



Национальный исследовательский
университет
«Высшая школа экономики»

Факультет экономики,
менеджмента и
бизнес-информатики

г.Пермь
2022

ПОИСК ПУБЛИКАЦИЙ О ЧЕЛОВЕКЕ ПО ЕГО ФОТОГРАФИИ

Командный проект

Команда: Атнагулова Ю., Лузина К., Невлер Т., Тарасова В., Фокина В

Руководитель: преподаватель кафедры информационных технологий в бизнесе
Лебедев В.В



План доклада

- Назначение проекта, цель и задачи проекта
- Роли и индивидуальные задачи каждого участника
- Результаты проекта



Постановка задачи

- Требуется разработать программу, с помощью которой можно получить отчеты о публикациях в СМИ
- Признак для поиска - фотография человека
- Перечень СМИ ограничивается конкретным перечнем
- Использование бесплатного программного обеспечения



Цель

Реализация программного продукта, позволяющего выполнять поиск публикаций о человеке по его фотографии в определенном списке СМИ

Задачи

Организовать команду, создать план реализации проекта

Провести анализ, составить требования

Выполнить проектирование архитектуры системы

Разработать систему на программном коде

Провести отладку и тестирование программного продукта



Атнагулова Юлия

- Менеджер проекта
- Аналитик



Организация процесса реализации проекта

Менеджер проекта

Работа с командой

- Распределение ролей
- Организация пространства для взаимодействия команды:
 - Площадки для хранения наработок, документов
 - Площадки для визуализации процессов
 - Площадки для общения
- Контроль работы команды

Работа с заказчиком

- Установление связи с помощью электронной почты
- Проведение консультаций в Zoom
- Доступ к пространству взаимодействия команды
- Демонстрация работы приложения

Работа с проектом

- Составление календарного плана проекта
- Контроль сроков выполнения задач



Существующие решения

	Яндекс Картинки	Гугл Картинки	TinEye	Search Face	search4face s
Доступность (бесплатность)	+	+	+	—	+
Поиск по лицам (распознавание лиц)	—	—	—	+	+
Поиск по неограниченным ресурсам	+	+	+	—	—



Анализ инструментов для распознавания лиц



1. Библиотека Face Recognition позволяет распознавать лица и работать с ними с помощью Python
2. Microsoft Azure Face API извлекает ряд атрибутов, связанных с лицом. Недостаток - платная
3. ПО Betaface сравнивает лица, но не позволяет распознать лицо на фотографии
4. FaceX позволяет провести сравнение лиц на двух изображениях, но у него нет бесплатных тарифов



Тарасова Валентина

- Аналитик
- Frontend-разработчик



Разработка требований





Разработка интерфейса

Используемые инструменты:

- JavaScript
- ReactJS
- React Bootstrap
- React-router-dom



Разработка интерфейса

Поиск статей по фото



Изменить фото

Найти статьи

Произошла ошибка.

Не удалось распознать человека на фотографии, попробуйте загрузить другое фото

Поиск статей по фото

ID Вашего отчета 2



Изменить фото

Найти статьи

Название	Ссылка на статью	Дата публикации	Описание
Закон формирует ответственность	https://transportrussia.ru/razdely/avtomobilnyi-transport/4600-zakon-formiruet-otvetstvennost.html	15.03.2022	Новости, обзоры, аналитика в сфере пассажирских и грузовых автомобильных перевозок. Региональные и международные перевозки.
Максим Решетников: «Один из приоритетов нашей работы – поддержка старшего поколения»	https://gazeta-perm.ru/article/maksim-reshetnikov-odin-iz-prioritetov-nashey-raboty-podderzhka-starshego-pokoleniya	18.10.2018	На очередном заседании Законодательного Собрания перед депутатами с ежегодным посланием выступил губернатор Пермского края &Максим Решетников.По словам главы региона, один из приоритетов

