.2	Функция модуля «Глоссарий».....	21
3.4.3	Функция модуля «ТМ-файлы».....	22
3.4.4	Функция модуля «Ключи МТ».....	22
3.4.5	Организация рабочего пространства переводчика в ПрЭВМ «Рукат» .....	24
3.5	Организация информационной безопасности Системы .....	26
3.5.1	Управление паролями .....	26
3.5.2	Обеспечение информационной безопасности ПрЭВМ «Рукат» .....	27
4.	Информация, необходимая для установки и эксплуатации программного обеспечения	29
4.1	Работа системы .....	29
4.1	Требования к программному обеспечению ПК пользователя .....	30
4.2	Рекомендуемые требования к серверному обеспечению .....	30
4.3	Требования к браузерам .....	30
5.	Заключение .....	32

## 1. Термины и сокращения

Обозначение	Описание
Аутентификация	Проверка подлинности пользователя путём сравнения введённого им пароля (для указанного логина) с паролем, сохранённым в базе данных пользовательских логинов
БД	База данных
БП	База переводов
Базы переводческой памяти (Translation Memory)	Память переводов (или ТМ) — это база данных, в которую в виде сегментов сохраняются переводы текстов с одного исходного языка на один или несколько языков перевода.
Глоссарии	Наборы терминов с вариантами перевода на один или несколько языков, толкованиями, комментариями и прочей дополнительной информацией.
Кэш-память	Промежуточный буфер с быстрым доступом к нему, содержащий информацию, которая может быть запрошена с наибольшей вероятностью.
Машинный перевод (Machine translate)	Процесс перевода текстов (письменных, а в идеале и устных) с одного естественного языка на другой с помощью специальной компьютерной программы.
ПО	Программное обеспечение
ПК	Персональный компьютер
РуКат	Системное ядро программного обеспечения
Система	ПрЭВМ «РуКат»
СУБД	Система управления базами данных
Токен	Представляет собой зашифрованную последовательность символов, которая позволяет точно идентифицировать объект и определить уровень его привилегий.
Хеш	Хеш-функция, или функция свёртки — функция, осуществляющая преобразование массива входных данных произвольной длины в битовую строку установленной длины, выполняемое определённым алгоритмом.
Экранирование	Замена в тексте управляющих символов на соответствующие текстовые подстановки.
API	Программный интерфейс приложения

Backend	Серверная часть системы
Frontend	Клиентская часть системы
gRPC	Представляет фреймворк, который использует протокол RPC (Remote Procedure Call) для обмена сообщениями между клиентом и сервером.
HTTP	Протокол прикладного уровня передачи данных.
JWT	Это открытый стандарт для создания токенов доступа, основанный на формате JSON. Как правило, используется для передачи данных для аутентификации в клиент-серверных приложениях. Токены создаются сервером, подписываются секретным ключом и передаются клиенту, который в дальнейшем использует данный токен для подтверждения своей личности.
NIST	Национальный институт стандартов и технологий
PBKDF2	Стандарт формирования ключа на основе пароля.
RBAC	Управление доступом по ролевой модели
REST	Набор архитектурных принципов построения сервис-ориентированных систем.
SHA256	Криптографический алгоритм хэширования

## 2. Общие сведения о программном продукте.

### 2.1 Основное положение

ПрЭВМ «Рукав» реализовано в виде веб-приложения и представляет собой полуавтоматический переводчик текстов построенный на системной обработке текста любого объема в общепринятых форматах (.docx, .xlsx, .pptx, .html, .xliff, .tmx).

Возможности системы:

- Синхронизацию совместной работы переводчиков и менеджеров;
- Планирование процесса перевода;
- Удобное представление информации по проекту;
- Верификация качества переводов на основании загруженных глоссариев;
- Использование машинного перевода;
- Подключение готовых глоссариев;
- Создание новых глоссариев;
- Объединение ресурсов переводчиков и редакторов;
- Обеспечение одновременного доступа к рабочим документам с установленным допуском конфиденциальности для различных категорий пользователей.

Пользователь получает доступ к системе через браузер или по программному интерфейсу приложения (далее — API).

Система состоит из двух частей: backend части, именуемой в дальнейшем «Рукав» и frontend части — интерфейс программного обеспечения.

«Рукав» — ядро системы, реализовывает полуавтоматический перевод текстов (далее — ядро).

Интерфейс Системы доступен из публичного сегмента сети Интернет и поддерживается большинством браузеров. Адаптивная верстка позволяет

использовать различные устройства для работы в Системе без потери информации. Язык интерфейса – русский.

В зависимости от требований к программному обеспечению, интерфейс ПрЭВМ «Рукат» может иметь разный дизайн и сама Система может оснащаться дополнительными модулями, но основные модули ядра Системы и его архитектурное решение останутся неизменными.

## **2.2 Принципы и подходы к созданию системы**

При создании Системы использовались следующие принципы и подходы:

- Российское ПО — единственная российская система полного контроля и управления потоком переводов.

Единственное российское решение для обслуживания масштабных корпоративных переводческих проектов. Система разворачивается на серверах клиента и разработчика на территории РФ. Таким образом обеспечивается полная независимость от иностранных систем;

- Конфиденциальность и интернет — полный контроль и обеспечение защищенности данных, содержащихся в переводе.

Обеспечение конфиденциальности при работе в условиях распределенной сети и нестабильных ресурсов, используемых через сеть Интернет. Защита паролем с использованием токенов. Для аутентификации пользователя используется технология JWT.

Возможность развернуть продукт на серверах клиента.

- Автоматизация работы с документами — прогрессивная автоматизация процесса корпоративных переводов без потери качества. Возможность подключения машинного перевода.

Система – платформа для автоматизации всех процессов документооборота компании (создание, перевод, редактирование, форматирование, верстка)

Мы предлагаем масштабируемое решение, которое можно настроить под потребности клиента. К системе можно подключить неограниченное число пользователей.

Программное обеспечение легко интегрируется в корпоративный интранет;

- Аппаратное накопление знаний — аппаратное накопление знаний и системное формирование корпоративного стандарта документов.

Программное обеспечение – ценное дополнение к экосистеме компании, укрепляющее ее позиции как одного из лидеров на рынке цифровизации.

Каждый новый документ обработанный переводчиками, редакторами и подтвержденный заказчиком становится важной частью корпоративного знания, которое автоматически воспроизводится системой при загрузке в нее новых схожих документов.

Система всегда предлагает использовать только те формулировки, которые были ранее подтверждены непосредственным заказчиком;

- Кастомизированное решение — кастомизированное решение, полностью встраивающееся в структурированный документопоток компании.

Система является полностью настраиваемым инструментом для управления процессами перевода, синхронизированным с управленческой архитектурой компании.

