

Сервис для распознавания документов

Лабораторная работа 1

Техническое задание к проекту

Написать краткое техническое задание к проекту. Описать только требования к основной функциональности системы. По этим требованиям вы далее будете строить систему.

Функциональные требования

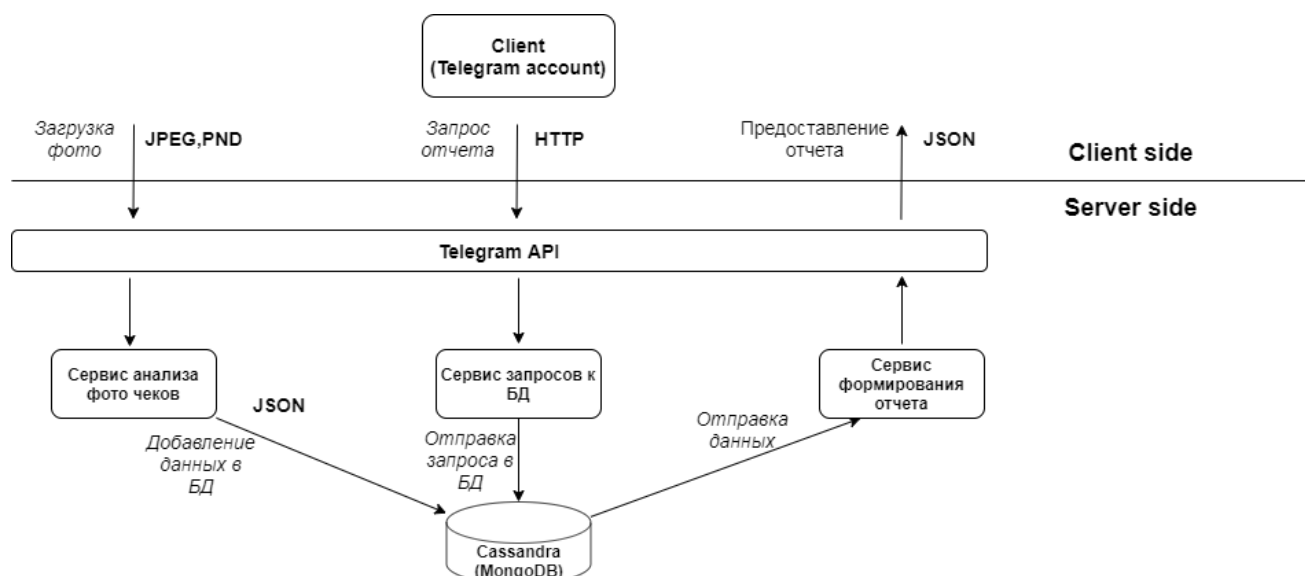
Описание требования
1. Система предоставляет пользователю загружать фотографию чека в телеграмм-бот.
2. Система распознает и сохраняет следующие данные с чека: <ul style="list-style-type: none">• дата и время покупки;• наименование товара и его количество;• общая сумма покупки;• сумма покупки отдельного товара.
3. Система сохраняет результат в базу данных в текстовом виде.
4. Система позволяет пользователю видеть статистику по покупкам за некоторый промежуток времени. Примечание: конкретный список возможных статистик будет определен на фазе дизайна.
5. Система сохраняет в базе данных текстовые данные чека с привязкой к ID пользователя.

Требования к нагрузке

1. Средняя нагрузка ~100 запросов с клиентской части в минуту. Один запрос с клиента соответствует одному запросу на обработку данных, сохранение результатов обработки в БД и одной операции помещения значения в in-темогу кэш.
2. Пиковая нагрузка ~1000 запросов с клиента в минуту. Длительность нагрузки около получаса. А также, промежуток в ~10000 запросов в минуту, длительностью в три минуты.
3. Latency при запросах с клиента не более 3 сек. Запись в БД и другие серверные процессы, инициированные клиентским запросом, происходят асинхронно и не заставляют клиента ждать.

4. Downtime системы не более минуты в неделю.
5. Запрос с клиента обрабатывает одну фотографию, объемом не более 1 мб, сохраняет обработанный результат - текстовую информацию, объемом не более 2 мб. Формат фотографий JPEG, PNG.

Архитектура проекта и технологии



Ссылка для редактирования и детального просмотра:
<https://drive.google.com/file/d/13rjow45xolEmnRi5OWvaWzZ2Xmge-7YS/view?usp=sharing>

1. Выбрать технологии для проекта. Кратко обосновать свой выбор.

Tesseract OCR	Технология для распознавания текста на изображении. Будем использовать библиотеку pytesseract .
Библиотеки Python	OpenCV - библиотека для работы с изображениями
Библиотека для взаимодействия с Telegram API	Библиотека telebot. Библиотека, которая позволяет достаточно просто реализовать своего бота на python

Оценка трудоемкости проекта

Name	Duration	Start	Finish	Predecessors	Resources	Custom 1
Распознавание чека	4 months	01/09/2020	31/12/2020			

<u>Milestones</u>						
Analysis and planning complete	15 days	01/09/2020	30/09/2020		Project Manager	
Design complete	15 days	01/10/2020	31/10/2020		Project Manager	
Development completed	1 month	01/11/2020	30/11/2020		Project Manager	
Stabilization completed	15 days	01/12/2020	15/12/2020		Project Manager	
Deployment completed	15 days	16/12/2020	31/12/2020		Project Manager	
<u>Builds Schedule</u>	5days					
Beta	0day					
Release	0day					
Project management	1day	01/09/2020	01/09/2020		Project Manager[25%]	
Team meetings and status reports	1day	02/09/2020	02/09/2020		Project Manager[25%]	
<u>Analysis and planning</u>	15 days	01/02/2017	01/04/2017			
<i>Functional Specification</i>	2days	01/02/2017	01/03/2017			
Requirements clarification	1day	01/02/2017	01/02/2017		Project Manager	
Functional specification writing	1day	01/03/2017	01/03/2017		Project Manager	
Approval with customer	0day	01/03/2017	01/03/2017		Project Manager, Customer	
<i>Estimation</i>	2 days	01/04/2017	01/04/2017			
Project plan update	1day	01/04/2017	01/04/2017		Project Manager	
Approval with customer	0day	01/04/2017	01/04/2017		Project Manager, Customer	
<u>Design</u>	2days	01/05/2017	01/06/2017	13		
<i>Technical design specification</i>	1day	01/05/2017	01/05/2017			
Technical specification writing	1day	01/05/2017	01/05/2017		BA Lead	
Approval with customer	0day	01/05/2017	01/05/2017	23	BA Team, Customer	
<i>GUI Drawing</i>	2days	01/05/2017	01/06/2017			
UI requirements clarification	1day	01/05/2017	01/05/2017		UI Lead	
UI mock-ups drawing	1day	01/06/2017	01/06/2017	26	UI Lead	

Approval with customer	0day	01/06/2017	01/06/2017		UI Team,Customer	
<i>Test plan & test cases</i>	2days	01/05/2017	01/06/2017			
Test plan document writing	1day	01/05/2017	01/05/2017		QA Lead	
Test cases document writing	1day	01/06/2017	01/06/2017		QA Lead	
Approval with customer	0day	01/06/2017	01/06/2017		QA Team,Customer	