

База географических названий GNS

[Обсудить в форуме](#) Комментариев — 0

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

Сводные данные и примеры для скачивания

[GEOnet Names Server \(GNS\)](#) - свободно распространяемая база данных имён географических объектов. Используется Федеральным правительством США, но "варианты названий тех или иных объектов, а также связанные с ними данные могут не соответствовать точке зрения правительства". Ограничения по использованию данных отсутствуют.

Каждая запись базы представляет собой текстовую строку, в которой описаны помимо всего прочего: широта/долгота географического объекта, его название в кодировке Unicode UTF-8 и классификационный признак типа объекта (реки, горы, населенные пункты и др.)

География покрытия включает в себя весь мир, за исключением США и Антарктики. GNS содержит порядка 4 миллионов объектов и приблизительно 5.5 миллионов имен. Ежемесячно обновляется в среднем по 20 000 записей данной базы. Последнее обновление произведено 20 апреля 2009 года. Используемая система координат - WGS84, координаты приближительны и предназначены только для целей поиска.

Получение данных

Полную версию базы GNS на весь мир можно скачать со страницы: <http://earth-info.nga.mil/gns/html/namefiles.htm>, воспользовавшись гиперссылкой, "Click here to Download a single...". Имя загружаемого архива соответствует дате последнего обновления базы. Архив занимает порядка 300 Мб, несжатые данные - около 900 Мб.

Там же можно получить данные по каждому государству в отдельности в текстовом формате с разделителями (разделитель - символ табуляции). Для этого необходимо выбрать требуемую страну (поле "Country Name") и перейти по гиперссылке "Reading Order". Ссылка для скачивания данных на территорию РФ: <http://earth-info.nga.mil/gns/html/cntyfile/rs.zip>.

Обработанные данные

Получить подготовленные данные можно выбрав одну из ссылок в таблице. Если вы работаете в ГИС скорее всего вам нужны shape-файлы в кодировке UTF-8. Нажав на дату, можно посмотреть историю обновлений.

Наличие обновлений проверяется ежедневно, непосредственно обновление происходит по мере обновления БД (примерно один раз в неделю).

На официальном сайте GNS расположены также инструменты [текстового поиска](#) объектов и [графического поиска](#) (в случае использования браузера Opera появляется сообщение "Invalid browser detected", с другими браузерами проблем не замечено) с использованием картографической подложки.

Поля базы данных

[Подробнее](#) на языке оригинала.