Ресурсы, задействованные в процессе перевода, управляются в соответствии с принципами документооборота заказчика. ПрЭВМ «Рукат» – единственная на рынке интегрированная система автоматизированного перевода и управления ресурсами предприятия, открытая для кастомизации и доработок в соответствии с уникальными требованиями заказчика;

- Управление ресурсами — точечное управление человеческими и электронными ресурсами в процессе перевода.



ПрЭВМ «Рукат» – платформа для управления всеми ресурсами компании, задействованными в процессе перевода.

Система является единым интегрированным решением – модуль для управления доступом к информации, а также кадровыми и временными ресурсами.

Программное обеспечение позволяет контролировать, согласовывать и планировать бюджет на перевод каждого документа и вести статистику всех действий над создаваемыми и переводимыми документами в защищенном облаке компании;

- Сокращение расходов — существенное сокращение расходов на поддержание любых лингвистических сервисов компании за счет накопленных данных.

ПрЭВМ «Рукат» создает базу знаний, которая позволяет максимально автоматизировать процессы работы со стандартными документами и формами. В результате отпадает необходимость проведения тендеров на оказание переводческих, редакторских и прочих лингвистических услуг.

### **3. Функциональное описание системы**

#### **3.1 Эксплуатационное назначение Системы**

Система предназначена для следующих категорий конечных пользователей:

1. Бюро переводов;
2. Переводческие отделы крупных компаний;
3. Частные переводчики (фрилансеры);

#### **3.2 Функциональные возможности Системы**

Система состоит из совокупности модулей, обеспечивающих выполнение следующих функций:

- Перевод исходного текста в защищенную облачную платформу на серверах клиента;
- Сопоставление выделенных строк с базой переводов (БП) в памяти системы;
- Базы данных (БД) структурированы по отраслям и хранятся в отдельных загружаемых файлах;
- Автоматический подбор из БД наилучших эквивалентов;
- Активизация инструментов для доводки текста до желаемого состояния;
- Компиляция переведенного текста в исходный формат;
- Бизнес-аналитика по проектам;
- Поиск терминологии с учетом морфологии;
- Проверка качества;
- Создание памяти переводов из имеющихся оригиналов и переводов;
- Использование машинного перевода;
- Создание глоссариев.

### 3.3 Общая архитектура системы

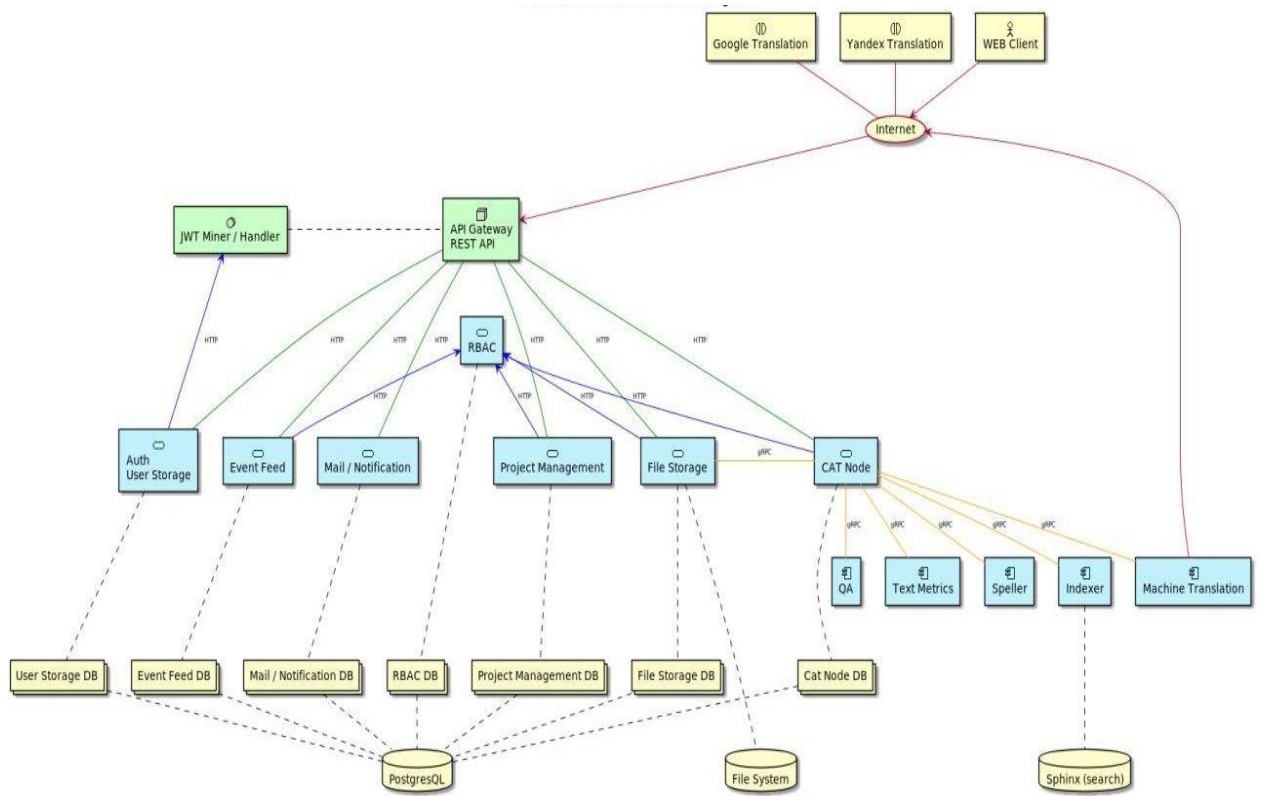


Рисунок 1 — архитектура ПрЭВМ «Рукат»

На рисунке 1 представлена общая схема архитектуры ПрЭВМ «Рукат».

Архитектура системы включает в себя:

- Система представляет собой многоконтейнерное приложение с микросервисной архитектурой.
- Архитектурный стиль Системы — Rest over HTTP.

REST — набор архитектурных принципов построения сервис-ориентированных систем.

HTTP — протокол прикладного уровня передачи данных.

- Общение между сервисами ядра (на схеме «CAT Node») осуществляется с помощью системы удаленного вызова процедур gRPC.

gRPC представляет фреймворк, который использует протокол RPC (Remote Procedure Call) для обмена сообщениями между клиентом и сервером. Цель фреймворка состоит в том, чтобы обеспечить высокую производительность в тех условиях, где это особенно критично, например, при интенсивном обмене информацией в режиме реального времени. Он использует HTTP / 2 для транспорта, буферы протокола в качестве языка описания интерфейса и предоставляет такие функции, как аутентификация, двунаправленная потоковая передача и управление потоком, блокирующие или неблокирующие привязки, а также отмены и тайм-ауты. Он генерирует кроссплатформенные привязки клиента и сервера для многих языков

- Многоконтейнерное приложение запускается с помощью технологии docker-compose и docker-swarm. Docker-compose используется для одновременного управления несколькими контейнерами, входящими в состав приложения.
- Каждый docker-compose разворачивает несколько docker-контейнеров внутри сервиса.
- Docker-swarm управляет кластерами, в которых располагаются несколько docker-контейнеров.
- Для каждого сервиса настроен свой веб-сервер на базе nginx.
- Общение со сторонними приложениями осуществляется с помощью API. К примеру, для сервиса машинного перевода используется API Yandex и Google переводчиков.

