

3geonames dot org

An open source Geocoding system for the simple communication of locations with a resolution of 1 m



ERVIN RUCI

I hit my laptop's keyboard repeatedly for fun and profit

Geocode.xyz

BRUSSELS-VOT-SHOOV

3071531887023

50.812375,4.38073

Fosdem Université libre de Bruxelles Campus du
Solbosch Avenue Franklin D. Roosevelt 50 1050
Bruxelles Belgium

Location Codes

From XY to Z

M L S, Geohash, Mapcodes, Plus codes, O P C, N A C, XADDRESS, What3words, Zippr, MapTags, OkHi, Geokey, FB ...

Address Codes

Alphanumeric string sets created by humans for communicating locations with other humans.

FRANKLIN ROOSEVELT	Brussels
Franklin Rooseveltlaan	Brussel
FRANKLIN ROOSEVELTLAAN	Brussels
Avenue Franklin Roosevelt	Ville de Bruxelles

Permutations e.g.,

Avenue Franklin D. Roosevelt 50 1050 Bruxelles Belgium

or

Avenue Franklin Roosevelt - Franklin Rooseveltlaan 50 1050 Brussel Belgium

or

Avenue Franklin D. Roosevelt 50 1050 Brussels Belgium

[https://geocode.xyz/Avenue Franklin D. Roosevelt 50 Brussel Belgium](https://geocode.xyz/Avenue%20Franklin%20D.%20Roosevelt%2050%20Brussel%20Belgium)

BRUSSELS-AAX-MONTESE / BRUSSELS-NILWK / 50.81136,4.38176

Geocode

A hashing function for locations.

$G\{\text{Latitude}, \text{Longitude} + \text{Geo Names}\} = \text{Geocode}$.

Some Geocode Attributes



FREE

as in Latitude,Longitude. (Eratosthenes
circa 3rd century BC)



SHORT

Optimal location encoding.



SPATIAL LOCALITY

2D <-> 1D



MEMORIZABLE / DISTINCT

For Humans



UNIQUE

One to One



DETERMINISTIC

Must be generated offline.

→ **Free**

And Open Source:

<https://github.com/eruci/geocode>



Short

50.81237, 4.3807 → SOLBOSCH-NILKP → BRUSSELS-VOT-SHOY

And even shorter still, by using the Acronymize extensions:

BRU-VOT-SHO

The extensions work by shortening the major location names first up to two letters, eg Los Angeles -> LA, Brussels -> BRU

↖ Spatial Locality

Alphanumeric Geocodes/Triple Name Geocodes at borderline areas will share most of the significant bytes/geonames.

```
45.00001,-64.36000 -> EHB105754C -> HALIFAX-GAZAH-DOMOU  
44.99999,-64.36000 -> EHB1056SH4 -> HALIFAX-GAZAH-NDITI
```

This solves borderline issues of the geohash algorithm.

Equivalent Geohashes are:

```
45.00001,-64.36000 -> f840p2n2p3  
44.99999,-64.36000 -> dxfpzryrzq
```

Although these points are only 1 meter apart. (see <http://geohash.org/f840p2n2p3> and <http://geohash.org/dxfpzryrzq>)

❑ Memorizable / Distinct

■ BRUSSELS-VOT-SHOY

BRUSSELS-NILKP → BRUSSELS-VOT-SHOY

Geonames have Levenshtein distance of at least 2.

Phonetic distance of at least 1 (Metaphone Algorithm).

■ I find it easier to remember than

R96J+X7 Brussels, Belgium

■ or ///luring.garage.asserts

■ or u1515djgw6j

 **Unique**

50.8,4.38 → BRUSSELS-PUNO-VACY
BRUSSELS-PUNO-VACY → 50.8,4.38

50.8,4.38 → BRUSSELS-NISLK
BRUSSELS-NISLK → 50.8,4.38

50.8,4.38 → 8A473WX5EX
8A473WX5EX → 50.8,4.38

Geohashes "u150gxv4" and "u150gxv5" (and others) all decode to (50.8,4.38).



