

# Полное описание задач конкурса на соискание премий Мэра Москвы «Лидеры цифровой трансформации»

## Задача 1

### Главное контрольное управление города Москвы

*Описание задачи «Формирование интерактивной карты благоустройства города Москвы»:*

1. Обработка исходных данных:
  - a. разобрать и загрузить excel-файлы титульных списков благоустройства на 2020 и 2021 года;
  - b. разобрать и загрузить json реестр объектов дорожного хозяйства (ОДХ);
  - c. пересечь реестр и титульные списки на предмет наличия ОДХ в титульных списках.
2. Визуализация данных на карте:
  - a. встроить карты ЕГИП в своё приложение;  
Данные для встраивания:
    - 1) Ридми по встраиванию  
<http://egiptest.mos.ru/jsapi/docs/#/README?id=%D0%A1%D1%82%D0%B0%D1%80%D1%82>  
и описание объектов  
<http://egiptest.mos.ru/jsapi/docs/#/examples/features/object-editor/object-editor>
    - 2) пример встраивания  
<http://egiptest.mos.ru/jsapi/docs/#/examples/basic/create-map/create-map>
    - 3) Лого/пасс от описания АПИ  
egipapi  
1FdPjl4UjwHMeZ
    - 4) лог пас для доступа к ЕГИПу  
hakaton1  
iN2m03A#
    - 5) Адрес боевого ЕГИПа  
<https://egip.mos.ru/>
  - b. разместить на карте Москвы объекты благоустройства сохраняя геометрию границ;
  - c. реализовать функционал маркировки объектов в зависимости от наличия в титульных списках 2020 и 2021 годов.
3. Распределить задания на обследования инспекторам и обеспечить обработку результата:
  - a. реализовать функционал группировки объектов титульного списка в зависимости от транспортной доступности (пешее обследование, либо обследование с использованием ТС), визуализировать на карте.

- б. реализовать функционал формирования заданий и распределения нагрузки между инспекторским составом в зависимости от количества задействованных инспекторов и их удаленности до объектов (пресет: есть X инспекторов, из которых Y на машине, которые распределены по городу и привязаны к инспекциям).
- с. реализовать возможность нужному инспектору ввести результат обследования: по шкале от 1 до 5 оценку необходимости и применимости благоустраивать территорию, визуализировать на карте.
- 4. Добавить систему отчетности по результатам обследования объектов
  - а. Выгрузить excel-файл титульного списка благоустройства с дополнительной колонкой – оценка сотрудника.
  - б. Создать дашборд начальника ОАТИ и начальников инспекций со статистической информацией.

## **Задача 2**

### **Департамент культуры города Москвы**

*Описание задачи «Разработка рекомендательного сервиса по контенту и активностям учреждений культуры»:*

#### **Общие вводные**

В настоящее время в учреждениях, подведомственных Департаменту культуры города Москвы, используется множество информационных систем, предназначенных для автоматизированного сбора информации. Отсутствие информационного обмена в рамках информационных систем, содержащих сведения о пользователях и потребленном ими "культурном контенте", не позволяют сформировать полноценный "культурный портрет".

#### **Цель**

Обеспечить привлечение новых пользователей и популяризацию услуг, предоставляемых учреждениями, подведомственными Департаменту культуры города Москвы, за счет релевантной рекомендательной системы.

#### **Термины и сокращения**

*БМ* - учреждения библиотечного типа, подведомственные Департаменту культуры города Москвы.

*КЦ* - учреждения культурно-досуговых типов, подведомственные Департаменту культуры города Москвы.

*КДФ* - культурно-досуговое формирование - вид деятельности, реализуемый в КЦ.

*Учреждения* - учреждения, подведомственные Департаменту культуры города Москвы.

*Пользователь* - физическое лицо, однозначно идентифицируемое по базе данных посетителей, которое намерено воспользоваться или воспользовалось услугами КЦ и БМ.

*АБИС* - единая автоматизированная библиотечно-информационная система на базе "MARK-Cloud", предназначенная для ведение библиотечно-информационного обслуживания населения (книговыдача, выдача справок и т.п.) и работы с книжным фондом библиотек.

*ID* - уникальный идентификатор пользователя АБИС.

*ИС БМ* - информационная система "Библиотеки Москвы", обеспечивающая взаимодействие АБИС с внешних информационными системами и сервисами.

*БС* - автоматизированная информационная система "Обеспечение прохода и учет посещаемости учреждений Департамента культуры города Москвы", в рамках которой осуществляется внесение сведений о проводимых активностях (мероприятиях, экскурсиях и т.п.), продажа билетов на них, а также осуществление резервирования помещений под активности в учреждениях.

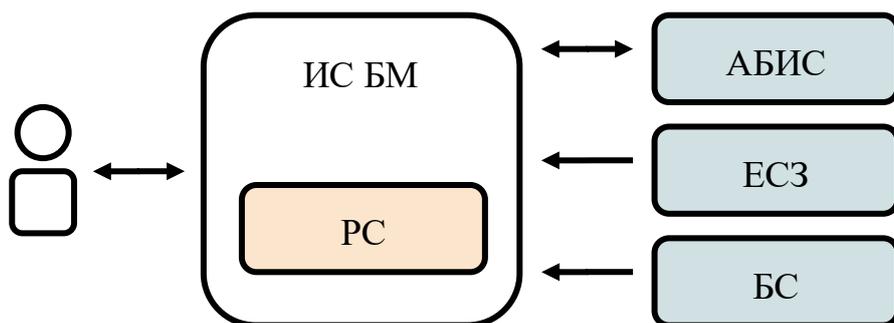
*ЕСЗ* - единый сервис записи, входящий в состав комплексной информационной системы "Государственные услуги в сфере образования в электронном виде",

обеспечивает внесение сведений о функционирующих КДФ, а также осуществление записи в них пользователей КЦ.

*API* - набор готовых классов, процедур, функций, структур и констант, предоставляемых приложением (библиотекой, сервисом) или операционной системой для использования во внешних программных продуктах.

*РС* - рекомендательный сервис, разрабатываемый в рамках хакатона "Лидеры цифровой трансформации".

### Примерная схема информационного взаимодействия



### Примечания к разрабатываемому решению

В рамках разрабатываемого решения должна быть произведена сборка семантического ядра, содержащего сведения по категориям из отраслевых информационных систем:

- по мероприятия - БС;
- по КДФ - ЕСЗ;
- по книги - АБИС.

На основе "культурного опыта" пользователя, идентифицированного по ID, РС должен направлять набор рекомендаций по каждой из категорий с учетом подходов Content-based и Collaborative Filtering.

Требования к порядок взаимодействия РС с ИС БМ не предъявляются.

### Пояснения к направляемым материалам

В качестве материалов для анализа и работы над разрабатываемым решением направляются датасеты по вышеуказанным категориям.

В датасете представлены вкладки:

- каталог - перечень атрибутов, характеризующих библиографическую запись - уникальное наименование издания;
- экземпляры - перечень изданий, привязанных к конкретным библиографическим записям, с указанием ряда дополнительных атрибутов;
- читатели - обезличенный перечень пользователей БМ;
- книговыдача - сводный реестр изданий выдаваемых читателям (с указанием сопутствующих атрибутов).

В датасете представлены файлы:

- Classifier\_hackaton - перечень направлений КДФ с указанием соответствующих идентификаторов;
- org\_hackaton - перечень организаций;
- services\_hackaton - перечень функционирующих КДФ;
- Pupil\_hackaton - обезличенный перечень пользователей КЦ;
- request\_hackaton - реестр заявлений на зачисление в КДФ;
- MegaRelation\_hackaton - реестр зачислений пользователей в КДФ.

Для удобства, в рамках данной категории приложены sql запросы.

В датасете представлена выгрузка мероприятий за текущий календарный год.

## Задача 3

### Департамент здравоохранения города Москвы

*Описание задачи «Разработка инструмента оценки качества работы алгоритмов разметки медицинских изображений»:*

ГБУЗ «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы» создан в 1996 году и является ведущей организацией в России в области развития и повышения эффективности службы лучевой и инструментальной диагностики. Осуществляет научную, технологическую, учебную, организационную и методологическую поддержку рентгенологов, радиологов и врачей функциональной диагностики. Центр ведёт научно-исследовательские и практические работы в области искусственного интеллекта в здравоохранении и телемедицины на национальном и международном уровнях.

#### АКТУАЛЬНОСТЬ ЗАДАЧИ

На сегодняшний день здравоохранение является одной из перспективных сфер для разработки решений и технологий, работающих на основе искусственного интеллекта (ИИ). Уже за последние годы в клиническую практику были внедрены многие инновационные продукты.

- Системы распознавания речи помогают врачам заполнять медицинскую документацию и взаимодействовать с интерфейсом медицинских информационных систем.
- Чат-боты позволяют регулярно взаимодействовать с пациентами, например напоминать о приёме лекарств и самостоятельном проведении диагностических манипуляций (измерение артериального давления, уровня сахара крови и т.п.), уточнять состояние здоровья пациента, записывать и сохранять результаты измерений.
- Автоматический анализ Big data позволяет определять на основе лабораторных анализов предикторы и риски возникновения заболеваний.
- Автоматический анализ медицинских изображений, например маммограмм и томограмм, позволяет выявлять патологии, проводить приоритизацию пациентов по тяжести состояния и в целом облегчает врачу-рентгенологу интерпретацию исследований.