1

Сортировать

Скачать отчет

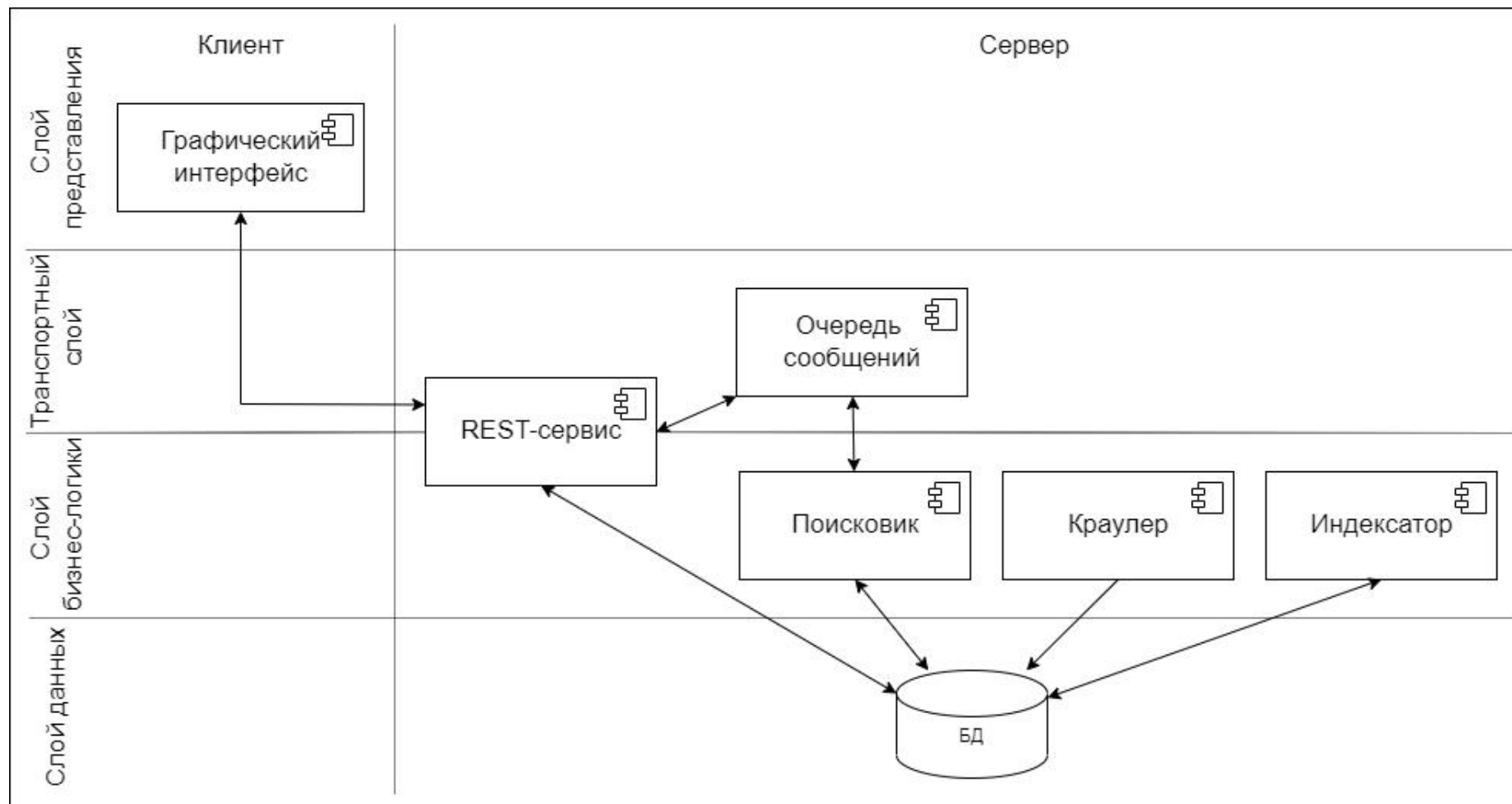


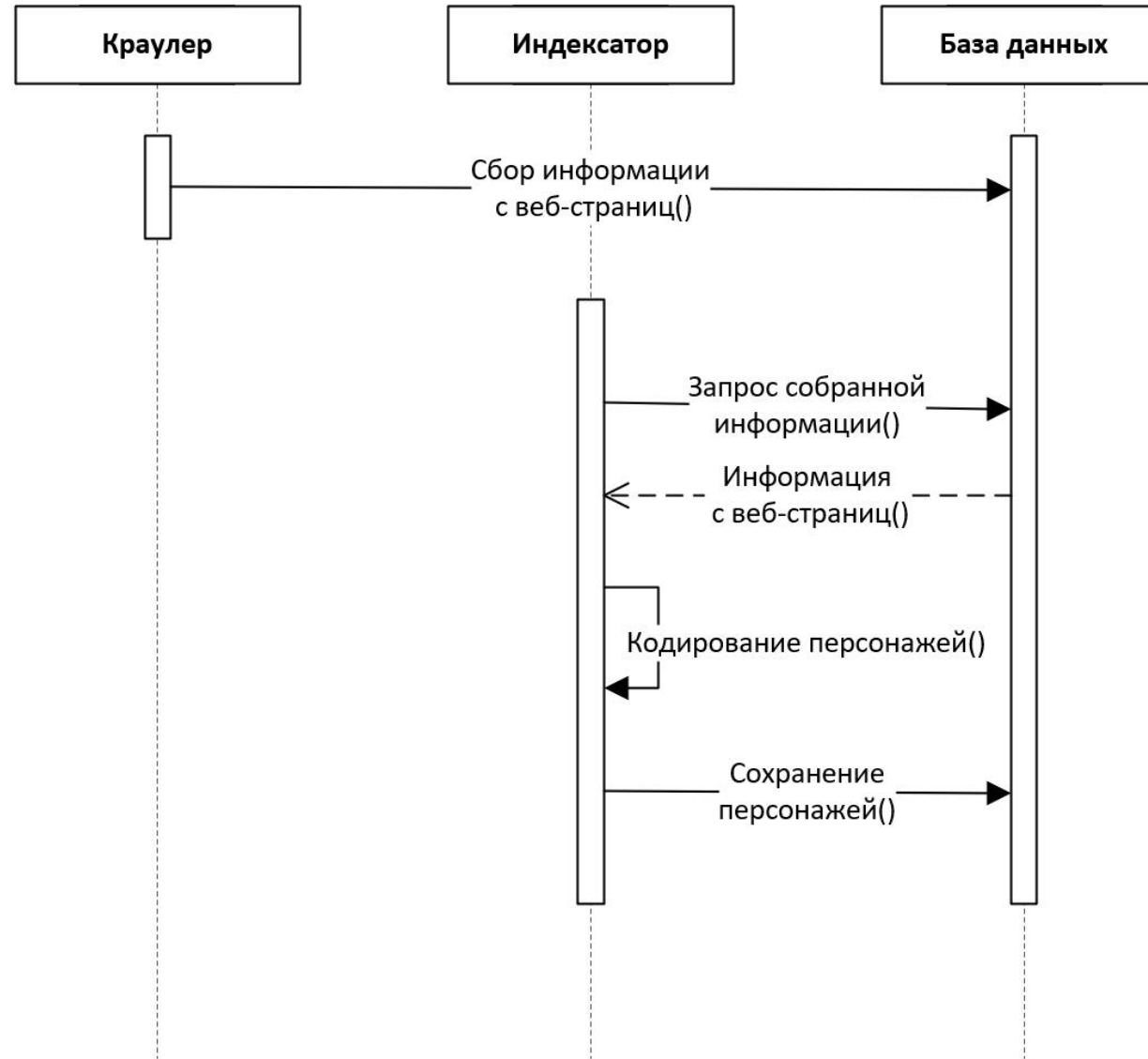
Фокина Виктория

- Архитектор
- Backend-разработчик