Deterministic

```
use Geo::Code;

my $g = Geo::Code->new();
my $geocode = $g->geocode(lat=>52.52699,lon=>13.40521);

my $xy = $g->geocode(gc=>$geocode);
my ($lat,$lon) = @{$xy};
```

API

api.3geonames.org

 AWS



aws.amazon.com/marketplace/pp/B07LB39JBB



Reverse geocode

<http://api.3geonames.org/50.812375,4.38073>

```
<elevation>99</elevation>
<timezone>Europe/Brussels</timezone>
<city>Brussels</city>
<name>Université Libre De Bruxelles / Campus Solbosch</name>
<prov>Flanders</prov>
<region>Provincie Antwerpen</region>
<state>BE</state>
<altgeocode>SOLBOSCH-NILKP</altgeocode>
<distance>0.075</distance>
<geocode>BRUSSELS-NILKP</geocode>
<geonumber>3071531887023</geonumber>
<threegeonames>BRUSSELS-VOT-SHOOY</threegeonames>
```



Reverse geocode

<http://api.3geonames.org/50.812375,4.38073.json>

```
{ "threegeronames" : "BRUSSELS-VOT-SHOY",
  "geonumber" : "3071531887023",
  "nearest" : {
    "distance" : "0.075",
    "timezone" : "Europe/Brussels",
    "elevation" : "99",
    "region" : "Provincie Antwerpen",
    "name" : "Université Libre De Bruxelles / Campus Solbosch",
    "state" : "BE",
    "city" : "Brussels",
    "prov" : "Flanders",
    "altgeocode" : "SOLBOSCH-NILKP" }
}
```

 **Cloud Performance**

'

100 to 200 reverse geocodes per second on a
T2.micro

1 vCPUs, 1 RAM (GiB)

Under the hood - The Algorithm

Divide Earth into 510100 'Hilbert' simple polygons → each polygon into squares → use skiplist data structures to name each square.

Each polygon is about 1000 km², each square is about 100 m² (Hilbert) or 1 m² (triple geoname version - Z-order)

(Overlaps make discontinuities less likely - necessary for accurate reverse geocoding at a small performance penalty)

No Database is needed for 3 geonames code.

Polygons

WAITANGI-USAKOS-IDIA

Shortening a 3 geoname code into a 2 or 1 geoname code produces simple polygons.

