|  |  |
| --- | --- |
| 1 | Добрый день уважаемые коллеги!  Представляю вашему вниманию открытие спринта по разработке инструмента проверки и классификации адресов.  Следующий слайд, пожалуйста. |
| 2 | Приоритетной целью для нас является создание инструмента для автоматизированной проверки адресов.  Для достижения данной цели:   1. Мы исследуем все возможные способы проверки и классификации адресов, как во внутреннем, так и во внешнем контуре 2. А для улучшения пользовательского опыта и повышения эффективности работы с геоданными разработаем программное средство, реализующее вышесказанные функции.   Следующий слайд, пожалуйста. |
| 3 | В рамках подготовки к проверке «Выявление мошеннических схем по потребительским кредитам» были выявлены случаи, когда клиенты в кредитных заявках указывали места жительства или регистрации, фактически не являющиеся жилыми домами.  Классификация и выявление недостоверных адресов проводилась сотрудниками вручную, что занимало большое количество времени. Тогда и появилась потребность в создании автоматизированного инструмента, выполняющего данную функцию.  Следующий слайд, пожалуйста. |
| 4 | Для решения этой задачи мы предлагаем рассмотреть следующие пути:   1. В первую очередь необходимо выполнить нормализацию адресов, чтобы привести их в единый вид. 2. После этого выполняется проверка адресов во внутренних витринах банка, а также во внешнем контуре в источниках 2ГИС и Open Street Map. 3. На третьем этапе определяется тип здания: административное, образовательное, жилое и т д. 4. Полученные данные необходимо проанализировать и сравнить. 5. На основе всех вышеописанных шагов будет реализован финальный инструмент для обработки.   Следующий слайд, пожалуйста. |
| 5 | На слайде представлена структура решения.  Мы планируем использовать инструменты SGD-Stream-client, язык программирования Python, библиотеки: pandas, pyspark, selenium и requests для выполнения поставленных задач.  Следующий слайд, пожалуйста. |
| 6 | Для анализа точности классификации зданий была отобрана выборка из 10 000 адресов, которые нормализованы с помощью SGD-stream-client.  Несмотря на хорошую детализацию, результаты Open Street Map для жилых домов оказываются менее надёжными, так как значительное количество зданий классифицировано неверно.  Выявлены ошибки - переклассификация жилых домов как магазинов, церквей, гаражей и других встроенных объектов, что значительно снижает качество данных.  Итог: OpenSteetMap к использованию не рекомендуется.  Следующий слайд, пожалуйста. |
| 7 | **Рекомендация по выбору данных**  На основании результатов исследования мы **рекомендуем использовать внутренние витрины данных (ФИАС, 2ГИС)**:  Для улучшения точности и детализации возможно объединение данных ФИАС и 2ГИС, учитывая, что ФИАС может служить базовым уровнем для проверки присутствия зданий, а 2ГИС — для уточнения типов. |
| 8 | **Инструмент проверки и классификации адресов**  Алгоритм работы инструмента проверки и классификации адресов:   1. Пользователь выбирает **Excel-таблицу** с адресами. 2. Адреса **нормализуются** с помощью **SGD-stream-client**. 3. Нормализованные адреса записываются в виде таблицы в **КАП** 4. В КАП адреса **обогащаются данными** из **ФИАС** и **2ГИС**. 5. **Объединённый файл** с классификацией зданий передаётся пользователю для дальнейшего анализа |
| 9 | Мы планируем в будущем продумать вопрос интеграции данного инструмента в AI помощник ЦК по геоданным. |
| 10 | Команда спринта представлена на слайде. Куратор – Хайретдинова Рената. Готов ответить на ваши вопросы.  \*\*\*КОНЕЦ\*\*\* |