Проектирование архитектуры



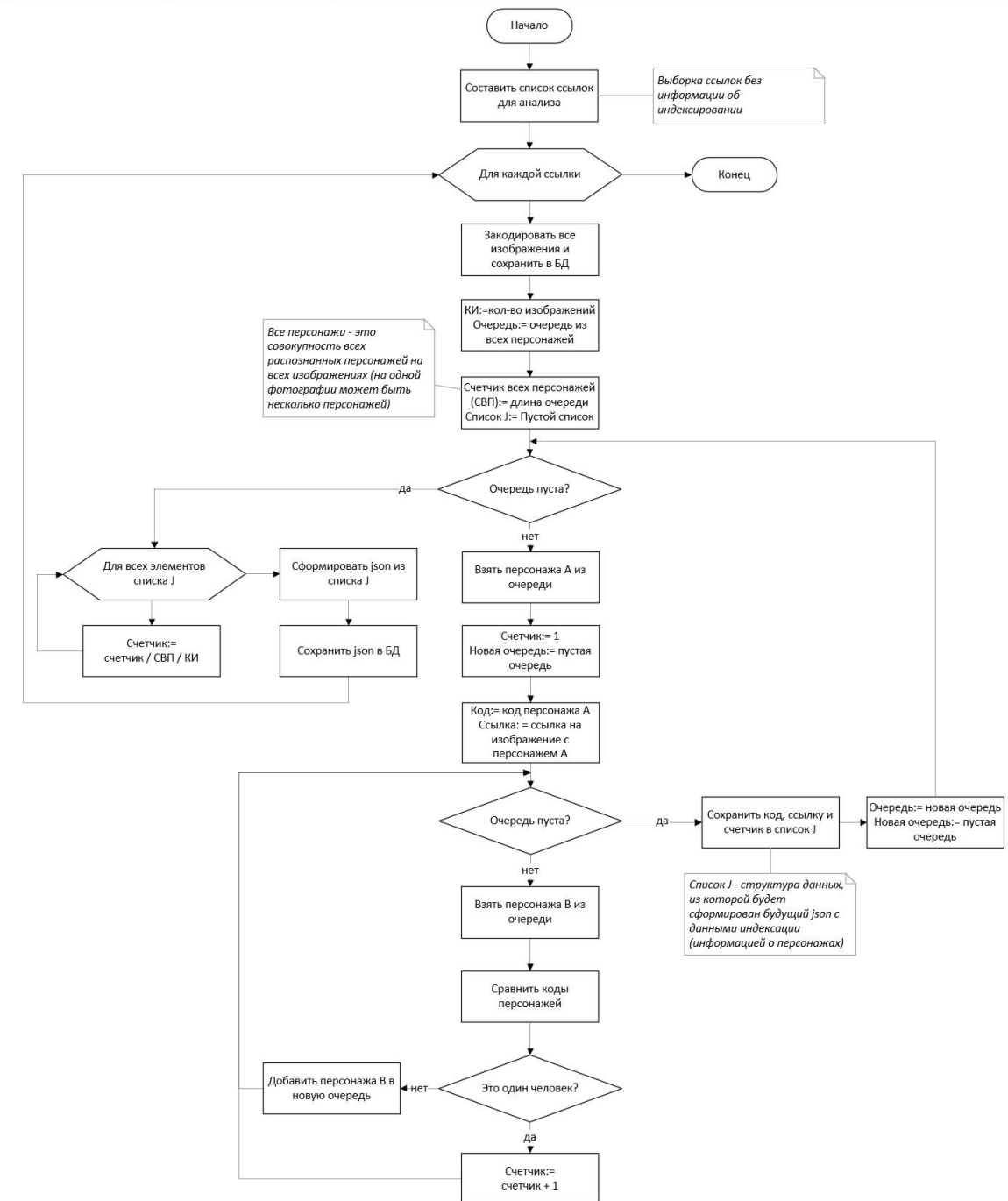


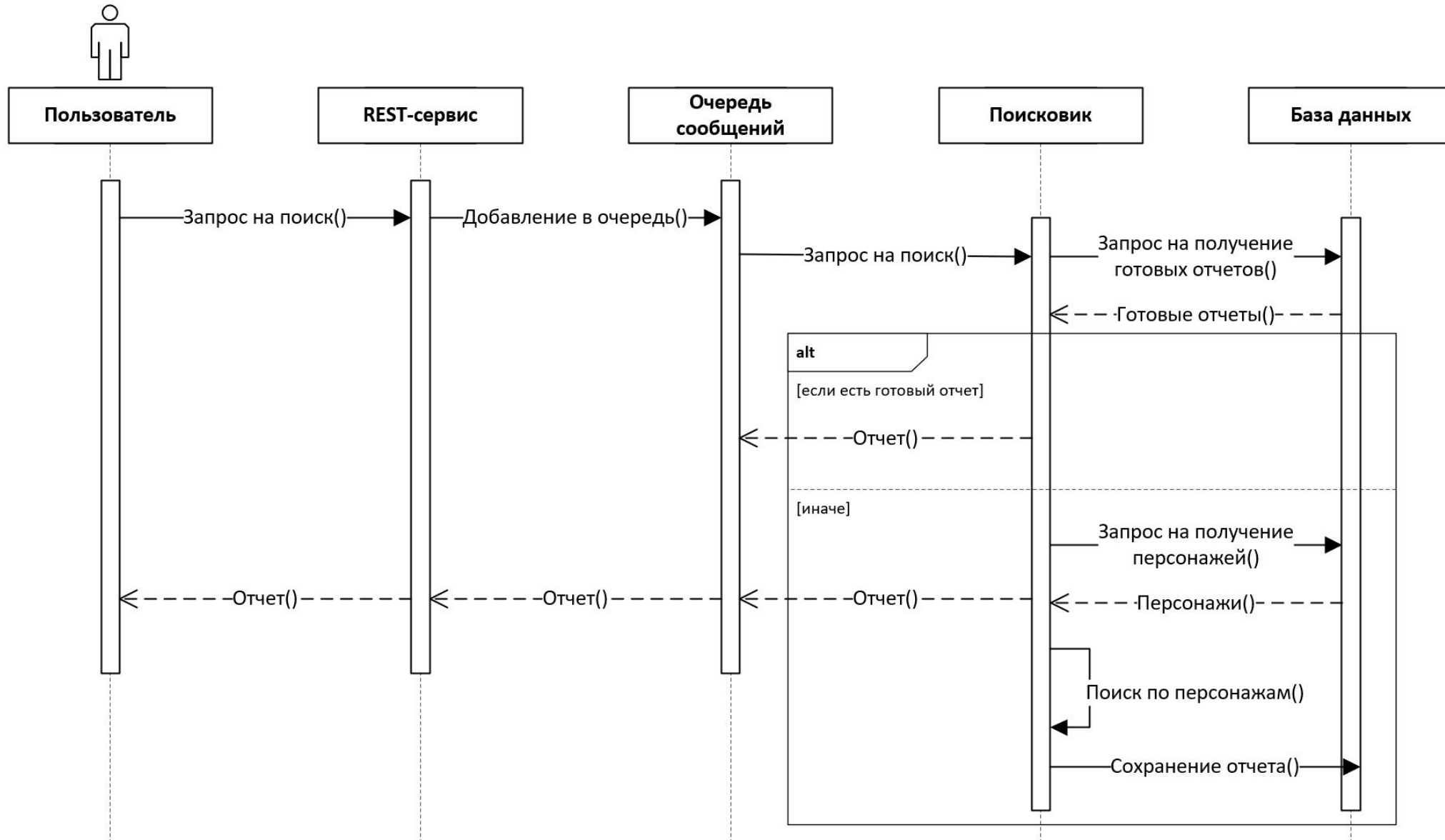


Поиск публикаций о человеке по его фотографии

