

Пакетное геокодирование в QGIS

[Обсудить в форуме](#) Комментариев — 21

Эта страница опубликована в основном списке статей сайта по адресу <http://gis-lab.info/qa/rugeocoder.html>

Геокодирование в QGIS с помощью сервисов Google, Yandex, OSM Nominatim.

Содержание

- [1 Введение](#)
- [2 Подготовка к работе](#)
- [3 Подготовка данных](#)
 - [3.1 Преобразование таблицы в SHP файл](#)
 - [3.2 Подготовка адреса](#)
- [4 Геокодирование](#)
- [5 После геокодирования](#)
- [6 Ограничения в использовании геокодеров](#)
 - [6.1 Google](#)
 - [6.2 Yandex](#)
 - [6.3 Nominatim](#)
 - [6.4 openstreetmap.ru](#)

Введение

Одним из способов получения геоданных из адресной информации является операция геокодирования. Геокодирование позволяет получить координаты объектов, адреса которых известны. В связи с большим объемом адресных БД, эту операцию обычно производят сервисы геокодирования. Наиболее известные сервисы на данный момент - [Google Geocoding](#), [Яндекс.Карты Геокодирование](#) и [OSM Nominatim](#).

QGIS не имеет встроенной возможности пользоваться сервисами геокодирования. Но гибкая система расширений позволяет добавлять практически любые функции. Для выполнения пакетного геокодирования можно использовать расширение [RuGeocoder](#).

Подготовка к работе

Для выполнения геокодирования понадобятся:

- QGIS версии 1.6 и выше
- Расширение RuGeocoder (Присутствует в [официальном репозитории](#))
- Соединение с Интернетом (желательно быстрое)

При подготовке данных также может понадобиться расширенный калькулятор атрибутивных полей. В примерах будет использоваться расширение [FieldPyculator](#).

Процесс установки QGIS подробно описан в статье [Установка QGIS/GRASS с помощью OSGeo4W](#).

Более подробно об установке расширений можно прочитать [тут](#).

Подготовка данных

Для работы геокодер требует точечный слой с адресами, находящимися в одном или нескольких

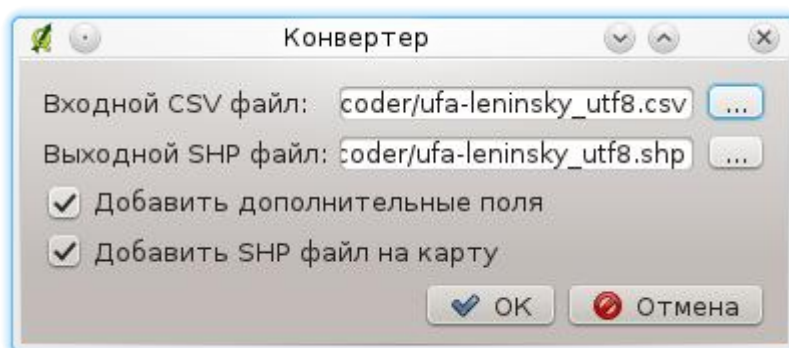
атрибутивных полях. Компонентная форма записи адреса (название города, название улицы, номер дома и т.д. находятся в отдельных полях) более предпочтительна и обычно позволяет добиться **значительно лучшего** результата при использовании любого геокодера.

Преобразование таблицы в SHP файл

Достаточно часто входные данные представлены в виде обычных таблиц. Наиболее удобный формат для обработки - CSV. С данным форматом могут работать практически все современные табличные редакторы: MS Excel, OpenOffice Calc и другие.

Для геокодирования таблицы её необходимо конвертировать в любой формат геоданных, которые можно редактировать в QGIS - SHP, sqlite, mapinfo и тд. Это преобразование можно сделать с помощью ogr2ogr (подробное описание можно прочитать в статье [Конвертация данных из CSV в SHP и обратно с OGR](#)). Но данный метод требует создания vrt файла и использование командной строки.

Для удобства RuGeocoder имеет встроенную функцию с подобным преобразованием. На панели "Модули" необходимо нажать кнопку "Конвертировать CSV в SHP" ^{CSV}_{SHP}. То же действие можно выполнить через меню "Модули" -> "RuGeocoder" -> "Конвертировать CSV в SHP". В открывшейся форме нужно выбрать входной CSV файл и указать путь, по которому будет сохранен результирующий SHP файл.



В результате конвертации, будет получен точечный shp файл с той же самой атрибутикой и нулевыми координатами. Если включена соответствующая опция, то к атрибутике будут добавлены пустые поля типа string: settlement, street, building_n и geocoded. Данные поля можно будет использовать в целях подготовки адреса для геокодера.

Кодировка данных выходного слоя будет такой же, как у входного. Если CSV файл был сохранен в UTF-8, то SHP будет иметь ту же кодировку. Возможно, после добавление слоя на карту у него будет выставлена неверная кодировка (зависит от настроек по умолчанию). Изменить кодировку можно в свойствах слоя на вкладке *Общие, Параметры источника*.