Для создания, управления, мониторингом, обслуживанием и обеспечением безопасности API используется API Gateway. API Gateway берет на себя все задачи, связанные с приемом и обработкой сотен тысяч одновременных вызовов API, включая управление трафиком, авторизацию и контроль доступа, регулирование количества запросов, мониторинг и управление

версиями API.

- JWT handler является обработчиком аутентификации.

ПрЭВМ «Рукат» использует *JWT* для проверки аутентификации пользователя следующим образом:

1. Сперва пользователь заходит на сервер аутентификации с помощью аутентификационного ключа (пара *логин/пароль*);
2. Затем сервер аутентификации создает *JWT* и отправляет его пользователю;
3. Когда пользователь делает запрос к API приложения, он добавляет к нему полученный ранее *JWT*;
4. Когда пользователь делает API запрос, приложение может проверить по переданному с запросом *JWT* идентифицируется ли пользователь в Системе. В этой схеме сервер приложения сконфигурирован так, что сможет проверить, является ли входящий *JWT* именно тем, что был создан сервером аутентификации.

- Для управления доступом по ролевой модели используется подход RBAC.

Суть подхода заключается в создании ролей, повторяющих бизнес-роли в компании, и присваивание их пользователям. На основе этих ролей проверяется возможность выполнения пользователем того или иного действия (лингвист, менеджер и прочее).

В зависимости от требований заказчика в Системе возможна организация разных типов ролей по RBAC подходу.

### 3.3.1 Структура хранения данных

Схема организации структуры хранения и управления данными:

- У каждого сервиса есть своя база данных;
- Базы данных сервисов управляются единой СУБД «PostgreSQL»;
- Отдельное хранение файловой системы;
- Используется система полнотекстового поиска «Sphinx», интегрируемая с СУБД «PostgreSQL».

### 3.3.2 Описание микросервисов

Auth User Storage является хранилищем информации об авторизованных пользователях ПрЭВМ «Рукат». Хранит токен пользователя переданный JWT обработчиком.

За взаимодействие между событиями в приложении в режиме реального времени отвечает Event Feed. Данные событий хранятся в базе данных PostgreSQL и заносятся в кэш-память Memcached.

Notification сервис отвечает за отправку сообщений и уведомлений.

Основная обязанность project management (управление проектами) сервиса состоит в:

- Поддержке сред перевода (Организаций);
- Внедрении иерархической модели представления бизнес-структур (Единица > Проект > Ресурс для перевода);
- Взаимодействие по Rest API с RBAC сервисом;
- Делегирование различных задач участникам организаций.

Сервис не управляет ресурсами, глоссариями и ТМ файлами.

RBAC управляет доступом по ролевой модели в Системе, созданной разработчиком по требованию заказчика.

RBAC взаимодействует с обработчиком аутентификации JWT. Схема взаимодействия представлена на рисунке 2.

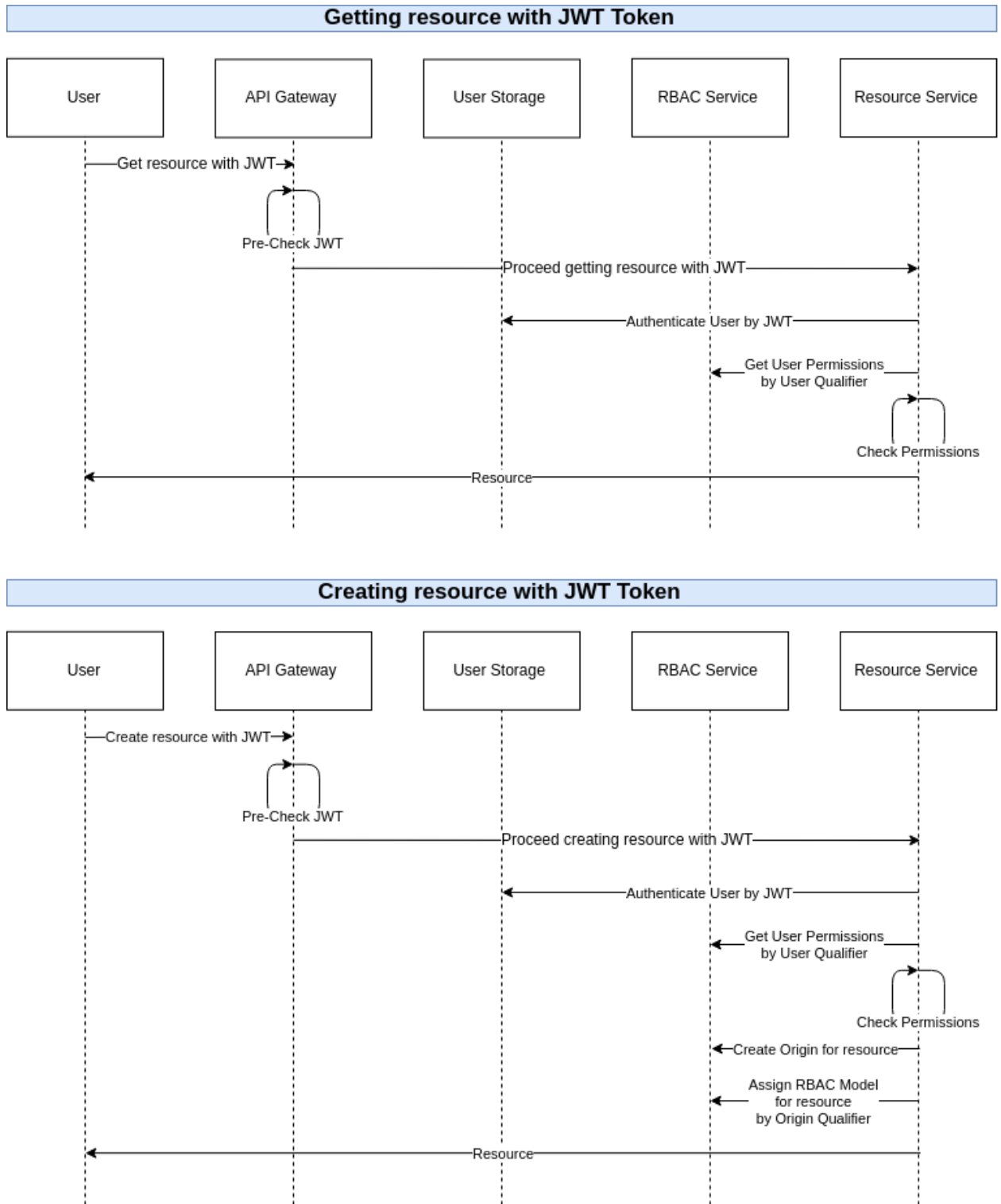


Рисунок 2 — схема взаимодействие RBAC с JWT

Работа схемы: пользователь проходит аутентификацию в Системе, ему выдается токен и выделяются права в соответствии с ролевой моделью. Токен пользователя и его права запоминаются в БД.