В конце 2019 года в Москве стартовал эксперимент по использованию технологий компьютерного зрения, работающий на основе ИИ, для анализа медицинских изображений. Предполагалось, что разработанные решения смогут автоматически анализировать и выявлять признаки патологии на диагностических изображениях. И уже сейчас внедрены ИИ-сервисы, помогающие врачам в диагностике COVID-19, рака лёгкого, рака молочной железы, туберкулёза. В ближайшем будущем запланировано расширение эксперимента и внедрение сервисов для диагностики заболеваний сердечно-сосудистой системы, желудочно-кишечного тракта, центральной нервной и опорно-двигательной систем.

## ОПРЕДЕЛЕНИЕ ПРОБЛЕМЫ

Перед тем как внедрить сервис, работающий на основе ИИ, в рутинную клиническую практику, его необходимо протестировать на техническую готовность, а также проверить, насколько он соответствует заявленным характеристикам диагностической точности. Это называется аналитической валидацией алгоритма. Прошедшие данную проверку ИИ-сервисы допускаются к интеграции с медицинскими информационными системами, в том числе городскими системами здравоохранения.

Сейчас аналитическая валидация проводится вручную. Это достаточно трудоёмкий процесс. К тому же неавтоматизированная валидация допускает случайные или намеренные отклонения от утверждённого алгоритма тестирования, позволяет манипулировать наборами данных, а также может потенциально ставить в неравные условия участников тестирования.

## ЗАДАЧА

Мы предполагаем, что возможно разработать единые подходы для автоматической аналитической валидации алгоритма и набора метрик для последующего сравнения с эталонным набором данных (результат экспертной разметки).

В рамках хакатона командам предлагается разработать инструмент, который будет оценивать качество разметки медицинских изображений ИИ-сервисами. Он должен автоматически определять набор метрик, а результаты его работы должны наиболее точно совпадать с мнением врача-эксперта.

## ЧТО ПРЕДОСТАВИМ ДЛЯ РЕШЕНИЯ ЗАДАЧИ

Командам будут предоставлены рентгенограммы органов грудной клетки (подвыборка из датасета NIH Chest X-rays), содержащие как патологию, так и «норму» (Origin), и наборы разметок. В наборах будет представлена разметка врача-эксперта (Expert) и трёх интеллектуальных алгоритмов (sample\_1, sample\_2, sample\_3).

Также врач-эксперт оценил корректность работы интеллектуальных алгоритмов по 5-ти бальной шкале (OpenPart.csv), где 5 – полностью согласен с разметкой интеллектуального алгоритма, 1 – полностью не согласен. Для усложнения задачи часть оценок была скрыта.

Структура данных:

- Корневая папка — Dataset
- DX\_TEST\_RESULT\_FULL.csv – описание находок
- Origin – входные изображения
- Expert — разметка врача-эксперта (Ground Truth)
- sample\_1 — первая разметка для сравнения
- sample\_2 — вторая разметка для сравнения
- sample\_3 — третья разметка для сравнения
- OpenPart.csv – оценка корректности работы интеллектуальных алгоритмов

Разметка предоставляется в двух форматах для простоты использования:

- Описание находок в виде набора прямоугольников и эллипсов. Описание для прямоугольников и эллипсов (DX\_TEST\_RESULT\_FULL.csv):

- Тип объекта Shape, бывает Rectangle — прямоугольник и Circle — эллипс.
- Эллипс имеет ориентацию параллельно осям.

○ Координаты центра x, y: xcenter, ycenter.

○ Половина ширины и половина высоты области w/2, h/2: rhorizontal, rvertical.

● Сегментационная карта находок. Отрисованные значения файла DX\_TEST\_RESULT\_FULL.csv. Файлы в формате png. При этом если значение пикселя 0 — область не аномальная, если значение 255 — аномальная.

## КРИТЕРИИ ОЦЕНКИ РАЗРАБОТАННЫХ РЕШЕНИЙ

### ОБЩИЕ ПРАВИЛА

- Не допускается использование исходников с закрытой лицензией.
- Не разрешается использовать закрытые данные от прочих участников (возможность предоставления своих данных другим участникам следует обсуждать с жюри).
- Если вы используете при разработке свои наработки, то они должны быть опубликованы и описаны.
- Запрещается ручная разметка датасета (для обучения, в качестве ответа и т.д.).

Оценивать предложенное решение мы будем по трем основным критериям:

Точность предсказания.

Сопроводительная документация.

Программный код.

### ТОЧНОСТЬ ПРЕДСКАЗАНИЯ

Точность предсказания — это основной критерий, по которому будет определен победитель. Ведь в дальнейшем именно точность предсказания поможет оценить качество работы ИИ-сервиса и принять решение о его внедрении в рутинную практику.

### СОПРОВОДИТЕЛЬНАЯ ДОКУМЕНТАЦИЯ

В сопроводительной документации необходимо максимально подробно описать гипотезу и алгоритм работы вашего решения. Что ещё важно отразить в документации:

- Должны быть описаны все применяемые методы. Если использовались готовые реализации — вам необходимо указать ссылку на этот метод.

- В случае проведения процесса обучения — он должен быть строго описан.

- Если вы будете использовать модифицированные классические методы оценки качества (mAP, IoU, Dice и т.п.) — все изменения должны быть описаны.

Главная задача сопроводительной документации — предоставить возможность воспроизведения вашего решения специалистами нашего Центра!

### ПРОГРАММНЫЙ КОД

К программному коду мы предъявляем три обязательных требования:

Исходные коды должны соответствовать сопроводительной документации.

Должен быть пример запуска приведённого кода, например с помощью консоли. Формат входа должен быть вида «папка с входными

исследованиями» → «отчёт в описанном формате».

По запросу жюри необходимо продемонстрировать запуск и исполнение кода.

Плюсами при оценке решений будут:

- Собраный докер-контейнер.
- Google Colab или аналог, позволяющий напрямую проверить работоспособность вашего алгоритма.
- Комментарии в коде.

Мы хотим обратить внимание, что в случае несоблюдения обязательных требований, участник не будет допущен в следующий этап конкурса.

## Задача 4

### Департамент предпринимательства и инновационного развития города Москвы

*Описание задачи «Разработка умного правового помощника для предпринимателей»:*

При поддержке Департамента предпринимательства и инновационного развития города Москвы функционирует **Штаб по защите бизнеса**, куда обращаются предприниматели по вопросам нарушения их прав органами исполнительной власти.

Самое большое количество времени для предпринимателя тратится на понимание того, как его обращение рассматривается соответствующим органом власти:

- Какие нормативные акты описывают, как его обращение должно быть рассмотрено
- Какие документы ему необходимо предоставить в заявке, чтобы обжаловать решение органа власти
- Все ли документы он приложил
- Нет ли ошибок в его документах
- По сути обращения согласно регламентам, регулирующим его ситуацию, прав он или орган власти. Почему (по каким нормам) он прав или не прав

С другой стороны, эксперты Штаба по защите бизнеса также тратят большое количество времени на анализ запроса предпринимателя, понимание, под какие нормы регулирования подпадает данный запрос, и первичную проверку комплектности документов.

Задача

Необходимо создать сервис, который на основании содержания обращения заявителя и/или его документов сможет проводить автоматическую первичную оценку обращения и выдавать рекомендации по составу и содержанию необходимых документов, и давать оценку степени «правоты» позиции заявителя, т.е., что заявитель прав или нет, на основании соответствия или не соответствия его требований правовым актам таким-то, п. таким-то.

Такую оценку должен иметь возможность видеть, как заявитель, так и эксперт, рассматривающий обращение, чтобы достигалась максимальная экономия времени с обеих сторон.

Требования к функционалу решения

Решение должно обеспечивать:

1. визуальное оформление текстового запроса с описанием сути обращения на web-сайте/чате в соц.сети
2. классификацию обращения – определять основную тему («класс») обращения

3. определять нормы законодательства, регулирующие обращение
4. подсказывать, какие подтверждающие документы должны быть представлены
5. запрашивать текст каждого документа и анализировать содержание
6. уметь распознавать текст из приложенных документов формата MS Word
7. уметь распознавать текст из приложенных скан-образов документов в PDF
8. подсказывать, какие из уже представленных документов соответствуют теме запроса, а, какие необходимо заменить
9. подсказывать, какие из уже представленных документов необходимо доработать или, какие документы необходимо добавить, чтобы максимально соответствовать требованиям регулирующего законодательства и обосновать, что заявитель - прав
10. давать оценку степени «правоты» позиции заявителя, т.е., что заявитель прав или нет, на основании соответствия или не соответствия его требований и/или документов правовым актам таким-то, п. таким-то
11. Решение может быть реализовано в виде интерактивного помощника/чат-бота с возможностью в любой момент подключить к «диалогу» эксперта от Штаба по защите бизнеса

#### Требования к технической реализации решения

Технически данное решение д.б. сделано в виде web-сервиса, позволяющего подключать к нему как формы web-страниц сайта, так и чат-бота.

Решение должно размещаться в закрытом контуре защищенного ЦОД и не требовать взаимодействия с внешними ресурсами, базами, сервисами и т.д.

#### Критерии оценки

Прототипы решений по степени их технической проработанности будут оцениваться согласно «Требованиям к функционалу решения» - технически более функциональными будут считаться решения, которые реализуют наибольшее количество требований п.1. - п.11.