Рассмотрим для примера CSV файл ([Файл:RuGeocoder Ufa-leninsky utf8.zip](#)), содержащий адреса УИК Ленинского района г. Уфы. После конвертирования получаем следующий SHP файл: [Файл:Ufa-leninsky utf8 shp.zip](#).

Подготовка адреса

Для повышения качества геокодирования необходимо разбить адрес на составные части: населенный пункт, улица, строение. Каждая из частей должна находиться в отдельном поле (название населенного пункта можно не помещать в отдельное поле, а указать вручную, если все адреса в слое принадлежат этому населенному пункту). Если исходные данные изначально были в такой форме, то этот пункт обработки можно пропустить.

Как примером воспользуемся ранее подготовленным SHP файлом. Добавим его в проект и с помощью [FieldPyculator](#) подготовим необходимые поля. Как видно по полю `addr_vote`, адреса указаны единообразно, но содержат лишнюю информацию: номера школ, название организаций и тд. Выделим название улиц, используя следующее выражение для Pyculator:

```
value = <addr_vote>.split(',')1.replace(u'ул.', '').strip()
```

Для получения номеров строений используем следующий код:

```
value = <addr_vote>.split(',')2.strip().split(' ')0
```


В результате, получаем SHP файл ([Файл:Ufa-leninsky utf8 shp attr.zip](#)) с заполненными полями *street* и *building_n*. Исключением является одна запись (УИК с номером 187), в поле *street* которой попало название поселка. Данный недостаток легко исправить вручную.

Больше примеров обработки адресов можно найти на странице в вики:

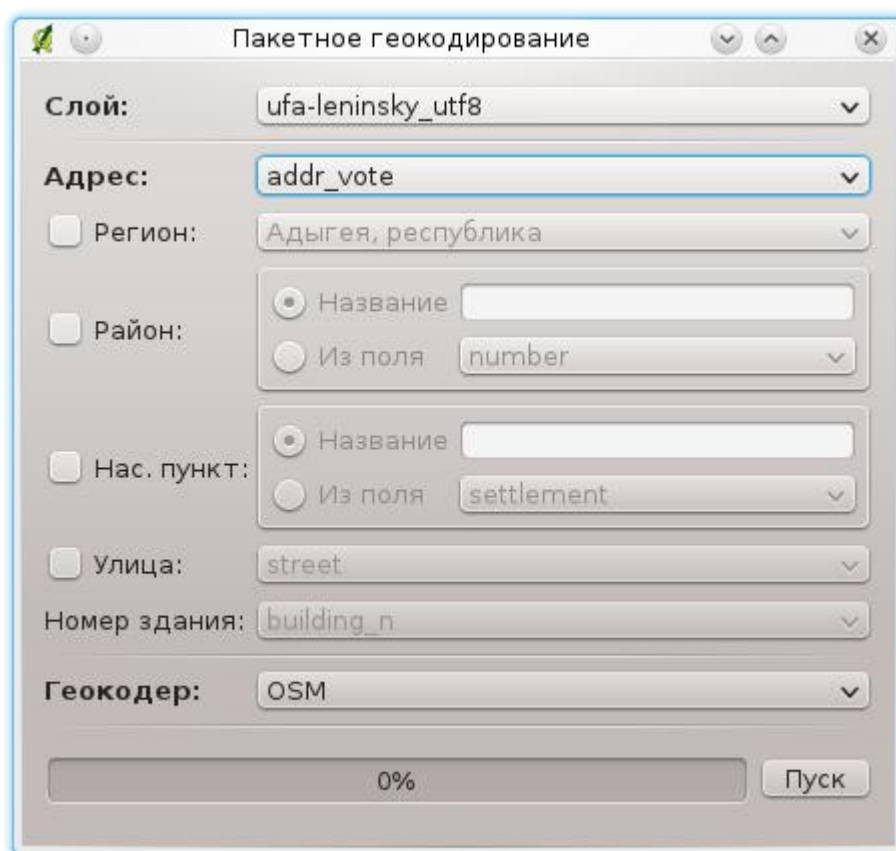
[Проект УИК Республики Башкортостан](#).

Геокодирование

После предварительной подготовки адресов можно приступить к самому геокодированию. Геокодировать можно любой точечный слой, редактирование которого возможно в QGIS (PostGIS, sqlite, MSSQL и т.д.). Слой для обработки должен быть добавлен в проект и открыт на редактирование.

Форма обработки вызывается нажатием кнопки "Пакетное геокодирование"  на панели "Модули". То же действие можно выполнить через меню "Модули".

В открывшейся форме прежде всего необходимо выбрать слой для геокодирования. Если адрес записан в одном поле, то достаточно выбрать это поле.