Сервис file storage (файловая система) отвечает за:

- Хранение различных крупногабаритных файлов;
- Загрузку файлов несколькими фрагментами.

Интерфейсы взаимодействия с файловой системой:

- Rest API;
- gRPC;
- Панель администратора.

### **3.3.3 Микросервисы ядра**

Сервис CAT Node состоит из следующих микросервисов:

1. QA (обеспечение качества).

Он обнаруживает ошибки в переводе, сравнивая исходные и целевые сегменты, проверяя терминологическую базу, орфографию, содержимое тегов и т.д. Сервис помогает переводчикам обнаруживать формальные ошибки и улучшает общее качество и согласованность перевода.

Производит:

- Автопроверку по различным фильтрам;
- Функцию проверки сегмента;
- Функцию проверки документа;
- Таблицу ошибок, найденных при проверке.



Интерфейс QA представлен на рисунке 3.

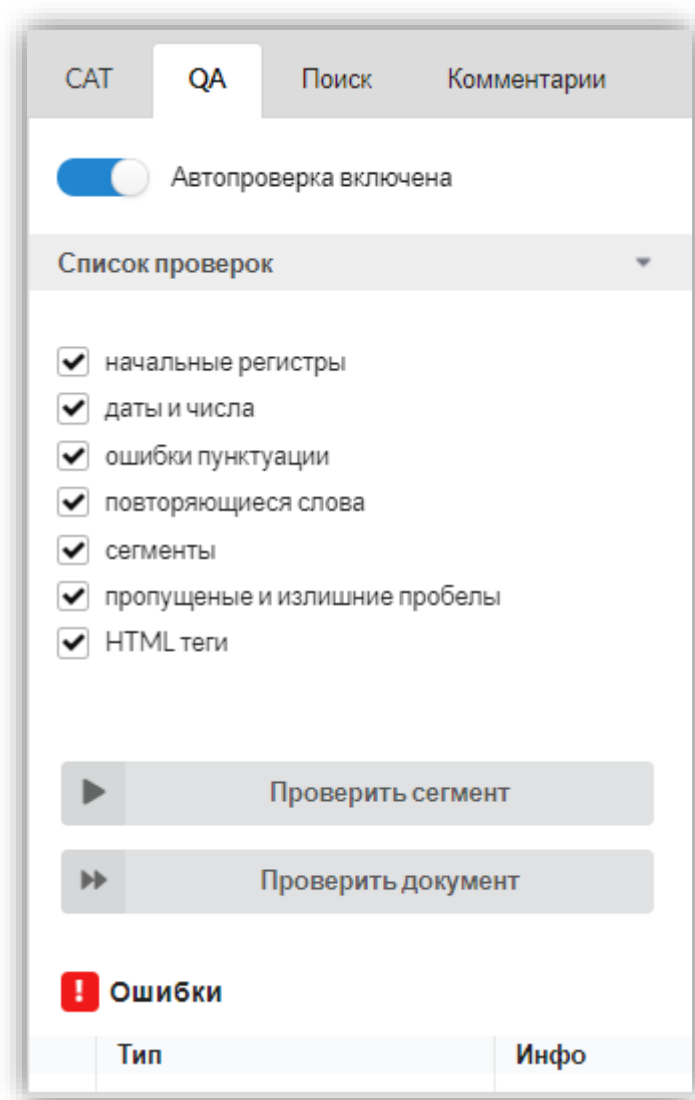


Рисунок 3 — интерфейс QA.

## 2. Text Metrics.

На вход микросервису передаётся текст, в ответ микросервис возвращает статистику слов, уникальных слов, символов, и так далее.

## 3. Speller

Speller — сервис для автоматической проверки орфографии. Он объединяет

общий набор функций, присутствующих в различных существующих продуктах / библиотеках для проверки орфографии, и предоставляет для этого стабильный API.

#### 4. Indexer

Принцип работы сервиса таков:

Все документы загружаемые в Систему индексируются (в индексе текстовой информации нет - данные представлены с помощью индекса) и с помощью специально составленных запросов ПО ищет термины и совпадения с существующей памятью переводов.

В качестве системы полнотекстового поиска используется sphinx.

#### 5. Machine Translation.

Для машинного перевода используются сторонние приложения (Bing Translator, Yandex Translate, Google Translate). Взаимодействие с внешними сервисами осуществляется с помощью API.

### 3.4 Модули ПрЭВМ «Рукат»

Рассмотрим реализацию основных модулей Системы в рамках интерфейса программного обеспечения ПрЭВМ «Рукат».

Следует учесть, что интерфейс Системы может меняться в зависимости от требований заказчика. При этом ядро Системы будет неизменно.