По степени функциональной проработанности решения будут оцениваться по качеству/глубине оценок и рекомендаций на любом тестовом примере, т.е.:

1. При наличии нескольких пунктов требований правового акта, регулирующих запрос, решение, определившее 1 требование, считается хуже решения, определившего 2 требования и т.д.
2. При наличии нескольких требуемых документов решение, определившее 1 требуемый документ, считается хуже решения, определившего 2 требуемых документа и т.д.
3. При наличии ошибок в документах заявителя, лучшим решением считается то, которое верно определило большее их количество
4. При предоставлении обоснования позиции заявителя лучшим решением считается решение, определившее большее количество требований, которым обращение соответствует или не соответствует

## Примеры данных

Наиболее часто в Штаб по защите Бизнеса запросы поступают по вопросам регистрационных действий с объектами недвижимости, поэтому составление прототипа требуется реализовать на основе регламента работы РОСРЕЕСТРа (Приказ Минэкономразвития России от 07.06.2017 N 278).

Для проверки будут представлены 7 примеров обращений:

- 5, где есть суть обращения, позиция (ответ органа власти с обоснованием) и результат.
- 2, где результат НЕ указан, на них должны быть показаны функциональные возможности прототипов

## Задача 5

### Государственная инспекция по контролю за использованием объектов недвижимости города Москвы

*Описание задачи «Автоматизированный анализ архивов фото- и скан- копий технических и контрактных документов»:*

1. Описание объекта/области решения задачи: в какой области/отрасли/сфере необходимо создать решение?

Решение должно представлять из себя систему классификации текстов и структуры неформализованных электронных документов в формате изображений, .pdf, .jpeg, .tiff и так далее.

Система анализа структуры текста и распознавания символов и объектов должна быть использована в дальнейшем в решении основной задачи Хакатона (смотри Приложение 1.).

2. Описание проблемы: какую отраслевую проблематику решает создаваемое решение (цифровой продукт)?

На текущий момент все еще существует множество сегментов деятельности государственных учреждений, в которых используется документация, формируемая на «бумажном носителе» или в общедоступных форматах текстовых редакторов. Как правило такие документы структурируются для анализа специалистами путем ручного переноса сведений со скан-копий документов или текстовых документов в структурированные таблицы. При этом специалисты используют существующие на рынке OCR решения по распознаванию текстов с дельнейшей выверкой и корректировкой, что отнимает колоссальные объемы ресурсов и несет за собой высокий процент ошибки.

В данной задаче решается конкретная задача по классификации документов на ограниченном наборе данных.

3. Описание образа результата: какой формат создаваемого решения предпочтителен (мобильное приложение, информационный сайт, информационная платформа или иной ИТ-продукт)?

Целевое решение должно:

- реализовать систему классификации текстов и структуры неформализованных электронных документов (автоматизированный парсер);
- иметь веб-интерфейс, позволяющий загружать исходные данные для классификации в штучном и пакетном режиме, задавать параметры классификации с возможностью их дальнейшего расширения (при этом состав параметров не должен ограничиваться конкретным решением), группировать документы по классифицированному типу документа;
- разработать и использовать в решении действующий прототип БД для использования при реализации основной задачи распознавания;

- учесть в концепции для каждого типа документа информацию о заранее определенных атрибутах для оценки качества распознавания этого типа документа, предусмотреть возможность задавать эти атрибуты в web-интерфейсе и сохранять в БД;

- предоставлять возможность выгружать структурированные .csv и .xml файлы, содержащие данные о типах классифицированных алгоритмом, их количестве и ссылках на них;

**4. - Применяемые технологии:** на базе каких технологий или с применением каких технологий необходимо создать цифровое решение? Решение может использовать широкий спектр технологий, разработанных для анализа сканированных документов (OCR) как собственной разработки, так и свободно-распространяемого программного обеспечения. Преимущественно приоритет отдается использованию решениям с открытым исходным кодом и отечественным разработкам.

**5. Специальные функциональные требования, предъявляемые к результатам работы.**

5.1. Автоматизированный парсер должен выявлять из перечня входящих данных документы с типами:

- Свидетельство об утверждении архитектурно-градостроительного решения
- Технический паспорт
  - Экспликация.
  - поэтажный план.
  - Ситуационный план.
- Договор аренды земельного участка.
- Разрешение на ввод объекта в эксплуатацию
- Разрешение на строительство
- Свидетельство о государственной регистрации права
- Выписка из ЕГРН
- Акт приемки в эксплуатацию законченного строительства
- Выписка из постановления
- Инвентаризационная карточка
- Извлечение из технического паспорта

5.2. Должна быть Создана Концепция решения основной задачи Хакатона

Структура Концепции решения (смотри Приложение 1.) , глубина детализации, количество и качество сопроводительных материалов – на усмотрение исполнителя. Оценка производится Экспертной комиссией.

5.3. Дополнительные требования к автоматизированному парсеру

Реализовать дополнительные возможности кроме основных возможностей решения, описанных в п.3 настоящего документа:

- Предусмотреть возможность автоматической группировки / фильтрации документов по заданному признаку: дате документа, реквизитам документа: номер, дата, выдавший орган.
- Необходимо автоматически находить в документе схемы, чертежи и сохранять их отдельно в формате jpeg с указанием в названии из какого они документа, его даты и номера.
- Определять, что в одном файле может быть несколько отсканированных документов, должна их автоматически разделять на разные файлы и определять тип каждого.

**6. Портрет аудитории, которая может выполнить поставленную задачу.**

Команда должна иметь практический опыт разработки решений распознавания скан-копий документов, перевода в структурированный машиночитаемый формат и последующего анализа. Опыт работы с системами машинного обучения.

**7. Передача результата. Критерии приемки и оценки.**

1. Все исходники должны быть выложены в единый репозиторий и переданы в полном объеме конкурсной комиссии.
2. Передаваемый результат в соответствии с «Описание образа результата» оценивается по критериям экспертами Заказчика работ:

№пп	Критерий	Вес
1	Оценка качества реализации классификации Экспертами	30%
2	Оценка Концепции Решения экспертной комиссией	20%
3	Техническое интервью и оценка технической архитектуры решения	20%
4	Реализация дополнительных требований	15%
5	HR интервью	10%

### 3. Формула расчета итоговой оценки.

Итоговая оценка = сумма (Критериальная оценка  $i$  \* Вес  $i$ )

Победителем признается та команда, решение которой набрало больше всего баллов Итоговой оценки.

Для участия в 3х недельном интенсиве выбираются несколько лучших команд. С этими командами будет подписываться NDA для работы на реальном наборе данных в рамках выполнения Основной задачи Хакатона.

## Приложение 1.

Задача «Автоматизированный анализ архива технических документов МосгорБТИ».

1. Описание объекта/области решения задачи: в какой области/отрасли/сфере необходимо создать решение?

Решение должно представлять из себя систему распознавания текстов и структуры неформализованных электронных документов в формате изображений, .pdf, .jpeg, .tiff и так далее, анализа распознанного текста, анализа структуры документа, поиска ключевых слов и вывода итоговых данных в табличной форме (далее – Классификатор).

Система анализа структуры текста и распознавания символов и объектов в изображениях должна использовать алгоритмы машинного обучения оператором.

2. Описание проблемы: какую отраслевую проблематику решает создаваемое решение (цифровой продукт)?

На текущий момент все еще существует множество сегментов деятельности государственных учреждений, в которых используется документация, формируемая на «бумажном носителе» или в общедоступных форматах текстовых редакторов. Как правило такие документы структурируются для анализа специалистами путем ручного переноса сведений со скан-копий документов или текстовых документов в структурированные таблицы. При этом специалисты используют существующие на рынке OCR решения по распознаванию текстов с дельнейшей выверкой и корректировкой, что отнимает колоссальные объемы ресурсов и несет за собой высокий процент ошибки.

В данной задаче решается конкретная задача по выявлению спецификаций технических паспортов и прикрепленных документов для оптимизации аналитической работы и последующей автоматизации деятельности БТИ. Решение должно оптимизировать процесс анализа отсканированных неформализованных документов и формирования структурированной базы данных из исходных документов в разрезе набора выявленных документов, даты составления, данных о составителе, об утверждающем лице, владельце, геопространственных координат и т.д.

3. Описание образа результата: какой формат создаваемого решения предпочтителен (мобильное приложение, информационный сайт, информационная платформа или иной ИТ-продукт)?

Целевое решение должно:

- иметь веб-интерфейс, позволяющий загружать исходные данные для анализа в штучном и пакетном режиме, а также обучать систему анализа текста и таблиц, распознанных неуверенно или некорректно;

- предоставлять возможность выгружать структурированный .csv и .xml файл, содержащий обработанные данные в структурированном виде;
- предоставлять отчет-протокол о работе алгоритма с расчетом применяемых мер качества распознавания  
вероятность правильного распознавания,  
вероятность ошибки второго рода,  
F-мера,  
ROC - кривая,  
коэффициент корреляции Мэтьюса и др.;
- предоставлять Web-интерфейс полнотекстового и контекстного поиска документа;
- предоставлять REST API для запроса документа(ов) по набору параметров;
- BI систему построения отчетов, позволяющую строить аналитические формы отчетности

**4. Применяемые технологии:** на базе каких технологий или с применением каких технологий необходимо создать цифровое решение? Решение может использовать широкий спектр технологий, разработанных для анализа сканированных документов (OCR):

- алгоритмов распознавания с различными методами обучения (с учителем, без учителя, с подкреплением)
- технологий анализа текстовой информации полнотекстового и контекстного поиска.

