



ГРИНАТОМ Хакатон

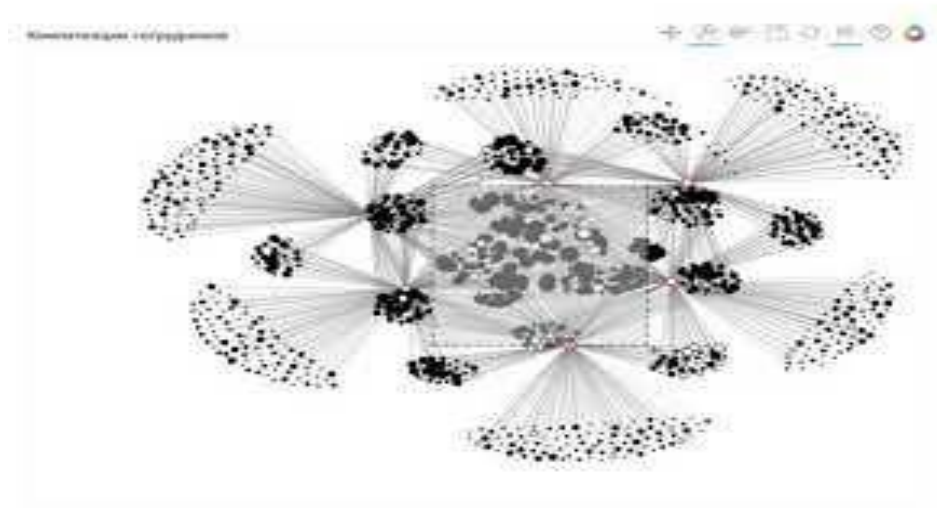
Трек: Машинное обучение



SmartHR

Команда: **Сителикс** (г. Саров)






<https://youtu.be/XLmeXgl5twk>

Web-интерфейс

Проект:

Сотрудник:

Абалдуева Наида Иосиповна



ФИО: Абалдуева Наида Иосиповна

Возраст: 27

Должность: ведущий научный

Рекомендации отправить по электронной почте

lovebirds1856@rosatom.ru

Подобрать экспертов

Увольнение

Рассчитать вероятность увольнения

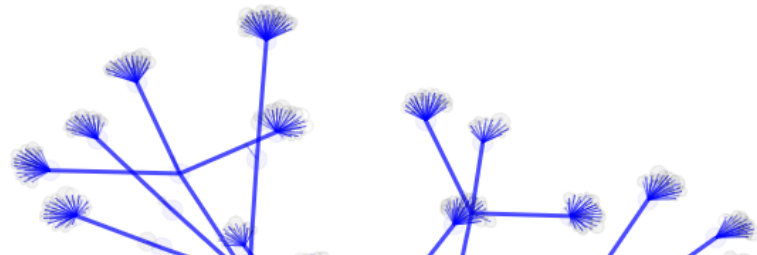
Графы взаимодействия

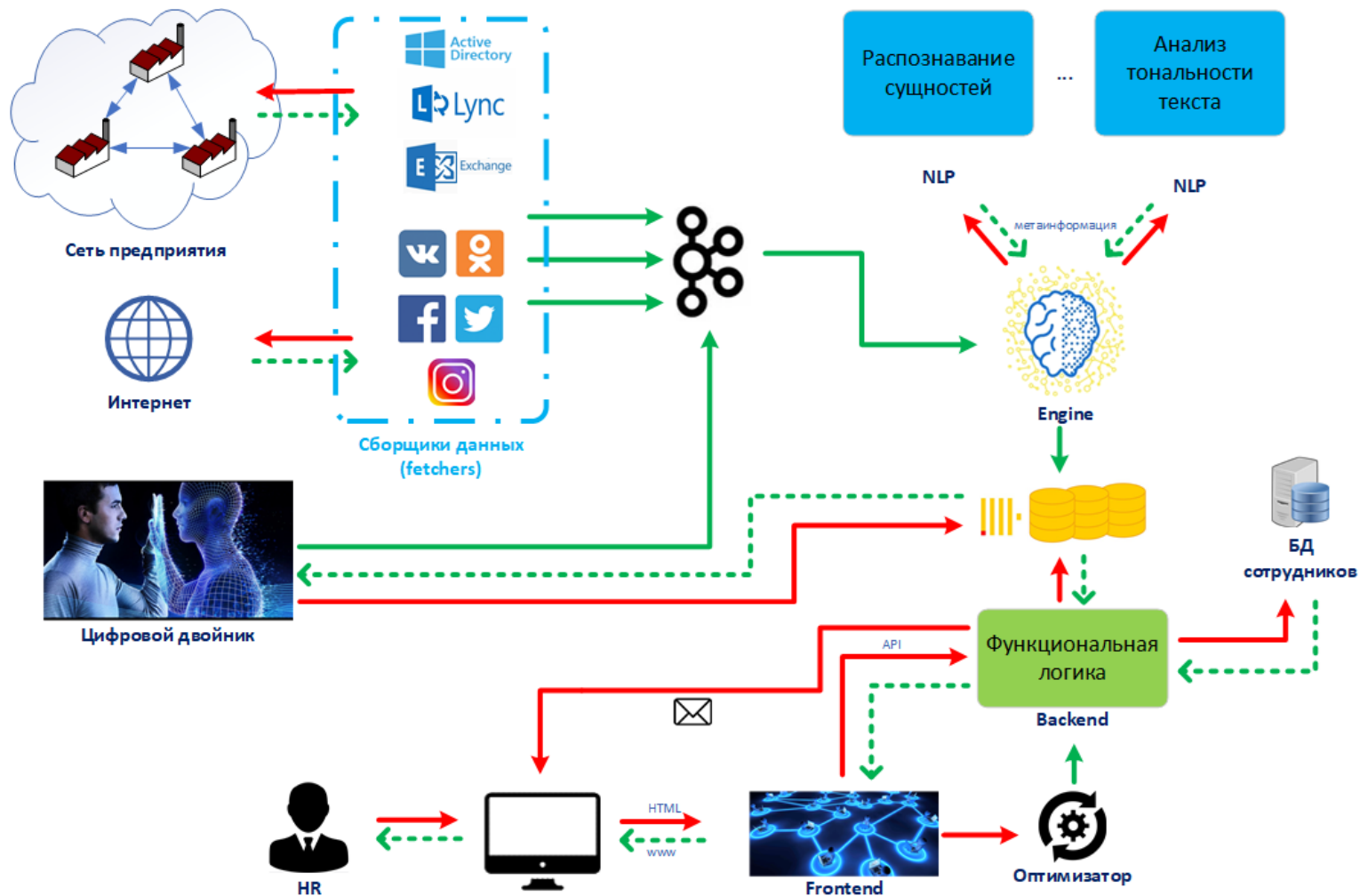


staff-unit



Сотрудник-подразделение





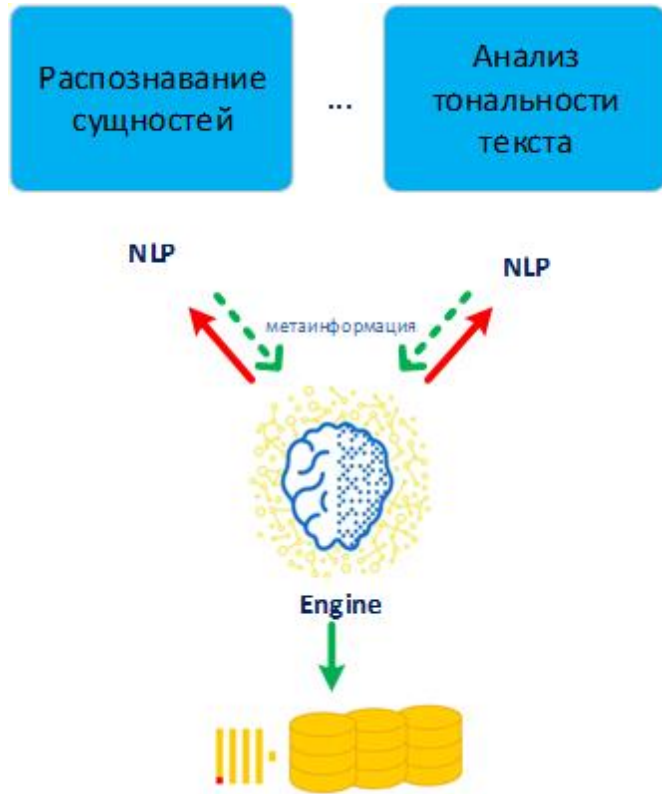
Принцип работы



Для каждого источника данных реализуется отдельный сервис по сбору данных (fetcher), который функционирует в непрерывном режиме для выявления новых данных.

Fetcher полученные новые результаты отправляет брокеру сообщений Kafka, с помощью которого осуществляется дальнейшая гарантированная доставка сообщений потребителю.

Принцип работы



Производится NLP-анализ текста писем и сообщений, анализ тональности текста, распознавание именованных сущностей, анализ графов взаимодействия с коллегами.

Полученная метаинформация с указанием источника данных (реальное предприятие или его цифровой двойник) отправляется на долговременное хранение в базу данных.