Алгоритм индексации

- Сокращает время поиска
- Рассчитывает релевантность результатов







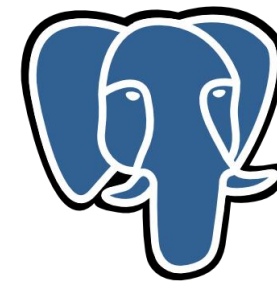
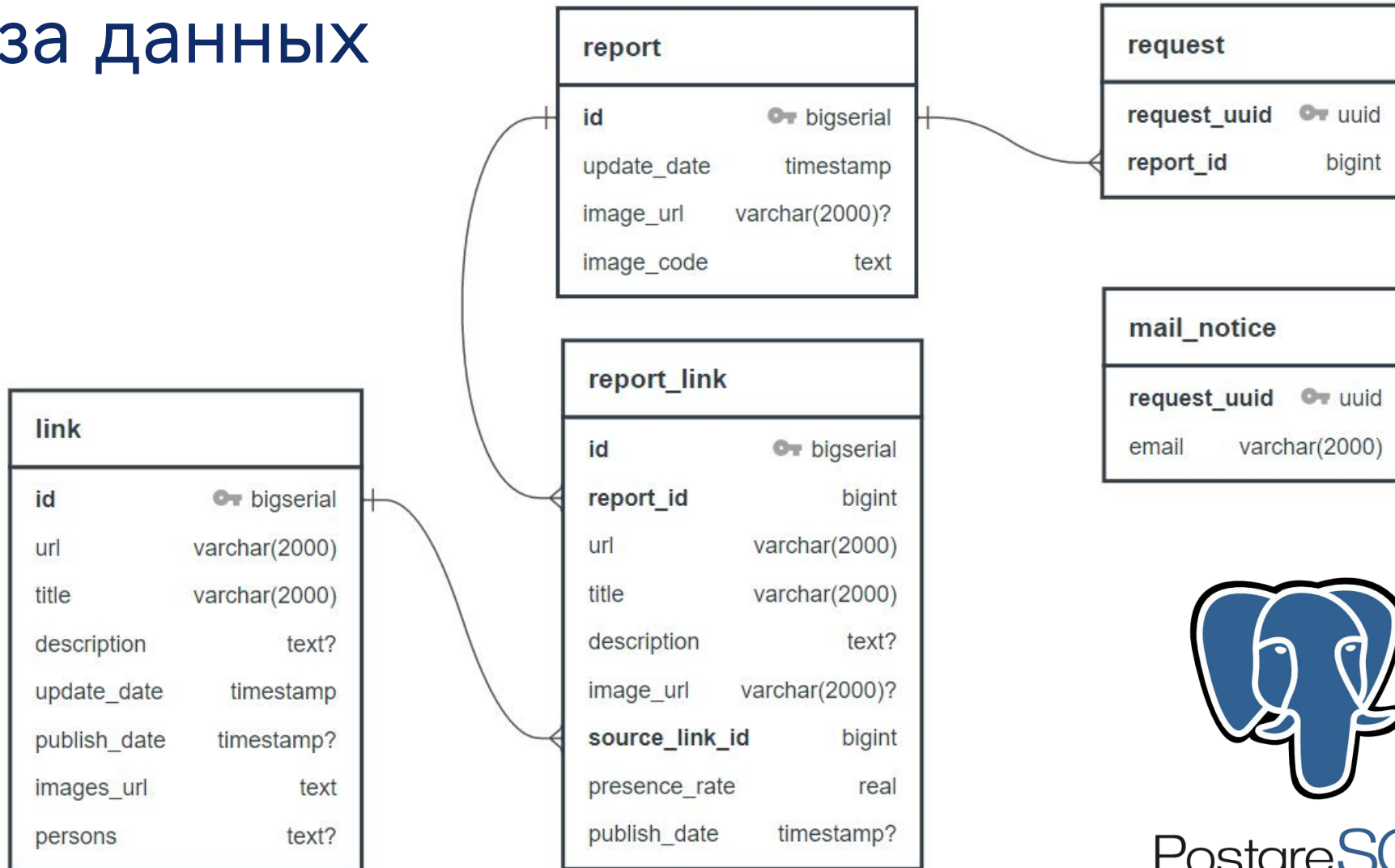
Выбор инструментов для реализации

- Языки: Python, Java, Kotlin
- Face Recognition
- Selenium, Scrapy и др.
- Spring Framework
- RabbitMQ