#### 3.4.1 Функция модуля «Проекты»

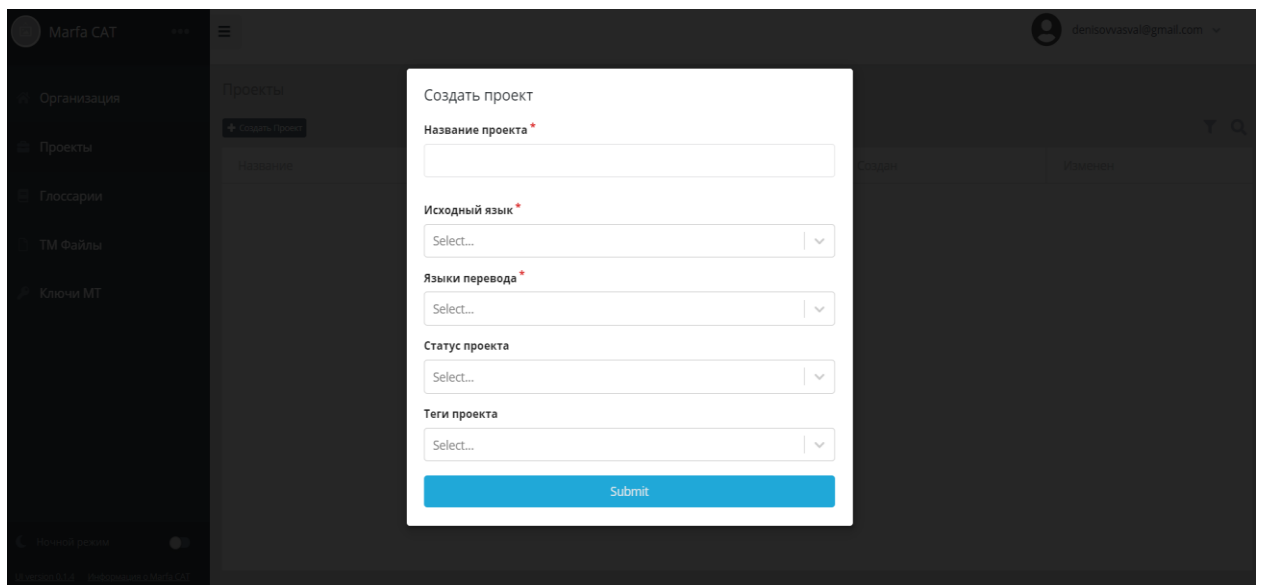


Рисунок 4 — модальное окно создания проекта

Пройдя аутентификацию, пользователь создает в Системе организацию и после этого попадает на страницу с интерфейсом модуля «Проект».

Для создания нового проекта нужно:

- Нажать на пункт меню «Проекты»;
- Нажать на кнопку «Создать проект»;
- Заполнить появившуюся форму.

На рисунке 4 представлено модальное окно создания проекта.

После заполнения модального окна пользователь попадает на интерфейс со списком текущих проектов организации (рисунок 5).

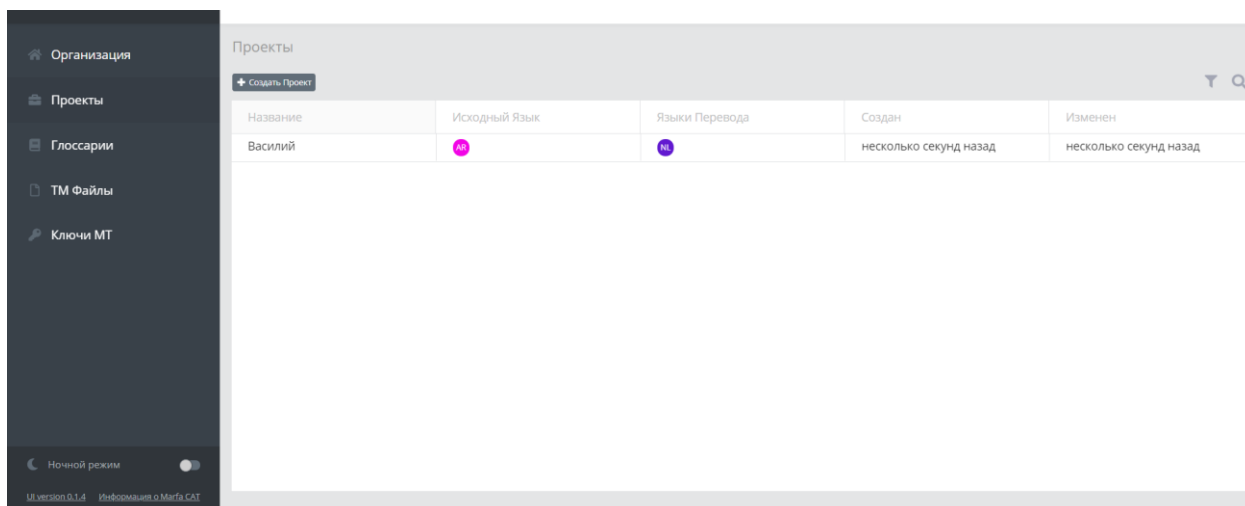


Рисунок 5 — интерфейс модуля «Проект» после заполнения модального окна

Далее пользователь попадает на страницу проекта (рисунок 6), где может:

- Посмотреть информацию о проекте;
- Создать ресурс для перевода по нажатию кнопки «Создать ресурс» во вкладке «Документы для перевода»;
- Посмотреть и изменить предпочтения перевода во вкладке «Предпочтения перевода» - добавить глоссарии, ТМ-базы и МТ ключи;
- Отредактировать данные по проекту: название, языки оригинала и перевода, статус и теги проекта;
- Добавить в проект исполнителей и распределить между ними роли.

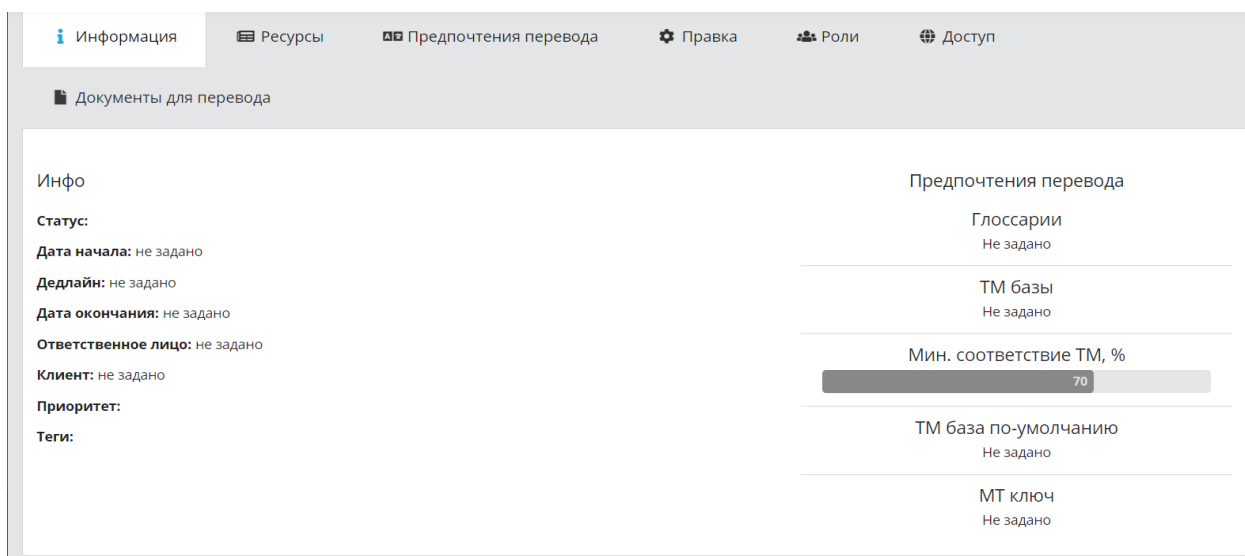


Рисунок 6 — информация о конкретном проекте.

Добавить документ в проект можно, перейдя во вкладку «Документы для перевода» и нажав кнопку «Загрузить файл». После добавления документа он появится в списке документов (рисунок 7).