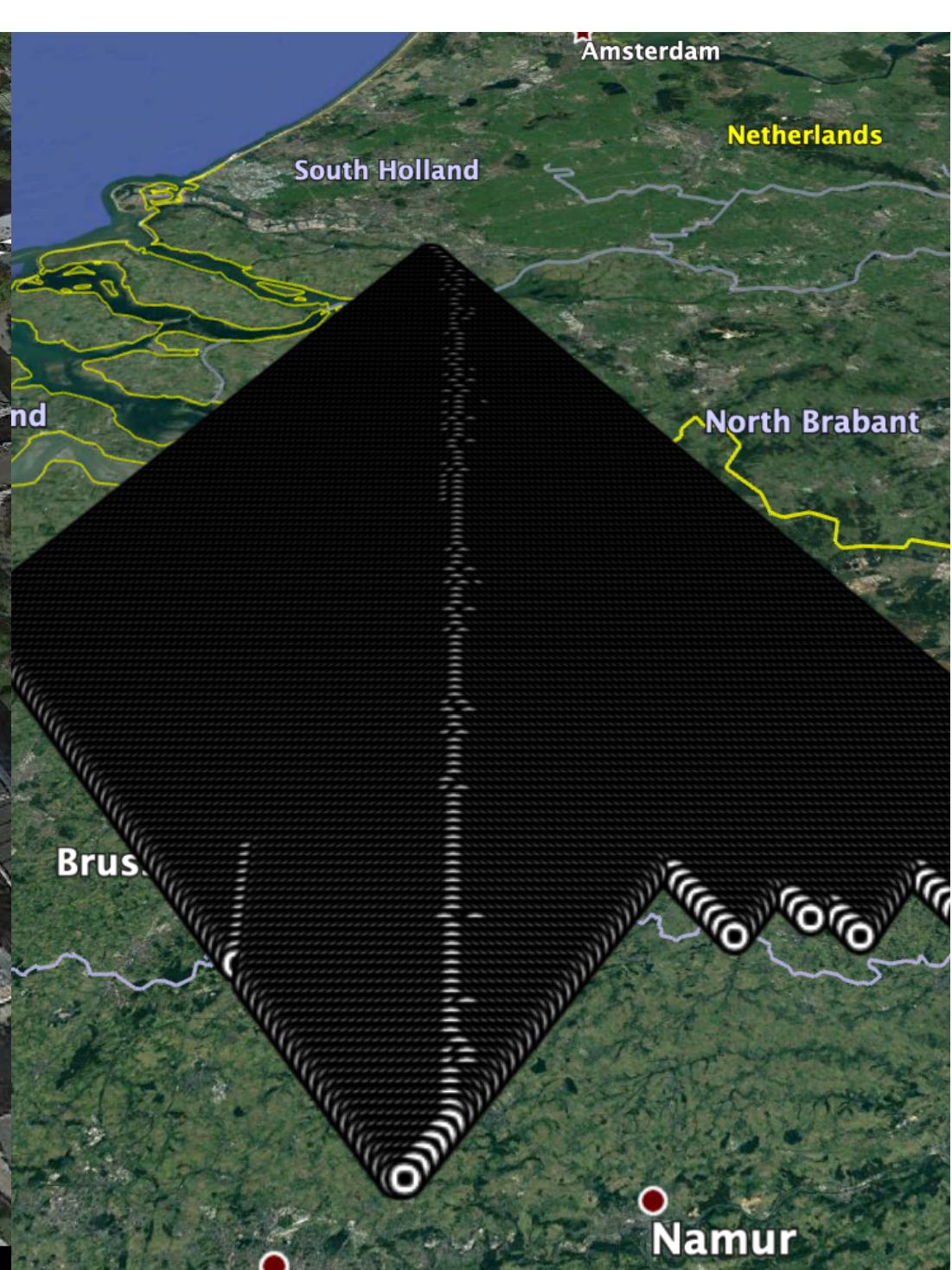
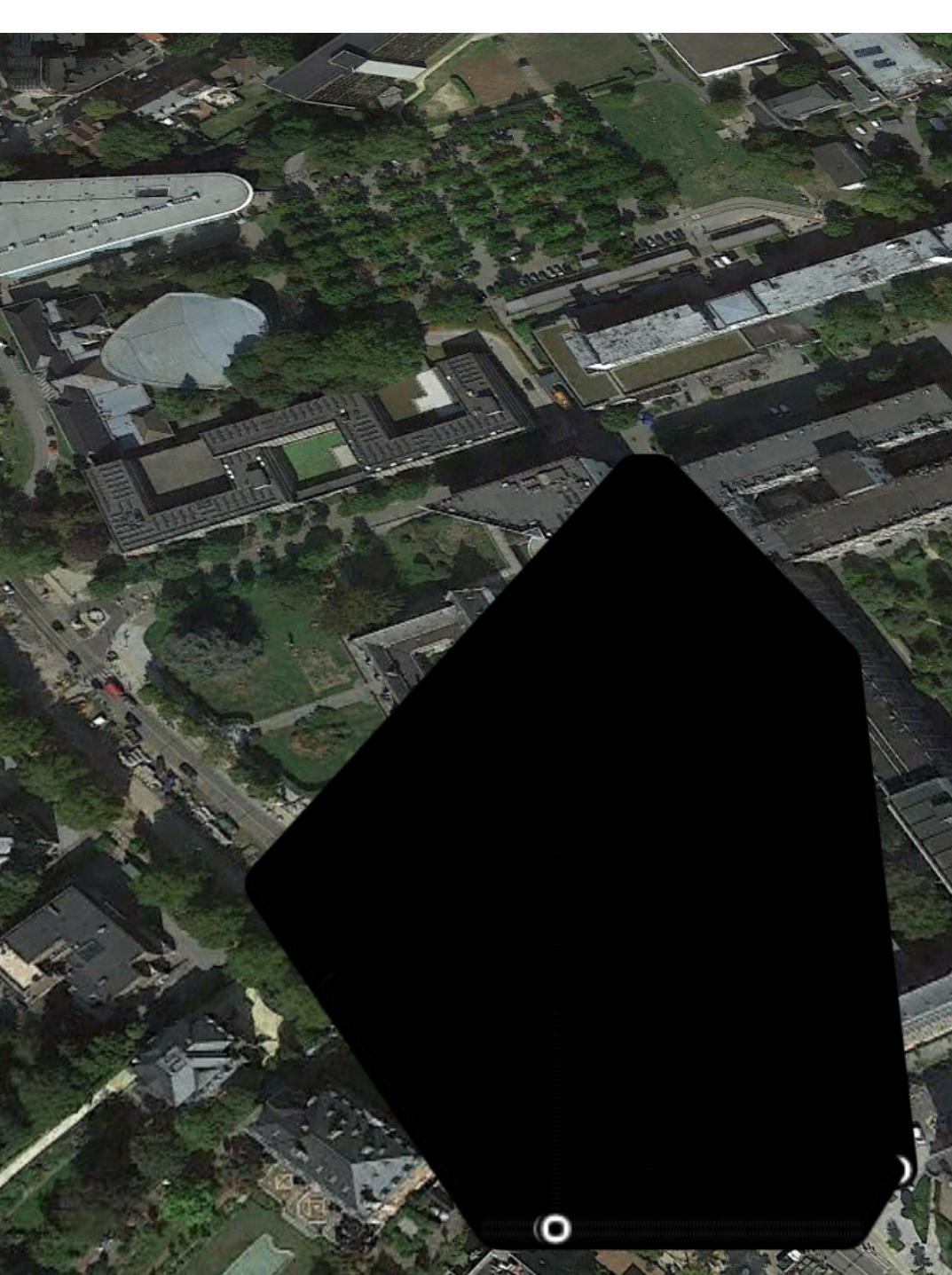
WAITANGI approx. 21,403 km²

WAITANGI-
USAKOS approx. 0.1463 km²

WAITANGI-
USAKOS-IDIA approx. 1 m²







DIRTY-WORDS

I used 146300 geonames from open sources.