База данных



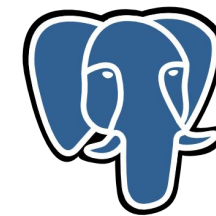
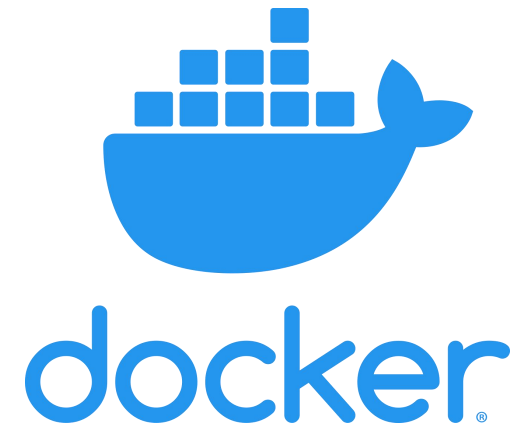
PostgreSQL



Развертывание

Контейнеры в Docker:

- База данных
- Брокер очередей
- Компоненты системы



PostgreSQL





Разработка краулера

Для www.perm.kp.ru

Использовались:

- Python
- Scrappy
- Selenium

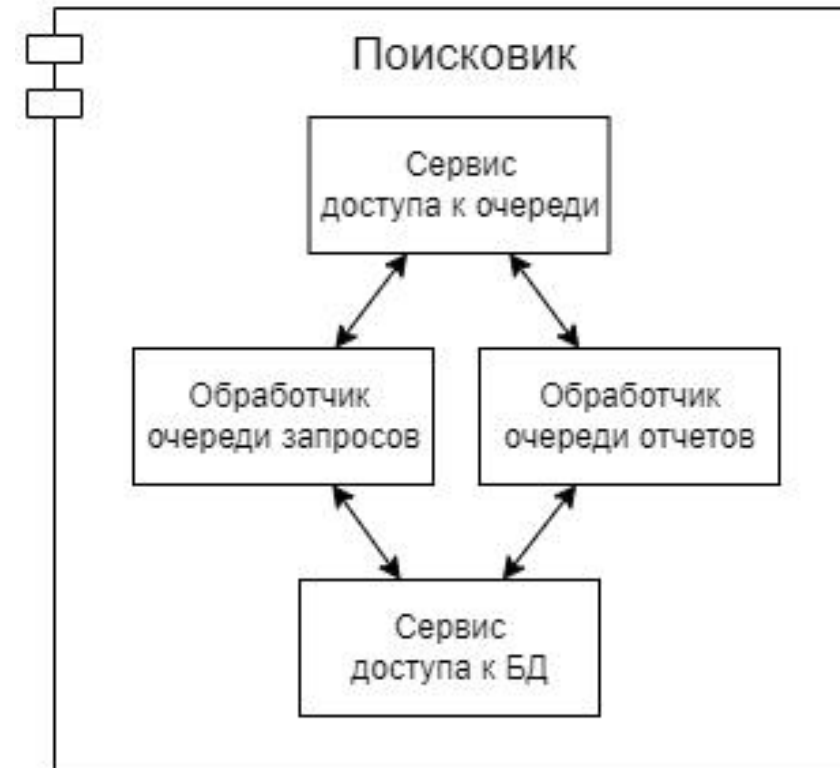




Разработка «Поисковика»

Использовались:

- Python
- Face Recognition
- Pika для обработки очереди
- Psycopg2 для доступа к БД