Для улучшения результатов геокодирования крайне желательно указать регион, включив соответствующий переключатель и выбрав из списка необходимый субъект. Если в слое присутствуют адреса из нескольких населенных пунктов, то необходимо выбрать поле, содержащее название поселения. Если же все адреса относятся к одному населенному пункту, то его название можно указать вручную. Аналогичная ситуация с районам.

Пакетное геокодирование

Слой: ufa-leninsky_utf8

Адрес: addr_vote

☒ Регион: Башкортостан, республика

☐ Район:
 ☒ Название
 ☐ Из поля number

☒ Нас. пункт:
 ☒ Название уфа
 ☐ Из поля settlement

☐ Улица: street

Номер здания: building_n

Геокодер: OSM

0% Пуск

Идеальным случаем является вариант, когда адрес компонентный. Включив переключатель "Улица", можно указать поле содержащее название улицы и поле содержащее номер строения.

Пакетное геокодирование

Слой: ufa-leninsky_utf8

Адрес: addr_vote

☒ Регион: Башкортостан, республика

☐ Район:
 ☒ Название
 ☐ Из поля number

☒ Нас. пункт:
 ☒ Название уфа
 ☐ Из поля settlement

☒ Улица: street

Номер здания: building_n

Геокодер: OSM

0% Пуск

Выбрав геокодер, можно запускать процесс обработки. Время выполнения геокодирования зависит от количества записей в слое, скорости соединения и загруженности выбранного сервиса. Геометрии записей, для которых геокодирование прошло успешно, будут изменены в соответствии с полученными результатами. В случае, когда геокодер не смог найти необходимое здание, будут возвращены координаты улицы, а при невозможности определить улицу - координаты населенного пункта. Для записей, адреса которых не

распознаны геокодером, будут записаны нулевые координаты. Также, если в слое присутствует поле "geocoded", то в него будет записан адрес, который вернул геокодер. Данное поле удобно использовать для анализа результатов обработки.

Если после геокодирования в слое много точек с нулевыми координатами (поле "geocoded" содержит много значений "Not found") или точки попадают в центр населенного пункта ("geocoded" содержит название поселения), то либо неудачно подобран геокодер, либо адреса подготовлены плохо и не могут быть распознаны.

В первом случае можно попробовать выбрать другой сервис из списка доступных. Следует учесть, что для больших и средних городов наиболее подходящим является Yandex. Для малых населенных пунктов чаще подходит OSM, хотя в некоторых случаях Yandex не отстает. Google хорошо справляется с задачей в основном в крупных городах.

Если смена геокодера не приводит к значительным изменениям результата, то следует приложить больше усилий для подготовки компонентного адреса.

После геокодирования

Результатом геокодирования является изменение геометрии. Координаты точек в таблицу не записываются. Если вам необходимо табличное представление координат, выберите в меню "Вектор->Обработка геометрии" пункт "Экспортировать / добавить поле геометрии".

Ограничения в использовании геокодеров

Коммерческие сервисы геокодирования (такие как Yandex или Google) имеют строгие лицензионные ограничения как на использование сервисов, так и на использование данных, полученных с помощью этих сервисов. **Плагин предполагает использование этих сервисов только в ознакомительных целях. Пользователь сам отвечает за соблюдение этих ограничений!**

Google

[Ограничение на использование](#)

Использование службы Google Geocoding API ограничено. Максимальное количество запросов геопозиционирования в день – 2500. Службу Geocoding API можно использовать только в совокупности с картой Google. Геокодирование без отображения результатов на карте не разрешается. Полная информация о допустимом использовании содержится в разделе [Ограничения лицензии на использование Maps API](#).

Yandex

[Пользовательское соглашение сервиса «API Яндекс.Карты»](#)

2.3.5. Геокодирование. Функция геокодирования может быть использована только для сайтов, использующих Сервис. Пользователь имеет право обрабатывать полученные координаты исключительно для отображения их посредством Сервиса. Количество обращений к функции геокодирования ограничено 25000 запросов для одного сайта в сутки.

Пользователь не имеет права:

2.3.6.5. Сохранять, обрабатывать и видоизменять полученные через Сервис Данные (включая результаты геокодирования), за исключением случаев временного хранения (кэширования) результатов геокодирования исключительно для целей улучшения функциональности и работоспособности Сервиса и только для использования в рамках возможностей, предоставляемых Сервисом.

Nominatim

[Политики использования сервиса Nominatim](#)

Обратите внимание, что данные, возвращаемые Nominatim, основаны на данных OpenStreetMap и попадают под действия лицензии [ODbL](#), хотя небольшие выгрузки могут быть использованы при условии честного использования (fair use/fair dealing).

openstreetmap.ru

В данный момент конкретной лицензии нет. Предполагается лицензионное соглашение, аналогичное Nominatim.

[Обсудить в форуме](#) Комментариев — 21

Последнее обновление: 2014-05-15 01:44

Дата создания: 10.09.2012

Автор(ы): [Евгений Никулин](#)