Обработка естественного языка

Введите текст для анализа

- Андрей, привет! Как дела? Что нового?
- Завтра едем в Москву для презентации проекта. Ура!
- Я ненавижу ездить в поездах РЖД, просто замучился!

- Андрей, привет! Как дела? Что нового?

Разговорная речь

	Слово	Тип
2	Андрей	персона

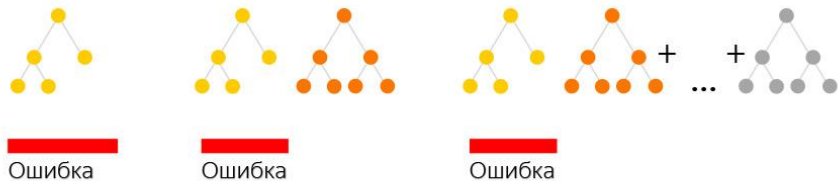
- Завтра едем в Москву для презентации проекта. Ура!

Положительная тональность

	Слово	Тип
5	Москву	место

Принцип работы

Градиентный бустинг



Yandex
CatBoost

Вероятность увольнения сотрудника рассчитывается с использованием алгоритма градиентного бустинга над решающими деревьями.

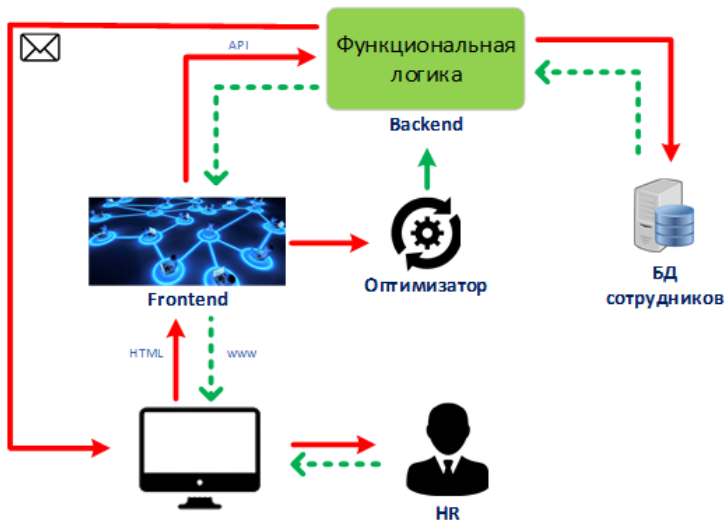
Решается задача логистической регрессии над двумя классами:

- 0 – сотрудник не уволится
- 1 – сотрудник уволится

Предсказание увольнения: выявленные важные признаки

Признак	Важность
Ипотека	29,7
Дата последнего повышения	25,3
Должность	12,4
Ранний уход с работы за последний месяц	12,3
Опоздание на работу за последний месяц	9,8
Уровень знаний JavaScript	3,6
Уровень знаний Python	1,1
Отгулы за последние 3 месяца	0,8