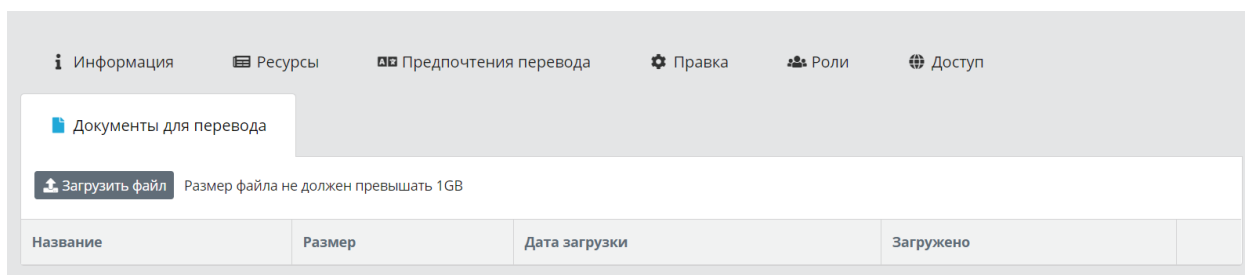


Рисунок 7 — добавление документа для перевода

### 3.4.2 Функция модуля «Глоссарий»

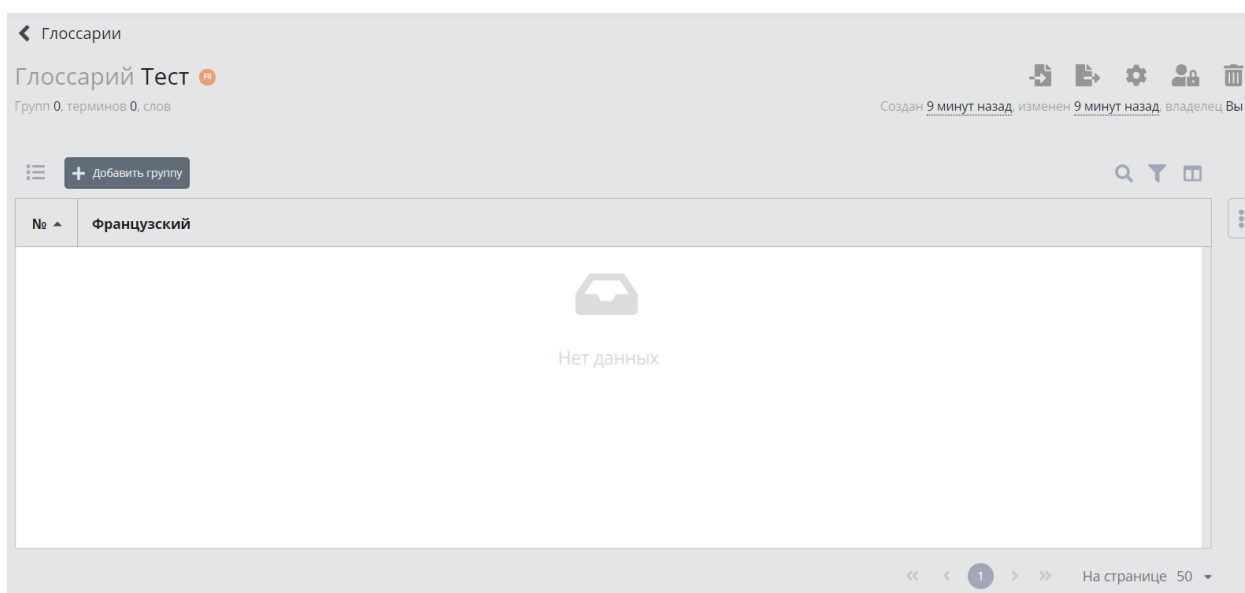


Рисунок 8 — интерфейс модуля «Глоссарий»

Глоссарий представляет собой таблицу, каждая строка которой является группой терминов (оригинал + перевод). Поиск нужных терминов можно осуществить с помощью формы поиска, сортировки и фильтрации.

С глоссарием можно совершать следующие действия:

- Экспортировать глоссарий (в формате `xlsx`);
- Изменять, добавлять, удалять группы терминов;
- Изменять свойства глоссария;

- Оставлять комментарии к группе терминов.

При создании глоссария в поле ввода «Языки глоссария» необходимо ввести и язык оригинала, и язык перевода.

### 3.4.3 Функция модуля «ТМ-файлы»

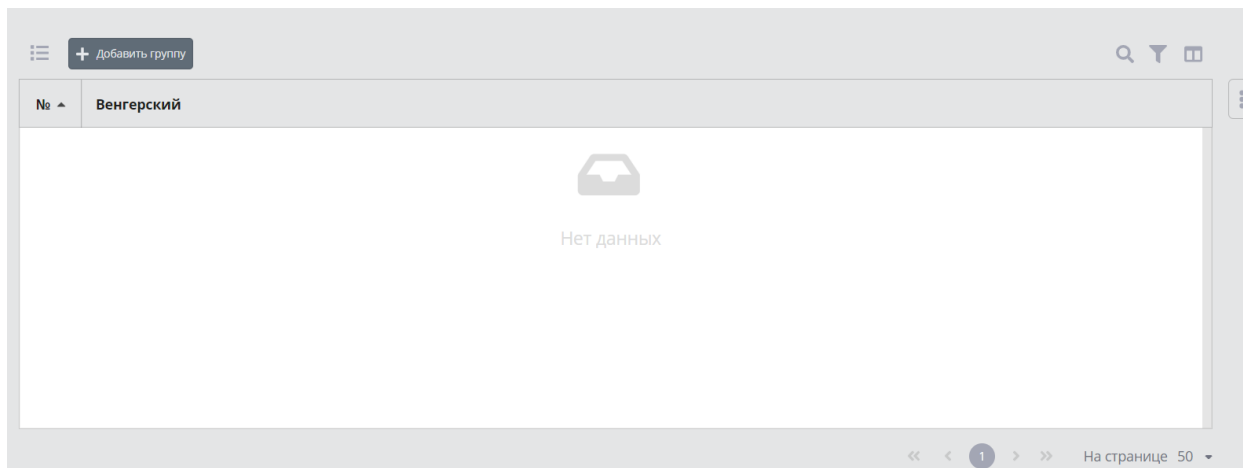


Рисунок 9 — интерфейс модуля «ТМ-файлы»

Алгоритм создания и работы с ТМ файлами аналогичен работе с глоссарием. Отличительной особенностью является возможность создания вариантов перевода сегмента текста. Интерфейс модуля представлен на рисунке 9.

### 3.4.4 Функция модуля «Ключи МТ»

Для того, чтобы воспользоваться машинным переводом в Системе, необходимо получить API ключ одной из внешних систем перевода.