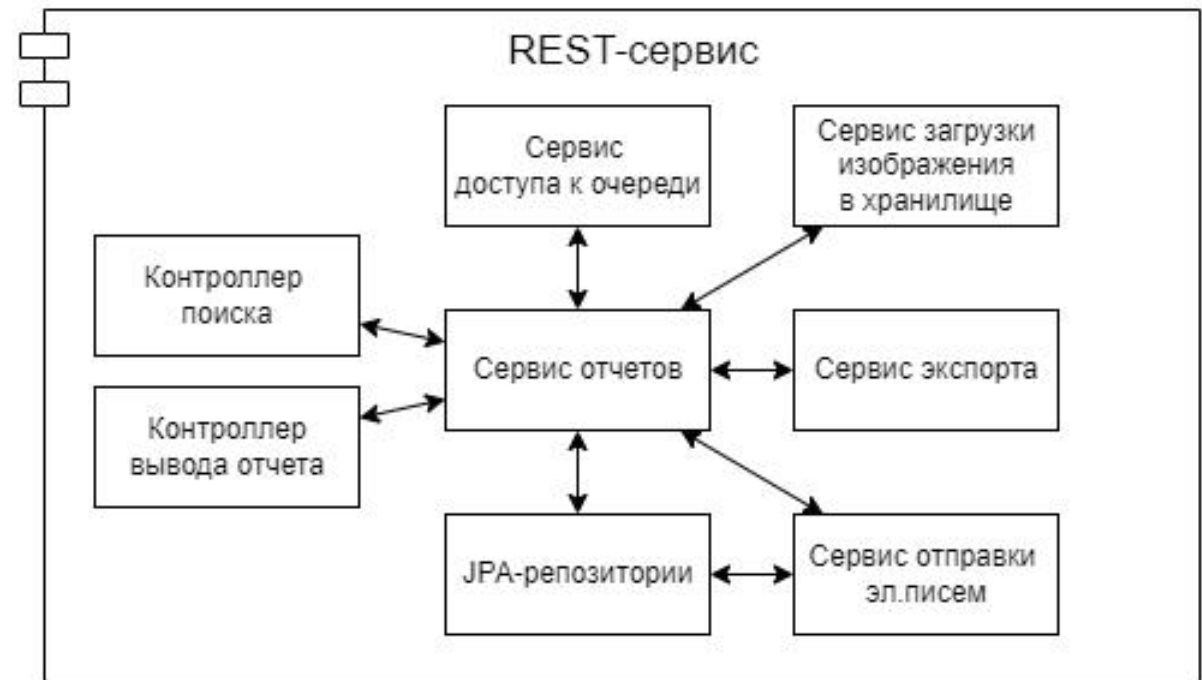




Разработка «REST-сервиса»

Использовались:

- Java, Kotlin
- Spring Framework
- Swagger
- Requests, poi-ooxml, postgresql и др.





Разработка «REST-сервиса»

report-resource Endpoints for getting search results

GET `/api/report/{id}` Get search results by the report id

GET `/api/report/{id}/download` Download search results as Excel file by the report id

search-resource Endpoints for finding web pages related to the person in the photo

POST `/api/search` Send a request to search for web pages related to the person in the photo

GET `/api/search/{uuid}` Get an id of a report with search results by the request uuid (waiting for a response from the queue)

POST `/api/search/email` Send a user email to notify the user of search results when the results are ready

GET `/api/search/request-count` Get the number of requests in the queue



Лузина Кристина

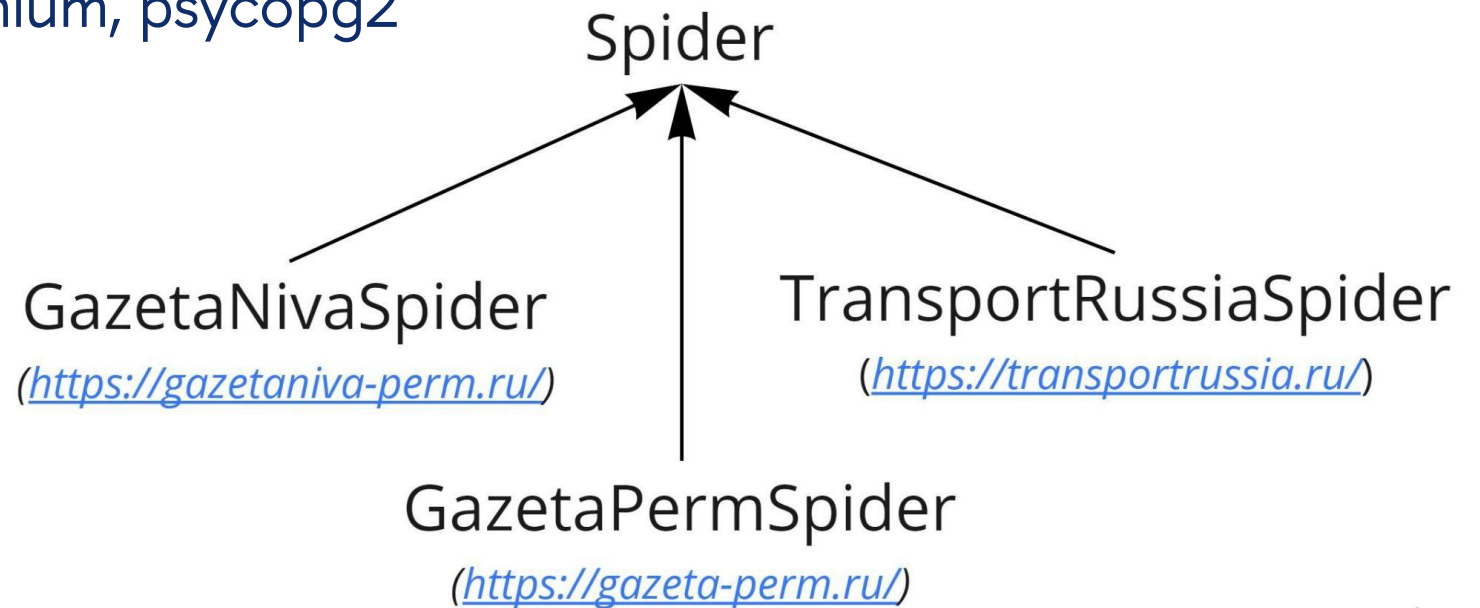
- Backend-разработчик



Разработка краулеров

Используемые инструменты:

- Язык программирования Python
- Библиотеки Selenium, psycorg2





Разработка «Индексатора»

Используемые инструменты:

- Язык программирования Python
- Библиотеки `face_recognition`, `psycopg2`, `urllib`





Невлер Татьяна

- Тестировщик



Тестирование

- Было составлено 44 тестовых сценария на 47 требований к системе.
- Была построена матрица покрытия требований тестами.
- Тестовое покрытие составляет 100%, т.е. все требования были покрыты тестовыми сценариями.
- В ходе тестирования было обнаружено 11 ошибок, которые затем были исправлены.

