

## Тестовое задание

### Автомаппер данных

Этапы:

#### **I** (веб-приложение на Django)

1. Микросервис выполняет чтение текстовых файлов (txt/csv) и определяет их структуру.

Включает следующие функции:

- чтение выбранного локального файла или файла, доступного по url
- идентификацию типов данных в тексте/таблице по определенным шаблонам. Имеет предустановленные шаблоны для следующих типов: дата, телефон, эл. почта, адрес, ip-адрес, ИНН.
- автоматическое определение категорий документа. Включает в себя два класса категорий: «Договор», «Публикация». Каждый класс имеет не менее 2 тегов, по обнаружению которых определяется категория.

2. Для указанных выше функций создать пользовательскую веб-страницу, которая предусматривает:

1. выбор расположения локального файла для чтения
2. окно для отображения структуры файла с указанием в каких колонках были распознаны предустановленные шаблоны типов. Возможность ручного изменения предложенных машиной типов и категорий
3. кнопка сохранения в Реестр (локальный файл) распознанной структуры файла с указанием его расположения, даты обработки и размера
4. окно для отображения содержимого Реестра на веб-странице с возможностью изменения всех полей

Для тестирования:

1. создать и наполнить (сгенерировать) текстовый файл с табличными данными (не менее 10 млн строк), содержащими следующие колонки (атрибуты): дата, серийный номер (содержит буквы, цифры), текст, стоимость, контактные данные (эл.майл, телефон), адрес, любые другие поля свободного назначения (не менее двух).

#### **II** (по согласованию)

Развернуть общедоступное веб-приложение с вышеуказанными сервисами для пользовательского тестирования.

#### **III** (по согласованию)

Создать файл сборки контейнера docker для автоматизированного выполнения второго этапа.