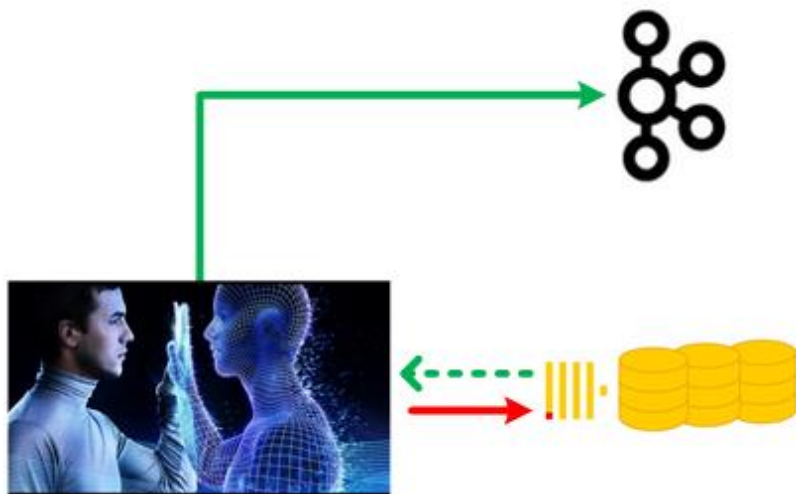
Принцип работы



HR взаимодействует с web интерфейсом с помощью браузера. Данные о сотрудниках-экспертах доставляется посредством E-Mail.

Функциональная логика взаимодействует с БД, где хранится обработанный цифровой след и тень сотрудника.

Перспективы развития



Развитие в направлении
цифрового двойника,
включающего возможность
моделирования поведения
сотрудников предприятия.

Используя технологии RNN, основываясь на данных,
собранных во время функционирования всей системы
в целом, со временем будут формироваться **точные**
«цифровые копии» сотрудников предприятия.

Killer-features

“Цифровой двойник” сотрудников – математическая модель наиболее полно описывающая взаимодействие между сотрудниками.

Возможность масштабирования под компании любого размера.

Что сделано?

Разработана архитектура сервиса

Реализованы прототипы ключевых компонентов.



Почему это нужно компании?

Кадры решают все! Уход ключевого сотрудника может отрицательно повлиять на ход выполнения проектов, вовлеченность других сотрудников и компанию в целом.

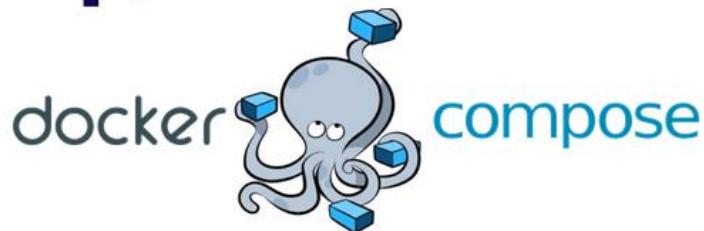
Правильный подбор кадров – ключ к эффективному развитию проектов компании.



Стек технологий



Streamlit



Планы на будущее

Реализовать сервисы по сбору данных из корпоративной сети и сети Интернет (fetchers), которые будут функционировать в непрерывном режиме для сбора и накопления новых данных.

Доработать "цифровой двойник" сотрудника.

Авторизация пользователей для разграничения доступа к функционалу.

Команда

Дмитрий Куликов

Компетенции:

Data Science

Backend

Project management

Product development

Public speaking

DataBase



Айдар Сайфуллин

Компетенции:

Frontend

Backend

UI/UX

Public speaking

DevOps

Благодарим за внимание



<https://github.com/jubbon/greenatom-ml-2019-11>

Дмитрий Куликов (+79063660240, kulikov@sarov.info)

Айдар Сайфуллин (+79053757057, aidar.saifouline@gmail.com)