Преимущественно приоритет отдается использованию решениям с открытым исходным кодом и отечественным разработкам, а также решениям с микросервисной архитектурой.

**5. Специальные функциональные требования, предъявляемые к создаваемому решению.**

#### 5.1. Основная функциональность.

Парсер должен:

1. Автоматически определять тип документа (технический паспорт, разрешение на строительство, инвентаризационная карточка и т.п.) Тип документа

должен определяться исходя из классификатора документов, одному типу может соответствовать несколько вариантов написания. Также система должна определять, что в одном файле может быть несколько отсканированных документов, должна их автоматически разделять на разные файлы и определять тип каждого

2. Автоматически определять реквизиты документа: номер, дата, выдавший орган
3. Автоматически определять объект, к которому относится документ: выделять и находить в тексте документа реквизиты объекта: адрес, кад номер, площадь и т.п.
4. Также если схема/чертеж перевернуты, система должна уметь это распознавать и сохранять файл в корректном положении.
5. Автоматически находить в документе положения по «заказу» (настройкам пользователя). Например, находить раздел договора аренды «особые условия» и выписывать все пункты из этого раздела структурированно в отдельное поле
6. Необходимо предусмотреть возможность автоматической группировки документов по заданному признаку.
7. Необходимо включить в систему распознавания и анализа механизм разделения результатов
  - на распознанные данные (распознано более 95%),
  - неуверенно распознанные данные (распознано более 70%),
  - нераспознанные данные (распознано менее 70%).
8. Интерфейс возможности уточнения неуверенно распознанных данных оператором (при наличии признаков некорректного распознавания, решения должно выдавать запрос оператору на уточнение сведений).
9. При уточнении сведений оператором, корректировки должны использоваться для повышения качества дальнейшей работы решения (обучение аналитической системы оператором).

## 5.2. Качественное распознавание документов.

Парсер должен выявлять из перечня входящих данных типы документов и определять перечисленные в каждом разделе атрибуты.

*Общие атрибуты к определению для всех типов документов*

Тип документа  
Номер  
Дата  
Выдавший орган

*Свидетельство об утверждении архитектурно-градостроительного решения*

Административный округ

Район

Адрес

Наименование объекта

Функциональное назначение объекта

*Согласованные технико-экономические показатели по объекту*

Площадь застройки

Объем

Этажность

Верхняя отметка объекта

Общая площадь

В том числе надземная

Подземная

*Градостроительное заключение*

Дата документа, номер документа

Адрес строительства

Вид разрешенного произведения комплекса научно-исследовательских и проектных работ (реконструкция здания, строительство здания и т.д.)

Вид работ по разрешению

Характеристика участка:

1. Площадь участка.

2. Краткая характеристика окружающей застройки

Характеристики объекта реконструкции или сноса (этажность, материал стен, год постройки, общая площадь).

*Акт государственной приемочной комиссии о приемке в эксплуатацию законченного строительством объекта*

Дата документа

Адрес местонахождения объекта

Тип работ

Показатели объекта (по проекту и фактически)

Площадь объекта (по проекту и фактически)

Этажность (по проекту и фактически)

Количество секций (по проекту и фактически)

Общий строительный объем (по проекту и фактически)

*Экспликация к архивному поэтажному плану*

Адрес здания

Этаж

Помещение  
Площадь помещения  
Общая площадь по этажу  
Общая площадь здания  
Высотность здания  
Дата составления поэтажного плана.

*Распоряжение префекта*

Дата и номер документа  
Адрес представления земельного участка  
Площадь земельного участка

*Разрешение на строительство*

Объекта капитального строительства  
Расположенного по адресу  
Строительный адрес  
Вид строительства  
Наименование объекта капитального строительства (этапа)  
Кадастровые номера земельных участков  
Срок действия документа  
Общая площадь  
Объем  
Высота  
Количества этажей  
Количество подземных этажей  
Площадь застройки  
Площадь участка  
Иные показатели

Разрешение на производство строительно-монтажных работ

Адрес строительства объекта  
Административный округ  
Район города Москвы  
Вид строительства  
Срок действия документа

Разрешение на производство подготовительных и основных строительно-монтажных работ

Адрес строительства объекта  
Район города Москвы (округ)  
Наименование объекта  
Вид строительства

Срок действия документа

*Разрешение на ввод объекта в эксплуатацию*

Объекта капитального строительства

Расположенного по адресу

Строительный адрес

Сведения об объекте капитального строительства

Строительный объем – всего (по проекту, фактически)

В том числе надземной части (по проекту, фактически)

Общая площадь - (по проекту, фактически)

Площадь встроенно-пристроенных помещений (по проекту, фактически)

Количество зданий (шт)

Нежилые объекты

Объекты производственного назначения

Количество мест (по проекту, фактически)

Количество посещений (по проекту, фактически)

Вместимость (по проекту, фактически)

Общая площадь подземной автостоянки (по проекту, фактически)

Общая площадь административных помещений (по проекту, фактически)

Общая площадь помещений культурно-досугового назначения (по проекту, фактически)

Торговая площадь (по проекту, фактически)

Этажность (по проекту, фактически)

Объекты производственного назначения

Мощность (по проекту, фактически)

Протяженность (по проекту, фактически)

Площадь (по проекту, фактически)

Общая площадь административных помещений (по проекту, фактически)

Общая площадь помещений культурно-досугового назначения (по проекту, фактически)

Торговая площадь (по проекту, фактически)

Этажность (по проекту, фактически)

Объекты жилищного строительства

Общая площадь жилых помещений (с учетом балконов, лоджий веранд и террас)  
(по проекту, фактически)

Количество этажей (по проекту, фактически)

Количество секций (по проекту, фактически)

Количество квартир (по проекту, фактически)

*Разрешение на производство подготовительных и основных строительномонтажных работ.*

Дата и номер документа

Адрес строительства объекта

Наименование объекта

Разрешение

На проведение работ (производство подготовительных работ, устройство, новое строительство и т.д.)

Срок действия разрешения на строительство (дата).

#### *Договор аренды земельного участка*

Номер документа и дата документа

Основная часть – предмет договора, а именно характеристики земельного участка его площадь адресный ориентир. Целевое назначение земельного участка.

Хозяйствующие субъекты при наличии разделение на доли.

Срок действия договора аренды

Особые условия (особые условия использования земельного участка):

1. К примеру, установлен запрет на возведение временных и капитальных строений сооружений.

2. Установлен запрет на строительство капитальных зданий строений и сооружений.

Также необходимо распознание архивной схемы земельного участка (плана земельного участка)

#### *Плановое (реставрационное) задание.*

На разработку научно-проектной документации для проведения ремонтно-реставрационных работ на памятнике истории и культуры.

Объект.

Адрес.

Основание для выдачи задания.

Краткие исторические сведения.

Характер планируемых работ.

Дата и номер документа.

Ситуационный план.

#### *Заключение Мосгосэкспертизы*

Дата и номер документа

Адрес объекта экспертизы по реконструкции здания

Технико-экономическое обоснование

Виды работ:

Дополнительное возведение:

1. Помещений

2. Этажей

3. Возведение надстроек

Полный демонтаж здания

Технико-экономические показатели здания:

1. Количество этажей

2. Площадь застройки
3. Общая площадь здания в том числе тех подполье, надземная и подземная часть
4. Объем здания надземная часть, подземная часть
5. Полезная площадь
6. Рабочая площадь

Общий вывод по объекту

### *Договор*

в том числе по формальным признакам (например, наличие обязательных блоков, «преамбула» и «реквизиты»), обнаруживать спецификацию (зачастую представленную в формате таблицы), выделять из спецификации информацию о поставляемом товаре, выполняемых работах, оказываемых услугах, включая:

- номер строки спецификации (если графа отсутствует – порядковый номер строки);
- наименование товара;
- страна происхождения;
- производитель;
- единица измерения;
- количество;
- ставка НДС (НДФЛ), %;
- цена за единицу, руб.;
- общая стоимость, руб.;

Парсер должен находить в тексте документа информацию о товаре, выполнять разбор таблицы

в соответствии с заголовками столбцов, сопоставлять данные с городскими классификаторами и справочниками (Классификатор предметов государственного заказа и Справочник предметов государственного заказа

<http://catalog.eaist.mos.ru/catalog>), собирать данные в предоставленную форму.

Необходимо предусмотреть возможность автоматической группировки товаров, в соответствии с классификатором.

### *Технический паспорт (МосгорБТИ, МособлБТИ, Ростехинвентаризация, ВИСХАГИ)*

Источник документации

Дата заполнения документа.

Наименование объекта

Инвентарный номер

Значение основной характеристики

Год постройки

Год ввода в эксплуатацию

Кадастровый номер ЗУ

Кадастровый номер ОКС

Распознавание штампа планов зданий,

дата заполнения штампа,  
даты внесения изменений.  
Наименование владельцев.  
Состав объекта,  
наименование,  
год постройки,  
площадь.  
Экспликация на земельный участок,  
площадь,  
адресный ориентир,  
назначение.  
Отметки об обследованиях,  
дата внесения данных,  
дата изменения данных.