	Отчет №6		Дата составления:	25.03.2022 12:36
№	Ссылка	Заголовок	Описание	Дата публикации
1	https://transportrussia.ru/razdely/ofitsialno/8554-my-r	Мы можем. Мы сделаем!	Решения Госдумы РФ, Правительства России, Минтранса и других официальных органов, касающиеся развития и функционирования транспортной отрасли. Оперативная информация.	09.03.2022 23:12
2	https://transportrussia.ru/razdely/vozdushnyj-transport	Поздравление Президента России Владимира Путина работникам и ветеранам гражданской авиации	Новости, обзоры, аналитика в сфере пассажирских и грузовых авиаперевозок. Региональные и международные перевозки. Субсидирование авиамаршрутов.	09.02.2022 20:38
3	https://transportrussia.ru/razdely/zheleznodorozhnyj-transport	Президент России утвердил перечень поручений по некоторым вопросам развития транспорта	Новости, обзоры, аналитика в сфере пассажирских и грузовых перевозок по железной дороге. Региональные и международные перевозки. Строительство магистралей. Обновление подвижного состава.	30.01.2022 19:12
4	https://transportrussia.ru/razdely/transportnaya-infrastruktura	Держать руку на пульсе	Аэропорты, мосты и мостовые переходы, станции метро, автостанции и автовокзалы, дорожные развязки. Инженерные решения и стратегические планы. Современные технологии в действии.	26.12.2021 22:01
5	https://transportrussia.ru/razdely/gorodskoj-passazhirskij-transport	Москва рассказала о планах транспортного развития	Газета «Транспорт России» - ведущее издание транспортного комплекса страны, официальный печатный орган Министерства транспорта Российской Федерации.	22.12.2021 22:17
6	https://transportrussia.ru/razdely/vozdushnyj-transport	Крепкие крылья Аэрофлота	Новости, обзоры, аналитика в сфере пассажирских и грузовых авиаперевозок. Региональные и международные перевозки. Субсидирование авиамаршрутов.	12.12.2021 21:05
7	https://gazeta-perm.ru/article/vse-pod-kontrolem-skriningovoe-obслеdovanie-na-ive	Все под контролем: скрининговое обследование на Иве	Для людей старшего поколения Мотовилихинского района регулярно проводятся скрининговые обследования. Сегодня осмотр «в шаговой доступности от дома» прошли жители микрорайона Ива	09.12.2021 02:39
8	https://transportrussia.ru/razdely/avtomobilnye-dorogi	Президент России Владимир Путин:	Национальные проекты — обзоры, аналитика. Новости регионов. Строительство, ремонт и содержание дорог. Статистика и финансы. Новейшие достижения науки и техники в дорожной отрасли.	15.10.2021 00:06
9	https://transportrussia.ru/razdely/vozdushnyj-transport	Дальневосточный прорыв	Новости, обзоры, аналитика в сфере пассажирских и грузовых авиаперевозок. Региональные и международные перевозки. Субсидирование авиамаршрутов.	09.09.2021 00:03
10	https://transportrussia.ru/razdely/vozdushnyj-transport	Президент России Владимир Путин призвал изабвить малые аэропорты от избыточных требований	Новости, обзоры, аналитика в сфере пассажирских и грузовых авиаперевозок. Региональные и международные перевозки. Субсидирование авиамаршрутов.	02.09.2021 02:57
11	https://transportrussia.ru/razdely/morskoj-transport/7	Президент России посетил Дальневосточный морской тренажерный центр	Новости, обзоры, аналитика в сфере перевозок морским транспортом. Развитие Арктики. Новые технологии в судостроении.	02.09.2021 02:07
12	https://transportrussia.ru/razdely/avtomobilnye-dorogi	В Росавтодоре подвели итоги кассового исполнения бюджета за семь месяцев	Национальные проекты — обзоры, аналитика. Новости регионов. Строительство, ремонт и содержание дорог. Статистика и финансы.	17.08.2021 22:18



Результаты

- Результатом выполнения проекта является разработанная **программная система** для поиска публикаций о человеке по его фотографии
- Разработанная **документация** к системе
- Система может быть **интегрирована** в систему заказчика или использоваться как самостоятельная
- Система имеет возможности **масштабирования**

Спасибо за внимание!

ПОИСК ПУБЛИКАЦИЙ О ЧЕЛОВЕКЕ ПО ЕГО ФОТОГРАФИИ



Атнагулова Ю., Лузина К., Невлер Т., Тарасова В., Фокина В

Пермь
2022