Для создания нового ключа МТ необходимо:

- Перейти в пункт «Ключи МТ»;
- Нажать «+ Создать МТ ключ».

В выпадающем модальном окне выбрать «Тип инструмента МТ». Можно выбрать только сторонний ресурс. В зависимости от типа инструмента необходимо получить соответствующий ключ.

Полученный API ключ вставить в поле «Ключ инструмента MT», применить изменения.

Чтобы добавить MT-ключ, нужно:

- Перейти во вкладку «Предпочтения перевода»;
- Нажать на стрелку в блоке выбора;
- Выбрать из списка.

На рисунке 10 показан интерфейс Системы после добавления информации о машинном переводе.

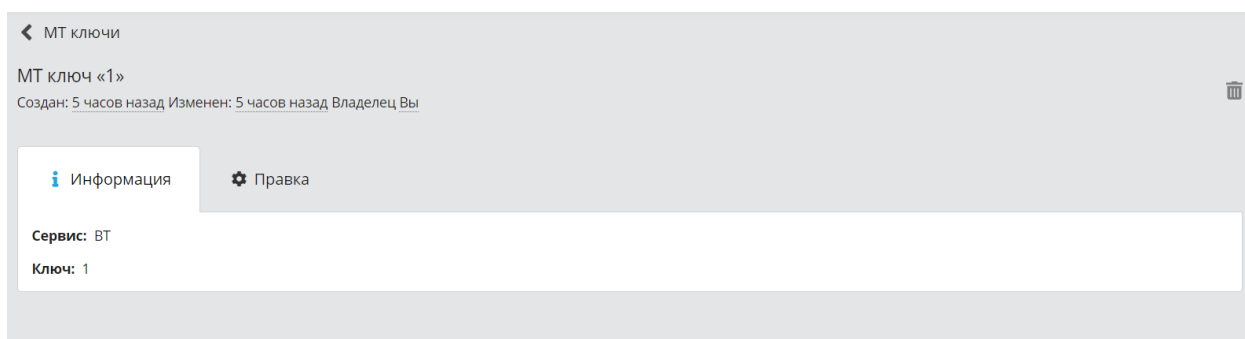


Рисунок 10 — информация о машинном переводе

В ПрЭВМ «Рукав» можно внести правки по использованию машинного перевода (рисунок 11).

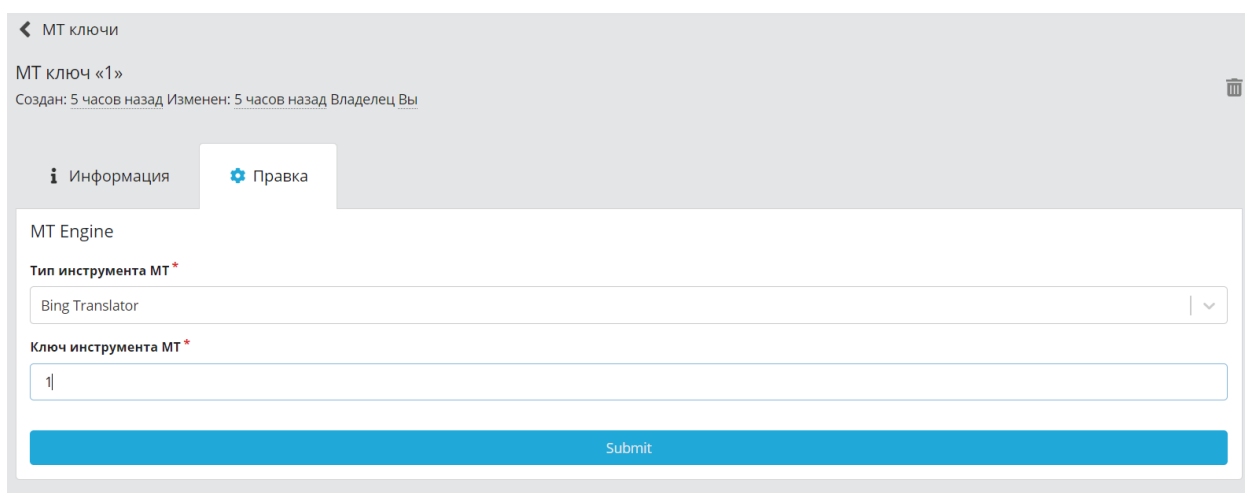


Рисунок 11 — правки машинного перевода

### 3.4.5 Организация рабочего пространства переводчика в ПрЭВМ «Рукат»

Чтобы перейти в интерфейс перевода, нужно открыть созданный ресурс в проекте и нажать на кнопку «Перевод».

Для создания ресурса нужно:

- перейти во вкладку «Документы для перевода»;
- нажать кнопку «Добавить документ»;
- после добавления документа нажать кнопку «Создать ресурс».

Интерфейс рабочего места переводчика представляет собой классический интерфейс CAT систем. В панели инструментов над таблицей располагаются следующие:

- Экспорт файла перевода;
- Машинный перевод;
- Перевод идентичных сегментов (применение перевода активного сегмента к идентичным ему сегментам);
- Перевод пустых идентичных сегментов (применение перевода ко всем пустым (непереведенным) сегментам, идентичным активному);
- Функция предварительного перевода;
- Добавление термина в глоссарий;
- Фильтры по сегментам.

Во вкладке «CAT»:

- Перевод сегмента с помощью онлайн-переводчика. Для подключения онлайн-переводчика нужно добавить ключ MT в предпочтения перевода;
- Термины (подгружаются из TM-файла, выбранного в предпочтениях перевода проекта);



- Память переводов (варианты перевода из ТМ-файла);
- История переводов (список изменений текущего сегмента с момента создания ресурса).

Вкладка «Поиск»:

- Конкордансный поиск по ТМ-базе, по языку оригинала и языку перевода.

Вкладка «Комментарии»:

- Написание, просмотр и удаление комментариев к сегменту.

## 3.5 Организация информационной безопасности Системы

### 3.5.1 Управление паролями

Для управлением паролями пользователей ПрЭВМ «Рукат» использует инструменты Django — фреймворка языка программирования Python.

В Django реализована гибкая система хранения паролей, а по умолчанию используется PBKDF2 (Password-Based Key Derivation Function — стандарт формирования ключа на основе пароля).

Атрибутом пароля объекта пользователь в этом формате является строка:

```
<algorithm>${iterations}${salt}${hash}
```

Это компоненты, используемые для хранения пароля пользователя, разделенные знаком доллара и состоящие из: алгоритма хэширования, количества итераций алгоритма (рабочего фактора), случайной соли и результирующего хэша пароля. Алгоритм является одним из нескольких алгоритмов одностороннего хэширования или хранения пароля, которые может использовать Django. Итерации описывают количество раз, когда алгоритм проходит через хэш. Соль является случайным семенем, а хэш - результатом односторонней функции.

