

Вы не представились системе

- [Войти](#)

Поиск

Адресный классификатор: ФИАС + Почта России

Материал из Data-in

Пространства имён

- [Статья](#)
- [Обсуждение](#)

Ещё

- [Ещё](#)

Действия на странице

- [Читать](#)
- [Просмотр кода](#)
- [История](#)
- [Экспорт в PDF](#)

Версия для печати больше не поддерживается и может содержать ошибки обработки. Обновите закладки браузера и используйте вместо этого функцию печати браузера по умолчанию.

Содержание

- [1 Общая информация о датасете](#)
- [2 История изменений](#)
- [3 Общее описание набора данных](#)
- [4 Структура набора данных \(Codebook\)](#)
- [5 Известные ограничения данных](#)
 - [5.1 Полнота данных](#)
 - [5.2 Описание проведенных проверок и их результатов](#)
- [6 Внесенные в набор данных изменения по сравнению с источниками](#)

- [7 Источники данных](#)
 - [7.1 Федеральная информационная адресная система \(ФИАС\)](#)
 - [7.2 Сведения портала kodifikant.ru о почтовых индексах и адресах России](#)
- [8 Ответственные за набор данных](#)

Общая информация о датасете

Набор содержит объединённые из двух источников (ФИАС и обработанные порталом kodifikant.ru данные Почты России) географические данные РФ с детализацией до дома.

Цель датасета — создание единого максимально полного классификатора географических данных, с помощью которого можно было бы осуществлять поиск и привязку географических объектов к официальным стандартизированным геоданным. Для осуществления этой привязки можно использовать библиотеку, доступную в [публичном репозитории на Github](#)

Единица наблюдения в датасете — адрес объекта недвижимости. Данные для каждого наблюдения приведены по следующим атрибутам: федеральный округ, субъект РФ, муниципальное образование, населенный пункт, локация, улица, дом.

Набор данных охватывает все регионы Российской Федерации. Всего набор содержит 25 569 964 наблюдений по 11 атрибутам.

Набор доступен для работы в формате CSV (кодировка: «UTF-8», разделитель: «;»)

История изменений

| Дата | Версия | Автор | Описание изменений |
|------------|--------|---|--------------------|
| 14.10.2021 | 1.0 | Валько Данила Валерьевич Глонин Константин Вячеславович Давыдов Николай Юрьевич | Документ создан |

Общее описание набора данных

| Атрибут | Значение |
|--|---|
| Общие сведения | |
| Полное наименование набора данных | Адресный классификатор: объединённые географические данные ФИАС и Почты России с точностью до дома |
| Сокращенное наименование набора данных | Адресный классификатор: ФИАС + Почта России |
| Краткое описание набора данных | Датасет, позволяющий осуществить поиск и привязку географических объектов к стандартизированным геоданным. Для осуществления привязки воспользуйтесь библиотекой по ссылке https://github.com/CAG-ru/geonorm |
| Краткое описание источника набора данных | 1) Федеральная информационная адресная система - хранит и обновляет (до 31.08.2021 г.) данные об адресных объектах, расположенных на территории РФ: субъектах РФ, районах, населенных пунктах, улицах, домах и др. 2) Обработанные порталом kodifikant.ru сведения из адресного справочника Почты России |