*Технический паспорт на здание*

Назначение объекта  
По адресу  
Владелец  
Год постройки  
Число этажей  
Кроме того, имеется: подвал, тех.подполье, цокольный этаж, мансарда, мезонин  
*«в данном разделе если подчеркнута то данные необходимо подтянуть»*  
Площадь крыши  
Объем  
Общая площадь по зданию  
В т.ч. общей без учета балконов и лоджий  
Жилые помещения: Общ. площ.  
В т.ч. общей площади без учёта балконов и лоджий  
Общая площадь, остающаяся к общему имуществу кондоминиума  
*Технический паспорт на домовладение (земельный участок)*  
Дата составления технического паспорта

Владения строениями в домовладении:

Дата записи  
Номер строения  
Наименование владельца здания (площадь права по документу)  
Основание владения (дата и номер документа)

Основные экономические показатели строений в домовладении:

Номер строения  
Назначение строения  
Материал стен  
Материал крыши  
Площадь крыши  
Число этажей

Площадь строения по наружному обмеру в кв.м.  
Строительная высота  
Объем (куб.М.)  
Даты внесения изменений

#### *Кадастровый паспорт*

Кадастровый номер  
Инвентарный номер  
Описание  
Наименование  
Назначение  
Адрес  
Местоположение  
Основная характеристика  
Кадастровый номер ЗУ  
Предыдущие номера  
Примечания  
Ситуационный план  
Масштаб  
Подписал

#### *Прочие типы документов*

- 5.2.1.1. Выписка из реестра Федерального имущества
- 5.2.1.2. Свидетельство о регистрации права
- 5.2.1.3. Распоряжение
- 5.2.1.4. Определение
- 5.2.1.5. Акт
- 5.2.1.6. Разрешение
- 5.2.1.7. Заключение
- 5.2.1.8. Проект
- 5.2.1.9. Письмо
- 5.2.1.10. Пояснительная записка

5.3. Портрет аудитории, которая может выполнить поставленную задачу. Команда должна иметь практический опыт разработки решений распознавания скан-копий документов, перевода в структурированный машиночитаемый формат и последующего анализа. Опыт работы с системами машинного обучения.

5.4. Передача результата. Критерии приемки и оценки.

1. Все исходники должны быть выложены в единый репозиторий и переданы в полном объеме конкурсной комиссии.
2. Передаваемый результат в соответствии с «Описание образа результата» оценивается по критериям экспертами Заказчика работ:

№пп	Критерий	Вес
1	Анализ отчет-протокола о работе алгоритма с расчетом применяемых мер качества распознавания и оценка качества распознавания	30%
2	Оценка качества реализации распознавания путем сопоставления с результатом распознавания специалистами БТИ г. Москвы	25%
3	Полнотекстовый и контекстный поиск	15%
4	Документирование и качество REST API	10%
5	Оценка полноты реализации функциональности решения	10%
6	BI отчеты	5%
7	Usability web интерфейса	5%

**3.** Несмотря на требования к решению задачи рассматриваться также будут и инновационные предложения, позволяющие оптимизировать указанный процесс иными методами. Инновационность предложения отражается в пояснительной записке. Оценка инновационности производится экспертами Заказчика путем анализа решения на основе заявленных тезисов в пояснительной записке.

**4.** Итоговая оценка. Формула расчета итоговой оценки.

Итоговая оценка = Критериальная оценка \* 0,8 + оценка инновационности \* 0,2

Победителем признается та команда, решение которой набрало больше всего баллов Итоговой оценки.

## **Задача 6**

### **Департамент информационных технологий города Москвы**

*Описание задачи «Разработка сервиса поддержки предпринимателей «Легкий Старт»:*

#### **Постановка задачи**

Основная задача – разработать сервис (модель или макет сервиса) поддержки принятия решения для предпринимателей. В рамках сервиса пользователю должна быть предоставлена возможность провести оценку перспектив реализации собственного проекта на территории города с учетом набора параметров, влияющих на его реализацию.

В рамках сервиса выделяются две основные ветки

1. Создание или развитие бизнеса на базе заданных услуг или товарных позиций.
2. Создание или развитие бизнеса исходя из территориальной привязки.

В целом под обоими ветками лежит сходный набор данных, различаются только связи и последовательность подачи. Формируемые в рамках предоставления сервиса зависимости дают разный результат при разных начальных условиях. Начальные условия выступают в роли переменных.

Результатом работы сервиса должна стать оценка (рекомендация) о перспективах реализации проекта пользователя в зависимости от выбранных исходных данных и параметров.

Результат работы сервиса должен быть представлен на карте города. Карта выступает ключевым агрегатором сведений и средством визуализации решения. Карта может выступать также средой формирования исходного запроса и его корректировки (уточнения).

#### **Целевая аудитория сервиса**

По умолчанию к целевой аудитории пользователей сервиса относятся юридические лица и индивидуальные предприниматели, рассматривающие возможность создания, масштабирования или реорганизации собственного бизнеса (деятельности). Под деятельностью понимаются активности, направленные на возмездное удовлетворение потребностей населения (потенциальных потребителей) в товарах или услугах, в целях получения прибыли.

В зависимости от выбранного подхода к решению задачи целевая аудитория может быть скорректирована, однако данные изменения должны быть отражены в описании задачи в сопроводительной документации к проекту решения.

## **Актуальность задачи**

В настоящее время нас окружает масса открытой и доступной информации. Зачастую информации больше, чем можно воспринять и обработать. При этом не смотря на избыточность доступной информации большая ее часть остаётся не востребованной, и, несмотря на то, что для более точного решения любой из актуальных задач требуется участь как можно больше переменных, подавляющее большинство решения принимаются на базе простых сопоставлений, количество переменных сводится к минимуму и множество факторов отбрасывается. Такой подход приводит к тому, что решения часто оказываются неверными, так как модели, на которые они опирались – не достаточно точными. Зачастую самостоятельно построить точную и максимально полную модель без использования специализированных инструментов и сложно, и дорого. Индивидуальные затраты на решение такой задачи не сопоставимы с ожидаемым результатом. Чем собирать для себя калькулятор, проще один раз посчитать «в столбик». Однако именно отсутствие надежного и доступного инструмента и приводит в последствии к ошибкам.

Государству, как регулятору, выгодна конкуренция, но не выгодна низкая стабильность бизнеса. Сопутствующие расходы, связанные с заменой одного бизнеса на другой, его ликвидацией или репрофилированием, чреваты выпадающими доходами из-за сокращения налоговой базы, расходами на переобучение и трудоустройство высвобождаемого персонала и иными организационными и финансовыми мероприятиями.

Очевидно, что полностью ошибок не избежать, и безусловно со стороны государства будет оказана вся необходимая поддержка и людям, и бизнесу, но ошибки, связанные с неверным расчетом входным параметров можно и нужно минимизировать.

Минимизация ошибок при принятии решения повысит стабильность бизнеса и положительно повлияет на рост доходов населения в целом, что в пределе приводит к росту потребления и росту налогооблагаемой базы, как следствие. Долгосрочные проекты являются более перспективными и позволяют в том числе повысить качество планирования на уровне государства в целом.

В этой связи создание инструмента, позволяющего обеспечить комплексный анализ, помочь принять правильное решение, является востребованной и актуальной задачей. Создание общедоступного калькулятора позволит сделать и старт, и дальнейшее развитие легким.

## **Подходы к решению**

В зависимости от выбранной категории товарной позиции или их набора пользователю предоставляется послойно информация о наиболее оптимальных местах реализации указанной позиции с разным уровнем детализации – в разрезе округов, управ, районов (административно территориальное деление города). Первый слой сведений – уровень обеспеченности выбранной товарной позицией в городе исходя из территориального признака. Далее, сведения о доступных торговых площадях (включая выставленные городом площади) в местах наиболее

оптимального развития бизнеса, стоимость аренды, площадь, дополнительные условия и ограничения (при наличии). Уже в привязке к помещению (потенциальной точке распространения) предоставляются дополнительные сведения относительно проходимости общей (поток населения) и потенциальных покупателей (сложный кейс, но решаемый через данные сотовых операторов по проходимости, наложенных на статистику среднего чека из налоговой по товарной позиции плюс сведения банка о владельце карты – все строго деперсонализированно). Так же предоставляются сведения о транспортной доступности территории (остановки общественного транспорта, станции метро, парковки). Дополнительно – сведения о потенциальных работниках – по информации о данных с рынка труда (городских и федеральных – в районе столько-то человек ведут поиск вакансии на такую-то позицию с ожидаемой оплатой такой-то). Карта криминогенной обстановки. Сведения по устойчивости бизнеса в выбранном районе (длительность аренды за одним арендатором, длительность существования фирмы, активность и иные открытые сведения из налоговой и имущественных органов).

Аналогичный характер связей может быть реализован и при реализации решения по территориальному признаку привязки бизнеса, но только в разрезе наиболее оптимальных услуг или товарных позиций к их реализации в зависимости от уровня удовлетворенности потенциального спроса населения района. В этой ветке дополнительно должны быть подсвечены сведения о налоговых льготах и преференциях для различных товаров работ или услуг, предоставляемых городом или федерацией и действующих на момент формирования запроса – дополнительный инструмент информирования бизнес сообщества о приоритетных и поддерживаемых направлениях развития на выбранной территории.

Отдельным направлением (решением) может выступить стык двух веток в режиме B2B – запрос по товарной позиции может быть сопоставлен с предложением предоставления помещения в аренду на указанной территории либо с заданными ценовыми параметрами и требуемой площадью.

