# SIpON no LK v4.002

# Формат исходных данных

После эксперимента с форматом файлов я вернулся к старому варианту, когда в первом столбце записана некая ключевая информация - идентификатор адреса или кадастрового номера (например, номер лицевого счёта). Это поле можно не заполнять, оно имеет значения только для последующего сопоставления полученной информации с рабочей базой данных. В качестве исходных данных для поиска используются столбцы со второго и далее. Первая строка – всегда заголовок.

Росреестр умеет искать по адресу, указанному почти в любом виде, однако, я счёт необходимым сделать минимальную предобработку исходных данных. Обработка включает два шага:

- если в начале строки встречается 6 цифр (индекс), они удаляются;

- удаляются некоторые мусорные данные (д., кв., им., ком.), а также буква «ё» заменяется на «е»; набор этих правил прописан во внешнем модуле *clear\_address.py* и может быть изменён на ваше усмотрение.

# Формат результатов.

При поиске по КН результат поиска выводится на тот же лист: столбцы с результатом дописываются справа от исходных данных в столбцы с 3 по 10. Информация собирается только для объектов со статусом «Актуально», для прочих – выводится только статус.

При поиске по адресу на первый лист рядом с исходным адресом записывается количество найденных ОН, либо «Информация не найдена», либо «Повтор адреса» в случае, если подряд идёт несколько одинаковых адресов.

Вся найденная информация записывается на лист «KNaddr#results». Кроме исходного адреса записывается обработанный адрес, по которому выполнялся поиск. Для удобства дальнейшей работы результаты раскрашиваются в разные цвета через один адрес.

К сожалению, сайт выдаёт в качестве результата поиска слишком много лишних объектов – например, меняет местами номер дома и квартиры, не учитывает корпуса и литеры номеров домов и т.д. – всего до 20 объектов, которые он считает подходящими. Разгребать всё это приходится руками. Кроме того, сайт игнорирует повторяющиеся элементы в строке поиска, например, «Ленинский район г. Ленинск ул. Майская» - подстрока «Ленинск» встречается два раза и будет проигнорирована, т.е. в качестве результата будет выдан, например, «Ленинский район, с. Заплавное, ул. Майская».

# Почему сразу 4 после 3 альфа:

- полностью переписал алгоритмы взаимодействия с сайтом, что привело к почти полному переписыванию кода;

- оба варианта поиска (по кадастровому номеру и по адресу) объединены в одну программу;

- поиск по адресу сразу же выбирает данные по найденным объектам;

- поиск по КН выполняется пакетами по 50 штук, в результате средняя скорость достигает 5-7 тысяч КН в час.

# Возможности

Запускать программу можно двумя способами: как скрипт (через IDLE или VisualStuio) и как самостоятельное приложение. При запуске из IDLE работа ведётся с браузером в обычном режиме, т.е. можно наблюдать процесс поиска. При этом работать на компьютере можно, но неудобно. В остальных случаях браузер запускается в скрытом режиме, после запуска можно свернуть окно программы и консоль вывода.

Можно запускать несколько копий программы, каждая будет выполнять поиск независимо от другой. Что скажет росреестр на такую нагрузку – не знаю.

Поиск можно почти в любой момент прервать, а затем продолжить с того же места.

Вид поиска – по кадастровому номеру или по адресу – определяется в момент открытия файла: если в ячейке B2 (первая строка первый столбец данных) находится кадастровый номер – включается поиску по КН, если строка не похожа на КН – включается поиск по адресу. Вид поиска отображается на кнопке.

Во врем работы выводится индикатор с количеством обработанных строк и средняя скорость (строк в час).

Поиск можно поставить на паузу – остановка произойдёт после окончания текущего цикла и может занять несколько секунд: при поиске по адресам остановка произойдёт после окончания обработки текущей строки (3-5 секунд), при поиске по КН – после окончания обработки пакета (20-30 секунд).

Нажатие в процессе поиска кнопок «Запустить Chrome» или «Обновить капчу» может привести к сбою, хотя в большинстве случаев не происходит ничего страшного – достаточно ввести новую капчу или запустить поиск ещё раз.

В крайне редких случаях, если закрыть программу через диспетчер задач или нажатием на крестик, происходит полная потеря файла. На такой случай в процессе работы каждые 10 минут сохраняется копия файла (*имя\_исходного\_файла.xlsx.backup*).

Капча выводится после предобработки: картинка делается чёрно-белой, удаляется некоторое количество визуального мусора.

# Copyright. Условия использования

## Начиная с версии 3.1 бота и загрузчика, а также с версии 4 программы SIpON

Идея, исполнение и программный код бота, загрузчика и библиотек, идущих с ними в комплекте, (всё вместе – КОД) принадлежат мне как автору.

Следующие права предоставляются:

- нескольким пользователям, писавшим мне на habr до 01.03.2021, без дополнительных условий;

- коллегам по работе без дополнительных условий;

- тем, кто уже заплатил за предыдущую версию какой-либо из указанных программ или за техническую поддержку или консультации по указанным программам;

- всем остальным на условиях «попробуй, если это окажется полезным, заплати автору любую сумму, которую сочтёшь адекватной полученной пользе»:

- вы можете изучать, анализировать и изменять КОД под свои нужды;

- вы можете использовать КОД в своей профессиональной деятельности, за исключением случаев, когда единственной целью использования является получение от третьих лиц платы за выписки из ФГИС ЕГРН, полученные при помощи КОДа.

Я отказываюсь от ответственности за последствия, вызванные применением вами КОДа. Если вы сомневаетесь или не уверены в результатах работы КОДа, напишите мне: я или развею ваши сомнения, или посоветую не использовать КОД. Также я не могу нести ответственность за совместимость КОДа с сервисами Росреестра.

Я не обязуюсь в той или иной мере развивать и поддерживать КОД.

По мере сил я буду оказывать техническую поддержку и помощь пользователям КОДа.

Если вы не согласны с чем-то, приведённым выше на этой странице, пожалуйста, не используйте КОД.

Со мной можно связаться:

- через https://github.com/0-6-1-7/rosreestr

- через https://habr.com/ru/users/0617/

- по электронной почте 0617@bk.ru

Автор будет рад получить «любую сумму» рублей на кошелёк Юmoney (YooMoney, бывший Яндекс.Деньги) 41001124561631

## Для более ранних версий и прочих скриптов, дата изменения которых в репозитории на github меньше 01.03.2021

Весь код свободен от любых ограничений, распространяется “as is” без всяких обязательств и ограничений. Код доступен по адресу https://github.com/0-6-1-7/rosreestr