По умолчанию Django использует алгоритм PBKDF2 с хэшем SHA256 (алгоритм шифрования), механизм растягивания пароля, рекомендованный NIST (Национальный институт стандартов и технологий).

Система использует этот алгоритм — он достаточно безопасен и требует большого количества вычислительного времени для взлома.

Django выбирает алгоритм, используя настройку `password_hashers` (список алгоритмов хэширования). Это список классов алгоритмов хэширования, которые поддерживает данная установка Django. Первая запись в этом списке (т.е. `settings.password_hashers[0]`) будет использоваться для хранения паролей,

а все остальные записи являются действительными хэшами, которые можно использовать для проверки существующих паролей. Это означает, что имеется возможность модифицировать `password_hashers` и менять список предпочтительных алгоритмов хэширования в Системе.

### **3.5.2 Обеспечение информационной безопасности ПрЭВМ «Рукат»**

Система использует внутренние инструменты Django для обеспечения информационной безопасности ПрЭВМ «Рукат». Django обеспечивает защиту конфиденциальных данных пользователей несколькими способами:

#### 1) Использование шаблонов и экранирования от XSS атак.

XSS атака — Это такой тип атак, который внедряет в веб-системы вредоносный код, заставляя её выдавать измененные данные, подменяет ссылки (видимые/скрытые) или выводит собственную рекламу на пораженном ресурсе.

Шаблоны Django избегают определенных символов, которые особенно опасны для HTML (язык разметки документов). Также, чтобы избежать внедрения вредоносного кода, следует использовать экранирование символов.

Экранирование символов — замена в тексте управляющих символов на соответствующие текстовые подстановки. Таким образом, Система не дает вторгнуться XSS атакам.

#### 2) Защита от CSRF

CSRF-атаки позволяют злоумышленнику совершать действия с использованием учетных данных другого пользователя без его ведома или согласия.

Django имеет встроенную защиту от большинства типов CSRF-атак. Защита CSRF в Системе работает путем проверки секретного ключа в каждом POST-запросе. Это гарантирует, что злоумышленник не сможет загрузить форму POST на ваш сайт и заставить другого зарегистрированного пользователя непреднамеренно отправить эту форму. Вредоносный пользователь должен знать секретный ключ, который является специфическим для пользователя.

### 3) Защита от SQL-инъекций

SQL — язык программирования структурных запросов к БД для различной работы с данными.

SQL-инъекция - это тип атаки, при котором злоумышленник может выполнить произвольный SQL-код в базе данных. Это может привести к удалению записей или утечке данных.

Запросы Django защищены от SQL-инъекции, так как их запросы строятся с использованием параметризации запроса. SQL-код запроса определяется отдельно от параметров запроса. Поскольку параметры могут быть предоставлены пользователем и, следовательно, небезопасны, они экранируются базовым драйвером БД.

### 4) X-Frame Options.

Clickjacking - это тип атаки, при котором вредоносный сайт обертывает другой сайт во фрейм. Эта атака может привести к тому, что ничего не подозревающего пользователя обманом заставят совершить непреднамеренные действия на целевом сайте.

Django содержит защиту от перехвата щелчков в виде промежуточного ПО X-Frame-Options, которое в поддерживающем его браузере может препятствовать тому, чтобы сайт оказался внутри фрейма. Можно отключить защиту для каждого просмотра или настроить точное значение отправленного заголовка.

## **4. Информация, необходимая для установки и эксплуатации программного обеспечения**

ПО работает по принципу тонкого клиента. Для работы с Системой необходимо устройство с функциями доступа в Интернет (персональный компьютер, планшет, телефон и пр.). Программное обеспечение, необходимое для работы пользователя с Системой, включает в себя операционную систему с графическим интерфейсом и браузер (веб-обозреватель). На сегодняшний день, браузеры, как правило, изначально установлены в операционной системе.

### **4.1 Работа системы**

- Пользователь проходит аутентификацию в системе с помощью JWT. Система разрешает пользователю доступ к нескольким приложениям, пройдя аутентификацию только один раз;
- Пользователю выдается токен;
- Пользователю выдаются RBAC права;
- Данные пользователя хранятся в БД user storage под управлением СУБД PostgreSQL;
- Пользователь начинает работу в Системе согласно выделенной ролевой модели (лингвист, менеджер и так далее в зависимости от требований заказчика);
- В дальнейшем происходит обращение к базе данных для авторизации пользователя в Системе. Компилируется код авторизации, реализованный на языке программирования Python. Информация об авторизованном пользователе передается в кэш-память браузера

(хранящиеся в браузере на компьютере пользователя временные файлы: логин, пароль, электронная почта) ответом на запрос с помощью nginx.

#### **4.1 Требования к программному обеспечению ПК пользователя**

Требования к ПО ПК пользователя:

- Windows Vista/7/8/10 или Mac OS X 10.5-10.8;
- Android;
- IOS;
- Ubuntu;
- Доступ к сети Интернет.

#### **4.2 Рекомендуемые требования к серверному обеспечению**

ПрЭВМ «Рукат» является self-hosted решением, что означает, что ПО располагается на арендованном физическом сервере, расположенном в дата-центре.

Также, в зависимости от требований к Системе, ПрЭВМ «Рукат» может использоваться как cloud-native приложение в соответствии с облачными моделями.

Требование к серверному обеспечению (согласно self-hosted решению):

- Установленный веб-сервер nginx версии не ниже 1.1;
- СУБД PostgreSQL;
- Бесперебойное питание;
- Сервер, расположенный в ЦОД.

#### **4.3 Требования к браузерам**

Рекомендуется использовать браузеры последних версий или, как минимум, перечисленные далее:

- Google Chrome версии 46.0 и выше;
- Mozilla Firefox версии 41.0 и выше;

- Internet Explorer версии 9 и выше;
- Microsoft Edge;
- Safari (для Mac OS X) версии 7.0 и выше;
- YaBrowser версии 15.9 или выше;
- Opera версии 32 или выше.

## **5. Заключение**

Программное обеспечение ПрЭВМ «РуКат» реализовано в виде веб-приложения и представляет собой полуавтоматический переводчик текстов построенный на системной обработке текста любого объема в общепринятых форматах.

Система имеет свое собственное ядро «РуКат» с неизменным основным функционалом и архитектурным стилем.

Frontend частью ПО может выступать как готовое решение от ПрЭВМ «РуКат», так и разработанное специально под требования заказчика с возможной доработкой backend части веб-приложения.

В Систему встроена защита конфиденциальных данных пользователей.

ПрЭВМ «РуКат» имеет гибкую систему и подходит для решения большинства лингвистических проблем.