В упрощенном варианте по решению команды сервис можно свести, например, к помощи в подборе помещения для заданного вида деятельности с учетом анализа по территории среднего чека клиента, плотности конкурентов (потенциальных), свободных помещений, длительности аренды (стабильность) и длительности бизнеса (частота сменяемости и род деятельности ушедшего с площадки арендатора).

Дополнительным бонусом будет являться реализация ленты времени по каждой из веток – статистика за различные отчетные периоды – как менялся спрос, состав населения (потенциальные покупатели) как назад, так и в перед по шкале (прогнозы с учетом строительства и ввода новых объектов – градостроительные планы города, как источник).

### **Источники данных и ограничения**

Для решения задачи понадобится, помимо московских баз данных, доступ к информации из федеральных баз данных, таким как налоговая (сведения по кассовым аппаратам, сведения из ЕГРЮЛ) и данные сотовых операторов и сведения банков эмитентов платежных инструментов. Часть таких данных является открытыми и могут использоваться на усмотрение команд. Все данные должны обрабатываться в строгом соответствии со 152 ФЗ и предоставляться конечным пользователям в деперсонализированном виде. Обязательным оговариваемым исключением является отказ от учета в статистике информации по онлайн торговле в районе – с одной стороны это может существенно повлиять на картину удовлетворенности, с другой – это иной вид деятельности и не является прямым конкурентом рассматриваемому. Если команды представят корректный алгоритм учета таких данных, то это будет рассматриваться как дополнительный бонус, но по умолчанию указанные данные для целей решаемой задачи рассматриваются как избыточные.

Предпочтительным к использованию при решении визуализации сервиса является штатный набор интерфейсов сайта mos.ru (ссылка на шаблоны может быть предоставлена по запросу команд в ходе решения задачи).

### **Документирование решения**

Обязательным условием надлежащего решения поставленной задачи является наличие сопроводительной документации к решению.

В сопроводительной документации необходимо описать детализированные требования в отношении задачи (описать ту область ограничений, которую команда себе определила и на чем сосредоточены усилия – указать итоговую целевую аудиторию и объем задач, решаемых сервисом). В сопроводительной документации необходимо максимально подробно описать гипотезу и алгоритм работы вашего решения. Так же должны быть описаны все применяемые методы. Если использовались готовые реализации — необходимо указать ссылку на такой метод.

Главная задача сопроводительной документации — обеспечить возможность воспроизведения реализованного решения сторонним исполнителем.

### **Требования к коду**

К программному коду предъявляем следующие требования:

1. Исходные коды должны соответствовать сопроводительной документации;
2. Должна быть обеспечена возможность выполнения процедур сборки и запуска приведённого кода;
3. По запросу жюри необходимо обеспечить проведение демонстрации запуска и исполнения кода.
4. Наличие комментариев в коде рассматривается как дополнительное преимущество.

## **Работа с данными**

Учитывая специфику решаемой задачи отдельным направлением, подлежащим оценке и анализу, является работа команды с данными. Требования в части работы с данными – отдельный объект анализа результатов работы команд. Не только алгоритм решения как таковой и качество кода в нем, но и методы и подходы к обработке исходных данных.

По причине широкой возможности выбора итогового решения и, как следствие, масштабного разброса по необходимым наборам данных командам будет предоставлен только базовый набор ключевых сведений. Допускается использование наборов открытых данных как города Москвы, так и федеральных наборов открытых данных. При наличии аргументированной потребности в данных, не представленных в открытых наборах, но доступных в системах и базах данных допускается использование тестовых массивов данных. Если с точки зрения команд требуется уникальный набор данных, сбор и обработка которых не ведутся – допускается его аргументированное использование, однако итоговая ценность предложенного сервиса будет оцениваться в том числе и с точки зрения затрат на его промышленную реализацию – как финансовых, так и организационных, а также технологических. Если ценность итогового решения будет не сопоставимой с затратами на организацию сбора, обработки и хранения данных.

### **Критерии, учитываемые при проведении итоговой оценки.**

Значимыми, с точки зрения итоговой оценки представленного решения являются следующие критерии:

1. Степень готовности решения к промышленному развертыванию (включая потребности в организации сбора данных и их ведения в отдельной системе);
2. Итоговый состав и размер целевой аудитории сервиса (предпочтение отдается сервису с максимальным охватом исходной целевой аудитории);
3. Степень востребованности предложенного сервиса (оценка экспертов);
4. Масштабируемость предложенных алгоритмов и решений (на частном примере предложен универсальный работающий алгоритм, который хорошо масштабируется на другие – более предпочтителен, чем слабо проработанный алгоритм, сервис, но работающий с широким спектром (номенклатурой) товаров и услуг);
5. Уровень проработки и описания методологии работы сервиса и качество сопроводительной документации.

## Задача 7

### Департамент образования и науки города Москвы

*Описание задачи «Разработка виртуальной лаборатории по программированию  
«Мой код»:*

#### 1. Требования к общим моментам

- 1.1. Требуется разработать автоматизированную тестирующую систему для автоматической проверки задач по программированию. Разработанная система должна быть гибкой и масштабируемой, то есть должна иметь возможность интеграции в уже готовые системы и расширения своего функционала при помощи дополнительных модулей.
- 1.2. Система должна иметь две части: клиентское и серверное приложение. Клиентское приложение представляет собой web-интерфейс для чтения, составления, редактирования заданий и отправки решений, проверка ответов и тестирование программного кода (компиляция и проверка на сэмплах и чекерах) осуществляется на серверной части.
- 1.3. В системе должна быть предусмотрена процедура регистрации и аутентификации пользователей:
  - При регистрации пользователь должен указать следующие данные: ФИО, Дату рождения (ЧЧ.ММ.ГГГГ), Образовательная организация, Адрес электронной почты, Логин и пароль для входа в систему;
  - Для аутентификации в системе пользователь должен указать свой логин и пароль;
- 1.4. В системе должны быть предусмотрены следующие роли пользователей: ученик, учитель (автор заданий), администратор.
- 1.5. Система должна предоставлять для учащихся:
  - возможность просматривать условия задач, сэмплов к ним;
  - возможность отправлять написанные решения на автоматическую проверку;
  - возможность видеть свои результаты после решения;
  - возможность видеть различные сообщения об ошибке (ошибка компиляции, ошибка времени выполнения, превышение лимита времени/памяти);
  - возможность повторной отправки решений.
- 1.6. Система должна предоставлять для учителей:

- возможность просматривать все послылки учеников;
- возможность объединять учеников в группы/классы;
- возможность смотреть мониторы - рейтинговые таблицы по курсам и/или соревнованиям для выбранных учеников;
- возможность изменять статус задачи - например, менять “ОК” на “Неверно оформлен код”.
- возможность загружать и/или создавать в редакторе на сайте условие задачи;
- возможность задавать ограничения для задачи по времени и памяти;
- возможность загружать и/или создавать в редакторе на сайте тесты для задачи в формате входной/выходной файл;
- возможность выбирать для задачи сэмплы - тесты, которые будут отображены в качестве примеров работы программы в условии задачи;
- возможность выбора из существующих или создания чекера - программы для автоматической проверки загруженного решения;
- возможность настройки постпроцессора - программы, определяющей количество набранных баллов за решение.

#### 1.7. Система должна предусматривать для администратора:

- возможность удаления пользователей;
- возможность смены роли пользователя;
- возможность удаления задач;
- возможность удаления курсов и соревнований;
- возможность отправки сообщений или уведомлений любым пользователям и их группам.

## 2. Общий вид задачи

### 2.1. Отображение окна с задачей должно содержать:

- Название задания;
- Легенда (необязательное поле - предыстория, описывающая подводящая к задаче);
- условие, описывающее постановку проблемы;
- примеры (сэмплы) - 1-3 теста, отображающиеся вместе с условием, содержат наборы входных и выходных данных к задаче;

- поле для ввода программного кода или возможность прикрепить файл с решением задачи, кнопку “Отправить”;
- строку статуса, в которой отображается вердикт по задаче и ссылка “Подробнее”. При нажатии на ссылку ученик должен иметь возможность увидеть характер ошибки в случае её наличия. Ученик не должен видеть тесты, на которых не работает его решение, но может видеть их количество и характер ошибки на каждом из них.

2.2. В рамках хакатона в системе должна быть предусмотрена обязательная возможность отправки задачи в систему на любых двух языках из списка: C/C++, Python, C#, Ruby, Perl, PHP, Java, Basic, JavaScript, R, Kotlin, Swift, Haskell, Алгоритмический. В дальнейшем система должна поддерживать добавление других языков программирования.

2.3. Использование готовых систем проверки решений и виртуальных обучающих сред в системе не допускается.

## Задача 8

### Департамент жилищно-коммунального хозяйства города Москвы

*Описание задачи «Разработка электронной базы животных без владельцев»:*

1. Описание объекта/области решения задачи: в какой области/отрасли/сфере необходимо создать решение?

В Москве функционирует 13 приютов для животных, в которых содержится около 17 тыс. собак и кошек. Все сведения о животных хранятся на бумажных носителях и в электронном виде в форматах word и excel: это карточки учета животных, акты отлова, приема-передачи в приют, ветеринарного осмотра. При этом данные систематизируются по видам мероприятий, а не по животным, что усложняет поиск нужной информации.