Поле	Описание поля	Тип поля
------	---------------	----------

	<p>Код шрифта, определяющий таблицу символов, используемую в полях SHORT_FORM, Generic и Full_Name (см. Character Set Conversion Table):</p> <p>1 = Америка/Западная Европа; 2 = Восточная Европа; 3 = Африка/Средний Восток; 4 = Россия/Центральная Азия; 5 = Азия/Тихоокеания; 6 = Вьетнам.</p>	
RC		Числовой
UFI	Уникальный идентификатор объекта. Одному объекту может соответствовать несколько записей. В этом случае поля UFI таких записей будут совпадать.	Числовой
UNI	Уникальный идентификатор названия объекта.	Числовой
LAT	Широта объекта в десятичных градусах ($\pm dd.dd...$). Для обозначения южной широты используется знак "-".	Числовой
LONG	Долгота объекта в десятичных градусах ($\pm dd.dd...$). Для обозначения западной долготы используется знак "-".	Числовой
DMS_LAT	Широта объекта в формате ГМС (ггммсс). Для обозначения южной широты используется знак "-".	Числовой
DMS_LONG	Долгота объекта в формате ГМС (ггммсс). Для обозначения западной долготы используется знак "-".	Числовой
MGRS	Военная система координат.	Текстовый
JOG	Номенклатура листа Joint Operation Graphics, 1:250000	Текстовый
FC	<p>Класс объекта. Девять основных типов:</p> <p>A = Административный район P = Населенный пункт V = Растительность L = Местоположение или область U = Подводный объект R = Улицы, шоссе, дороги, железные дороги T = Объекты рельефа H = Гидрографические объекты S = Точечные объекты</p>	Текстовый
DSG	Код обозначения объекта. Двух или пятисимвольный код используемый для обозначения типа объекта GNS. В перечень кодов входит более 600 типов объектов.	Текстовый
PC	Классификатор населенных пунктов. Числовая шкала, обозначающая относительную важность населенного пункта (от 1 - высокая до 5 - низкая). Значение NULL соответствует неустановленной важности.	Числовой
CC1	Основной код страны. Двухсимвольный код уникальным образом идентифицирующий геополитическую единицу (страну, подчиненную территорию и т.п.).	Текстовый
ADM1	Код административного деления первого уровня. Двухсимвольный буквенно-числовой код стандарта FIPS 10-4 описывающий первый уровень деления геополитической единицы, например штат (state) в США.	Текстовый
ADM2	Код административного деления второго уровня. Название единицы второго уровня деления геополитической единицы, например округ (county) в США.	Текстовый
POP	Численность населения.	Числовой
ELEV	Высота в метрах.	Числовой
CC2	Дополнительный код страны. Двухсимвольный буквенный код стандарта FIPS 10-4 уникально описывающий код страны объекта Geoname если он отличен от основного кода страны этого объекта (международные объекты).	Текстовый
NT	<p>Тип имени:</p> <p>C = Обычное название N = Название объекта в стандарте BGN</p>	Текстовый

NS = Название объекта в стандарте BGN в нероманском написании

P = Временное название

PS = Временное название в нероманском написании

H = Историческое название

HS = Историческое название в нероманском написании

D = Непроверенное название

DS = Непроверенное название в нероманском написании

V = Альтернативное название

VS = Альтернативное название в нероманском написании

LC	Код языка в стандарте ISO 639-3	Текстовый
SHORT_FORM	Сокращенная форма полного названия.	Текстовый
GENERIC	Описательная часть полного названия (гора, река, море).	Текстовый
SORT_NAME	Форма полного названия позволяющая осуществлять корректную сортировку в режиме географического справочника (gazetteer). Все символы заменены на латиницу в верхнем регистре, числа заменены на буквы в нижнем регистре (0-9 = a-j), пробелы удалены, запятые заменены на пробелы. Замена производится согласно Character Set Conversion Table Char Sets 20060703.pdf .	Текстовый
FULL_NAME	Полное название объекта. Представлено в порядке чтения, то есть имеет место запись вида "Гора Эверест", вместо "Эверест, Гора", как хранится в базе данных.	Текстовый
FULL_NAME_	То же, что и FULL_NAME, но все символы заменены на латиницу, регистр и пробелы сохранены. Замена производится согласно Character Set Conversion Table	Текстовый
MODIFY DATE	Дата добавления или изменения объекта (ГГГГ-ММ-ДД).	Дата

Недостатки

- База GNS построена по избыточному принципу, что является плюсом, так как информация не удаляется, но это значительно усложняет работу. Как только то или иное название помещено в базу, оно никогда не удаляется, за исключением очевидных повторений. Поэтому в базе присутствует большое количество вариантов имен объектов, которых на настоящий момент не существует.
- Географические координаты объектов зачастую очень неточны, расхождения могут достигать 2-3 км.
- Присутствует некоторое количество некорректных или просто неверных записей: некоторые места, описанные в базе GNS как принадлежащие одной стране, фактически располагаются совершенно в другой.

Ссылки по теме

- [База географических названий GeoNames](#)
- [Перевод названий населенных пунктов VMap0](#)
- http://earth-info.nga.mil/gns/html/gis_countryfiles.htm
- [GNS на вики OpenStreetMap](#)

[Обсудить в форуме](#) Комментариев — 0

Последнее обновление: 2014-05-15 01:36

Дата создания: 27.05.2009

Автор(ы): [Максим Дубинин](#)