| | |
|---|---|
| Покрываемый временной период | Актуален на 14.10.2021 |
| Доступные форматы | CSV (кодировка: «UTF-8», разделитель: «;») |
| Размер набора данных | 3,2 ГБ |
| Тематика набора данных | Геоданные |
| Режим доступа | Открытый (у Вас есть возможность работать с данными на своем личном ПК, в том числе и скачивать данные) |
| Периодичность публикации и обновления | |
| Дата размещения в каталоге | 14.10.2021 |
| Периодичность обновления в каталоге | Ежеквартально |
| Дата последнего обновления набора данных в каталоге | отсутствует |
| Характер последнего обновления | |
| Дата следующего обновления набора данных | 14.01.2022 |
| Дополнительные сведения | Адресный классификатор: объединённые географические данные ФИАС и Почты России с детализацией до дома // ФИАС; kodifikant.ru; обработка: Валько Д.В., Глонин К.В., Давыдов Н.Ю., Инфраструктура научно-исследовательских данных, АНО «ЦПУР», 2021. Доступ: Лицензия CC BY-SA. Размещено: 14.10.2021. URL: http://www.data-in.ru/data-catalog/datasets/179 |
| Цитирование набора данных на русском языке | Address classifier: geodata of FIAS and Russian Post with detailing to the building // Federal Tax Service of Russia; kodifikant.ru; data-processing: Valko D., Glonin K., Davydov N., The Research Development Infrastructure (RDI), CAG, 2021. Access: License CC BY-SA. Posted: 14.10.2021. URL: http://www.data-in.ru/data-catalog/datasets/179 |
| For references (English) | |
| Ссылка на открытый репозиторий | отсутствует |
| Геоданные (пространственная привязка набора данных) | Есть привязка по адресам |
| Лицензия, под которой публикуется набор данных | CC BY-SA (Creative Commons) https://creativecommons.org/licenses/by-sa/4.0/ Допускается использование, копирование и распространение данных в научно-исследовательских, коммерческих и любых иных целях. При преобразовании материала разрешается распространять переделанные части материала на условиях этой же лицензии — CC BY-SA |
| Стандарт публикации | http://opendata.gosmonitor.ru/standard/3.0 |
| Параметры запросов API | Интеграция через API не предусмотрена |

Структура набора данных (Codebook)

Датасет представлен в виде плоской таблицы, содержащей 11 атрибутов, 25 569 964 наблюдений.

Формат данных: CSV (кодировка: «UTF-8», разделитель: «;»)

| Атрибут | Описание | Количество пропусков (NaN) (шт/%) | Единица измерения | ОКЕИ | Формат |
|-------------------|--|-----------------------------------|-------------------|------|--------|
| region | Уровень федерального округа | 0,0% | | | string |
| municipality | Уровень муниципального округа | 40,7% | | | string |
| municipality_type | Тип уровня муниципального округа | 40,7% | | | string |
| settlement | Уровень населенного пункта | 3,8% | | | string |
| settlement_type | Тип уровня населенного пункта | 3,8% | | | string |
| location | Уровень локации | 77,7% | | | string |
| location_type | Тип уровня локации | 77,7% | | | string |
| street | Уровень улицы | 16,7% | | | string |
| street_type | Тип уровня улицы | 16,7% | | | string |
| house | Уровень дома | 0,0% | | | string |
| source | Код источника данных об адресе 1 - ФИАС 2 - Почта России | 0,0% | | | int |

Известные ограничения данных

Полнота данных

Представлен наиболее полный набор актуальных адресных элементов ФИАС на 28 августа 2021 г. и адресов Почты России. Пропуски в атрибутах связаны с разреженным характером почтового адреса. В качестве базы данных для сравнения были выбраны наборы [OpenStreetMap](#), поддерживаемые открытым сообществом картографов и содержащие до 97% адресных объектов по большинству крупных городов России.

Описание проведенных проверок и их результатов

- Тест корректности размещения адресных элементов по атрибутам для стратифицированного тест-сета, собранного по ФИАС и Почте России (n=963). Общая точность с глубиной до 6 уровней (region, municipality, settlement, location, street, house) — 0.973; до 5 уровней (region, municipality, settlement, location, street) — 0.994.
- Есть незначительное число ошибок унаследованных от первоисточника ФИАС (см. подробнее по [ссылке](#))

Внесенные в набор данных изменения по сравнению с источниками

- ФИАС предусматривает 10 адресных уровней, их свертка в компактную структуру датасета производилась по схеме:
 - 1 - регион → region
 - 3 - район → municipality
 - 35 - уровень городских и сельских поселений (фактически не задействован)
 - 4 - город → settlement
 - 6 - населенный пункт → settlement
 - 65 - планировочная структура и др. территории (ГСК, СНТ, лагеря отдыха и т.п.) -> location
 - 7 - улицы → street
 - 75, 8, 9 (не задействованы)

Для каждого уровня создается колонка с типом адресного элемента *_type (например, settlement_type = деревня, поселок, ...).