Информация о животных, готовых к пристройству в добрые руки, размещается в специальных онлайн-каталогах и на страницах приютов в социальных сетях.

Единый каталог животных отсутствует.

Участникам хакатона предлагается разработать прототип информационной системы «База животных без владельцев» (далее – Система).

2. Описание проблемы: какую отраслевую проблематику решает создаваемое решение (цифровой продукт)?

Система предназначена для решения следующих задач:

- контроля выполнения мероприятий по регулированию численности животных без владельцев в г.Москве;
- регистрации и учета движения животных без владельцев в городских приютах;
- ведения карточек учета животных без владельцев;
- автоматизации процесса пристройства животных в добрые руки.

К числу основных процессов, подлежащих автоматизации с использованием Системы, относятся:

- сбор, обработка и хранение данных о животных без владельцев;
- ведение Общего реестра животных без владельцев;
- ведение Реестра животных, готовых к пристройству в добрые руки;
- формирование сводных отчетных материалов.

3. Описание образа результата: какой формат создаваемого решения предпочтителен (мобильное приложение, информационный сайт, информационная платформа или иной ИТ-продукт)?

Целевое решение должно:

- реализовать базу данных учета животных без владельцев, а также обеспечить первичное наполнение базы данных сведениями, представленными в дата-сети (подробная информация приведена в п.6 ТЗ);
- иметь веб-интерфейс, позволяющий осуществлять регистрацию животных в городских приютах, а также учет их движения, в том числе выбытие из приюта;
- обеспечивать доступ в систему в соответствии с уровнями доступа пользователя;

- обеспечить возможность ведения справочников системы (создание, хранение, редактирование);
- обеспечить возможность заполнения отдельных позиций из справочников системы;
- предоставлять возможность выгружать сводные отчеты в соответствии с требованиями, приведенными в п.5 ТЗ;
- иметь общедоступный веб-интерфейс реестра животных без владельцев, готовых к пристрою в добрые руки, обеспечивающий доступ к просмотру данных по таким животным, возможность подбора питомца по заданным параметрам;
- содержать предложения по алгоритму пристрою животных из приютов в виде блок-схемы с описанием;
- иметь открытый API-интерфейс, обеспечивающий передачу данных по животным без владельцев, готовых к пристрою в добрые руки, во внешние системы в соответствии с требованиями, приведенными в п.6 ТЗ.

4. Применяемые технологии: на базе каких технологий или с применением каких технологий необходимо создать цифровое решение?

При разработке программного кода должны применяться методы безопасного программирования, включающие контроль версионности исходного кода. Реализация технических, программных, организационных и иных решений, предусмотренных проектом прототипа Системы, не должна приводить к нарушению авторских и смежных прав третьих лиц.

При использовании в Системе программ (программных комплексов или компонентов), разработанных третьими лицами, условия, на которых передается право на использование (исполнение) этих программ, не должны накладывать ограничений, препятствующих использованию Системы по ее прямому назначению.

Разработка программного обеспечения Системы должна выполняться с использованием стандартных и унифицированных методов разработки программных средств, входящих в состав используемых систем автоматизированного проектирования и разработки. Система проектирования должна обеспечивать унификацию функциональных задач, операций и интерфейсов в части их реализации.

Необходимо максимально использовать типовые поставляемые решения, лицензионное и сертифицированное программное обеспечение (операционные системы, СУБД, системы формирования и ведения отчетности, средства защиты информации и разграничения доступа и т.д.). Преимущественно приоритет отдается использованию решений с открытым исходным кодом и отечественным разработкам.

5. Этапы выполнения работ.

Этапы должны выполняться участниками последовательно.

Состав этапов:

- 5.1. Реализация прототипа базы данных животных без владельцев.
- 5.2. Реализация веб-интерфейса реестра животных без владельцев.

5.3. Реализация открытого API-интерфейса, обеспечивающего передачу данных по животным без владельцев, готовых к пристрою в добрые руки, во внешние информационные системы.

5.4. Реализация предложений по алгоритму пристрою животных в добрые руки в виде блок-схемы с описанием.

5.5. Реализация реестра животных без владельцев, готовых к пристрою.

5.6. Реализация отчета о мониторинге животных в городских приютах. Данный этап является дополнительным для участников.

6. Специальные функциональные требования, предъявляемые к результатам работы.

6.1. Требования к базе данных учета животных без владельцев.

- База данных должна содержать данные о животных, представленные в дата-сете (приложение 1). При ведении сведений о животных должны использоваться атрибуты учета, представленные в дата-сете.

- Прототип системы должен обеспечивать ведение (создание, хранение, редактирование) следующих справочников:

- справочник приютов;
- справочник эксплуатирующих организаций;
- справочник видов животных;
- справочник пола;
- справочник пород;
- справочник окрасов;
- справочник типов шерсти;
- справочник типов ушей;
- справочник типов хвостов;
- справочник причин смерти;
- справочник причин выбытия из приюта;
- справочник причин эвтаназии.

Справочники приведены в приложении 2 к настоящему ТЗ.

При ведении базы данных учета животных без владельцев должны использоваться вышеуказанные справочники.

Прототип системы должен обеспечивать формирование карточек учета животного (приложение 3).

6.2. Требования к интерфейсу реестра животных без владельцев

- Просмотр реестра животных, фильтрация животных реестра по следующим параметрам и формирование выборок животных по различным атрибутам, в том числе по № карточки,

№ идентификационной метки, кличке, а также по административному округу, приюту, периоду регистрации карточки, виду животного, возрасту, размеру, причине выбытия, статусу животного.

- Возможность просмотра информации по выбранному животному с возможностью перехода к просмотру Карточки учета животного (приложение 3). Система должна обеспечивать возможность редактирования данных карточки учета, а также возможность вывода карточки учета на печать в формате А4 или ее

экспорта в формате docx. Допустимо графы таблиц «ФИО/подпись» выводить только при отправлении карточки учета на печать.

- Постановка животных на учет в приюте и их выбытие из приютов.
- Простановка состояний, указывающих на готовность передачи животных владельцев (для последующей выгрузки данных об этом животном во внешнюю ИС).
- Внесение сведений по проведению обязательных мероприятий по чипированию, обработке от экто- и эндопаразитов, вакцинации, стерилизации (кастрации) по животным.
- Сводный отчет (реестр карточек учета животного) должен выгружаться в формате xlsx по форме приложения 4.

Доступ к просмотру, редактированию данных должен осуществляться в соответствии с уровнями доступа. Реализовать несколько ролевых моделей доступа:

- Департамент (ДЖКХ) – доступ ко всей информации;
- префектура – к информации приютов ведомственной принадлежности;
- эксплуатирующие организации (ГБУ «Автомобильные дороги» АО и ГБУ «Доринвест»):
  - ГБУ «Доринвест» - доступ ко всей информации;
  - ГБУ «Автомобильные дороги» АО - к информации приютов ведомственной принадлежности;
- приюты – к информации о животных, закрепленных за данным приютом.

При этом вносят сведения о животных исключительно сотрудники приютов для животных.

Для тестирования возможностей разрабатываемого прототипа Системы – достаточно сформировать по 1 учетной записи (далее – УЗ) уровней доступа Департамент, префектура и эксплуатирующие организации, а также по 1 УЗ для каждого приюта.

6.3. Требования к открытому API-интерфейсу, обеспечивающему передачу данных по животным без владельцев, готовых к пристрою в добрые руки, во внешние информационные системы

Состав данных API-интерфейса (далее – сервис) определяется участником самостоятельно, исходя из следующих задач:

- Пользователи конечной информационной системы должны получать основную информацию о животных.
- Сервис должен обеспечить передачу информации только по животным, готовым к пристрою в добрые руки (с проставленной отметкой о социализации).
  - Документация сервиса должна быть описана с использованием Swagger и представлена в соответствии с шаблоном, приведенным в приложении 5.
  - Данный сервис должен использоваться участником при последующей реализации реестра животных, готовых к пристрою в добрые руки.

6.4. Требования к алгоритму устройства в добрые руки животных из приютов с использованием сервиса (п.6.5 ТЗ) в виде блок-схемы с описанием; Участник должен предложить схему взаимодействия потенциального владельца – пользователя реестра животных без владельцев, готовых к устройству в добрые руки и приюта, в котором содержится интересующее пользователя животное, посредством функционала реестра животных без владельцев, готовых к устройству в добрые руки.

Алгоритм получения животных из приюта должен быть простым и понятным для пользователя сервиса.

Требования к используемой нотации не предъявляются.

6.5. Требования к реестру животных без владельцев, готовых к устройству в добрые руки

На данном этапе участникам предлагается реализовать реестр животных, готовых к устройству в добрые руки, куда автоматически должны попадать карточки животных из реестра животных без владельцев, где стоит специальная отметка об их социализации. Желательно, чтобы интерфейс реестра был выполнен в стилистике mos.ru.

Интерфейс должен содержать следующую информацию:

Перечень животных, находящихся в реестре животных, готовых к устройству в добрые руки, на текущую дату, с краткой информацией о животных. Состав отображаемой информации определяется участником самостоятельно.

- Информацию о приютах, в котором содержится выбранное пользователем животное. Состав отображаемой информации определяется участником самостоятельно.
- Фильтры для подбора животного пользователями по определенным параметрам. Состав параметров определяется участником самостоятельно.
- Карточка просмотра подробной информации о животном с информацией о приюте, в котором содержится животное

Интерфейс реестра должен содержать элементы и информацию, необходимые для реализации предлагаемого участником алгоритма устройства в добрые руки животных из приютов.

Реестр животных, готовых к устройству в добрые руки, в качестве источника данных должен использовать реализованный на одном из предыдущих этапов API-интерфейс, обеспечивающий передачу данных по животным без владельцев.

6.6. Требования к составу отчет о мониторинге животных в городских приютах. Отчет о мониторинге животных в городских приютах должен формироваться в реестре животных без владельцев с возможностью выгрузки его в формате docx. Форма отчета приведена в приложении 6.

## Задача 9

### Департамент строительства города Москвы

*Описание задачи «Разработка мобильного приложения для предупреждения несчастных случаев и учета рабочего времени на строительной площадке»:*

#### Общие сведения

##### 1.1. Наименование сервиса

Сервис предупреждения несчастных случаев и мониторинга времени нахождения строителей на строительных площадках объектов капитального строительства города Москвы (далее – сервис).

##### 1.2. Заказчик

Департамент строительства города Москвы.

##### 1.3. Термины и сокращения

Оконечное оборудование обработки данных или терминальное оборудование (переносимое техническое устройство)	оборудование, преобразующее телеметрическую информацию в данные для передачи по сетям связи
ФИО	фамилия, имя, отчество (при наличии)
ГИС	государственные информационные системы Правительства Москвы
Информация	сведения (сообщения, данные) независимо от формы их представления
Информационная среда	совокупность программного обеспечения необходимых для работы сервиса
Телеметрическая информация	информация, передаваемая окончательным оборудованием, включающая в себя информацию о физических параметрах, состоянии контролируемого субъекта и иных событиях
Нотация BPMN	система условных обозначений (нотаций) для моделирования бизнес-процессов в виде диаграмм
API	описание способов (набор классов, процедур, функций, структур или констант) необходимых для интеграции информационных средств, баз данных и(или) иных источников информационных сведений
Субъект	физическое лицо, в отношении которого применяются электронные технологии мониторинга местоположения

ТЗ	техническое задание на оказание услуг по мониторингу местоположения субъекта с использованием электронных технологий с целью реализации государственных функций в городе Москве
----	---

## **Назначение и цели сервиса**

### **2.1. Назначение оказания сервиса**

Назначением сервиса является автоматизация мероприятий по мониторингу местоположения субъекта (геолокация) с идентификацией субъекта посредством непрерывного сбора телеметрической информации с технических средств, необходимой для последующего автоматизированного сопоставления с данными в государственных информационных системах Правительства Москвы, с целью выполнения Комплексом градостроительной политики и строительства города Москвы и иными органами исполнительной власти города Москвы возложенных на них функций.

### **2.2. Цели оказания услуг**

Основными целями оказания услуг являются:

- автоматизация сбора и идентификации телеметрической информации о местоположении субъекта на территории города Москвы;
- автоматизация функций координации участников контрольно-надзорных мероприятий органов исполнительной власти города Москвы и подведомственных им организаций.

### **2.3. Места оказания сервиса**

Сервис предоставляется в соответствии с ТЗ на территории города Москвы.

## **Требования к сервису**

### **3.1. Общие требования к сервису**

Сбор телеметрической информации проводится приложения установленного в смартфон, носимого субъектом.

Сбор, агрегация и передача в государственные информационные системы Правительства Москвы информационных сведений сервиса происходит посредством программного обеспечения Информационной среды. Требования к программному обеспечению Информационной среды представлены в пункте 3.2 ТЗ.

Настройка и интеграция Информационной среды и ГИС осуществляется Исполнителем самостоятельно в соответствии с требованиями ТЗ.

В зону ответственности Исполнителя входит реализация следующих мероприятий, осуществляемых им в рамках разработки сервиса:

- разработка механизма регистрации новых носимых устройств в Информационной среде и передача сведений в ГИС;
- разработка механизма сбора, автоматизированной обработки телеметрической информации с последующим формированием аналитической отчетности;

Права на сервис, включая информационные данные, принадлежат Заказчику.

### **3.2. Требования к размещению Информационной среды**

Информационная среда обеспечивается Исполнителем. Интеграция Информационной среды и ГИС обеспечивается Исполнителем посредством предоставления Заказчику API (сторонами подписывается Протокол интеграции Информационной среды и ГИС.). Перечень ГИС, с которыми необходимо обеспечить интеграцию, Заказчик предоставляет Исполнителю на подготовительном этапе.

Информационная среда должна обеспечивать выполнение следующих функций:

- Формирование контролируемых зон объектов капитального строительства
- Функцию определения и регистрации работы смартфона с установленным приложением в пределах контролируемой зоны.
- Ведение реестра объекта капитального строительства;
- Ведения реестра субъектов;
- Ведение карточки субъектов;
- Ведение карточки объекта капитального строительства;
- Функцию определения совпадения траекторий двух смартфонов и более;
- Функцию определения активности субъекта на основе данных получаемых от смартфона, с установленным на нем приложением, а также с интегрированных с ним фитнес-браслетов и умных часов.
- Отображение на карте контролируемых зон и активных устройств, на которых активирована функция «страт начала работы на строительной площадке», с возможностью перехода в карточку субъекта;

- Функцию разграничения прав доступа к просмотру информации;
- Функцию оповещения о срабатывании кнопки SOS или события приравненного к срабатыванию кнопки SOS, на устройствах с активированной функцией «старт начала работы на строительной площадке»;

Информационная среда должна представлять собой следующий набор программных решений:

- Операционная система: LinuxRedHat (актуальная версия);
- Система управления базами данных: PostgreSQL (актуальная версия);
- Сервер приложений ApacheTomcat (актуальная версия).

Информационная среда представляет собой набор трех подсистем (модулей), совместимых на программно-аппаратном и информационном уровне, и обеспечивает следующий функционал:

- Интерфейс пользователя (веб-интерфейс Информационной среды);
- Файловое хранилище информации, в том числе программный интерфейс информационного взаимодействия с государственными информационными системами Правительства Москвы;
- Интерфейс построения аналитической отчетности.

Дополнительные требования к Информационной среде со стороны Заказчика не предъявляются.

### **3.3. Требования к приложению для смартфонов на операционных системах Android/IOS**

Приложение должно обеспечивать следующие функции:

- Установку на смартфоны с операционными системами Android, IOS;
- Функцию регистрации в информационной системе, с внесением основных данных о субъекте (ФИО, Специальность, Работодатель, телефон для связи);
- Функцию согласия с обработкой персональных данных, передаваемых приложением в информационную среду.
- Функцию страта начала работы на строительной площадке;
- Функцию окончания работы на строительной площадке
- Сбор телеметрической информации с встроенных в смартфон средств и передача в Информационную среду с частотой раз в минуту через мобильный интернет и через Wi-Fi с применением средств защиты информации, предотвращающих возможность перехвата и расшифровки передаваемой информации.
- Идентификация пользователя с помощью встроенных в смартфон средств (отпечаток пальца, FaceID, пинкод, пароль).
- Функцию отправки сигнала SOS в информационную среду.
- Передача в информационную среду сигнала SOS при резком изменении показаний со встроенных в смартфон средств (показания акселерометра) в пределах контролируемой зоны.
- Установку прошивки на часы/браслеты, которые могут быть интегрированы со смартфоном.

### **3.4. Требования к микропрограмме для фитнес-браслетов и умных часов**

Микропрограмма должна обеспечивать следующие функции:

- Установку на фитнес-браслеты и умные часы;
- Сбор телеметрической информации и информации об активности с встроенных в устройства средств и передачу в приложение на смартфоне, с последующей передачей в информационную среду.

## **Задача 10**

### **Московский транспорт**

*Описание задачи «Разработка модуля интерактивной карты метро города Москвы с возможностью построения маршрута для мобильного приложения iOS и Android»:*

Исходные данные

#### **1.1.1 SVG файл с картой**

Файл содержит данные о линиях метро, остановках метро.

Файл содержит данные о станциях, линиях и пересадках.

Описание задачи

Участникам предлагается разработать интерактивную карту метро, которая будет доступна пользователю в мобильном приложении на платформах iOS и Android.

**Возможности карты:**

#### **1.1.2 Зумирование**

Возможность приблизить или отдалить карту метро.

#### **1.1.3 Просмотр информации об остановке**

Возможность нажать на любую остановку метро и просмотреть основную информацию о ней: название, название и цвет линии, а также любую другую информацию, которую придумают участники задания в открытых источниках.

#### **1.1.4 Построение маршрутов по карте метро в режиме offline**

Реализовать возможность поиска оптимального маршрута по карте метро для пользователя. Возможность построить маршрут должна быть доступна без использования интернета. Выбор алгоритма для построения маршрута определяется участниками.

Требования к UI/UX

Внешний вид модуля и механика его взаимодействия с пользователем определяется командой самостоятельно. Необходимо уделить внимание скорости и плавности работы интерфейса.

### Требования к технической реализации

Разработанное решение должно иметь возможность интеграции в существующее приложение, написанное на «нативных» языках (Swift/Kotlin). Часть модуля может содержать кроссплатформенную реализацию (например, модуль построения маршрута может быть написан на C++ и использован для обеих платформ), или полностью реализован на нативных языках платформ iOS/Android (Swift, Kotlin).