Объекты на уровнях 3 (за исключением районов), 4, 6 раскладываются начиная с размещения поселения нижнего уровня в settlement и затем вверх по иерархии (например, поселения в г. Москва и городские округа указанные, как города размещаются в municipality, если в settlement уже указана деревня или посёлок).

Адреса Почты России извлечены html-парсером и перенесены в том виде, в котором они представлены в первоисточнике.

Источники данных

Федеральная информационная адресная система (ФИАС)

| Атрибут | Значение |
|---|--|
| Полное наименование источника данных | Федеральная информационная адресная система |
| Сокращенное наименование источника данных | ФИАС |
| Владелец источника данных | Федеральная налоговая служба (ФНС России) |
| Краткое описание источника данных | Федеральная государственная информационная система, обеспечивающая формирование, ведение и использование государственного адресного реестра (https://fias.nalog.ru/) Данные представлены в DBF и XML формате. |
| Ссылка на источник данных | https://fias.nalog.ru/DataArchive |
| Перечень используемых классификаторов | нет |
| Понятия, используемые в источнике | |

Методология и изменения методологии для источника данных Формат сведений об адресах, содержащихся в Государственном адресном реестре утвержден Приказом ФНС России от 13.05.2020 № ЕД-7-6/329@ <https://fias.nalog.ru/docs/%D0%9F-%D0%983490.doc>

Ссылки на методологию Документация и нормативные акты по ФИАС
<https://fias.nalog.ru/FiasInfo>

Сведения портала kodifikant.ru о почтовых индексах и адресах России

| Атрибут | Значение |
|--|---|
| Полное наименование источника данных | Сведения портала kodifikant.ru о почтовых индексах и адресах России |
| Сокращенное наименование источника данных | - |
| Владелец источника данных | Портал Бизнес-справочников kodifikant.ru |
| Краткое описание источника данных | Публичный клон адресных данных Почты России в html-структуре |
| Ссылка на источник данных | https://index.kodifikant.ru/ru/ |
| Перечень используемых классификаторов | нет |
| Понятия, используемые в источнике | |
| Методология и изменения методологии для источника данных | |
| Ссылки на методологию | |

Ответственные за набор данных

| № | ФИО ответственного лица | Электронная почта ответственного лица |
|----|--------------------------------|--|
| 1. | Валько Данила Валерьевич | d.valko@data-in.ru |
| 2. | Глонин Константин Вячеславович | k.glonin@data-in.ru |
| 3. | Давыдов Николай Юрьевич | n.davydov@data-in.ru |

Источник —

https://wiki.data-in.ru/index.php?title=Адресный_классификатор:_ФИАС_%2B_Почта_России&oldid=364

Категория:

- [Описания датасетов](#)

Навигация

Навигация

- [Главная](#)
- [База знаний](#)
- [Описания датасетов](#)
- [Дизайны исследований](#)

ИНИД

- [Каталог датасетов](#)
- [Новости](#)
- [Исследования](#)
- [Мероприятия](#)
- [Лицензия Продвинутого режима](#)
- [Лицензия обезличенных датасетов без возможности скачивания](#)
- [Лицензия обезличенных датасетов с возможностью скачивания](#)

Вики-инструменты

Вики-инструменты

- [Служебные страницы](#)

Инструменты для страниц

Инструменты для страниц

Инструменты страницы участника

Ещё

- [Ссылки сюда](#)
- [Связанные правки](#)
- [Постоянная ссылка](#)
- [Сведения о странице](#)
- [Журналы страницы](#)

Категории

Категории

- [Описания датасетов](#)



- Эта страница в последний раз была отредактирована 25 октября 2021 в 11:23.
- [Политика обработки персональных данных](#)
- [Пользовательское соглашение](#)
- [Правила использования информации, размещенной на Платформе ИНИД](#)