Go ahead and upload your own word lists.

Considerations

There are 658.84 trillion latitude,longitude points limited to the 5th decimal, if we allow latitude range to be [90.00000,-90.00000] and longitude [180.00000,-180.00000].

(hence around 146300 geonames are needed)

Latitude and Longitude or Longitude and Latitude

- The fourth decimal place is worth around 10 m: it distinguishes small houses from each other. (Hilbert)
- The fifth decimal place is worth around 1 m: it distinguishes small trees from each other. (Z-order)

Nearby codes and bearing

GEONAME-AAAAA is roughly southwest of
GEONAME-ZZZZZ

(Latitude indicates north-south bearing, longitude east-west;
when they are combined in a single number, the movement of
the number is northeast-southwest)

Nearby codes and bearing

Also

BRUSSELS-VOT-SHOOY is roughly southeast of

GENT-VOT-SHOOY

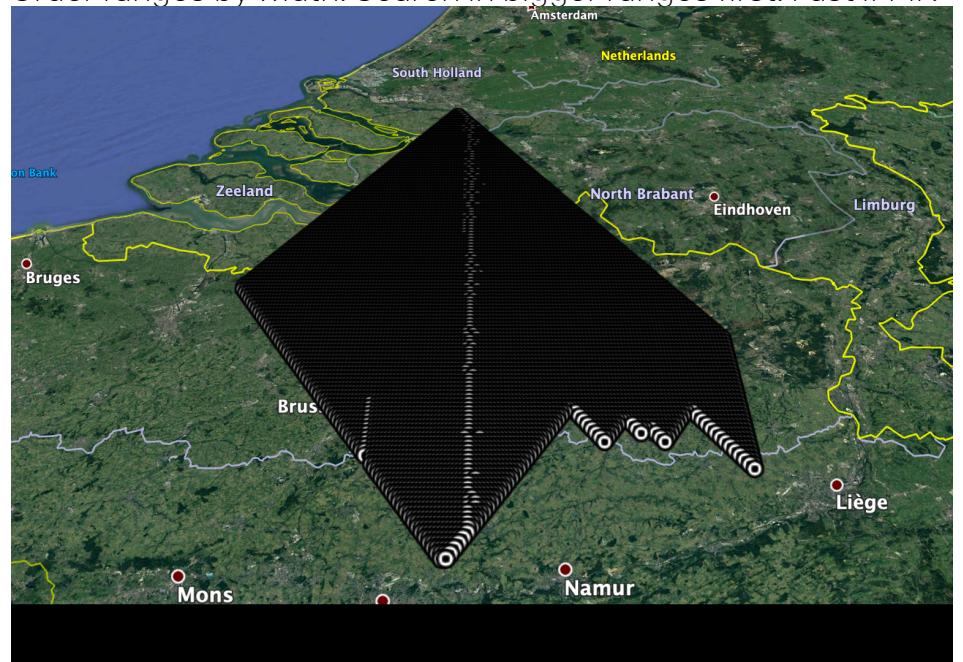
For those with knowledge of local geography

A faster PiP Algorithm using geocodes

Point in Polygon (PiP) queries are very fast with geonumber ranges.

- 3071532023036 .. 3071326459900
- 3071217416551 .. 3071149910022
- 3071130124612 .. 3071109992815
- ..

Order ranges by width. Search in bigger ranges first. Fast if PiP.



⚡ O(1) expected time complexity. O($\log n$) worst case.

Geocodes and Geoaliases

BRUSSELS-VOT-SHOOY

BRUSSELS-NILKP

50.812375,4.38073

SOLBOSCH-NILKP

3071531887023

Recap

An open triple geoname geocode for the world.



Resolution

1x1 & 10x10.



Hilbert space of 658.84 trillion squares

latitude range of [90.00000,-90.00000]

and longitude range of [180.00000,-180.00000]



3GeoName / Geocode

A unique 3 geoname code for each roughly $1 \times 1 \text{ m}^2$ cell and a unique locality context geocode for each $10 \times 10 \text{ m}^2$ cell.

Other Attributes

In addition to being precise, geocodes are also:



Fast

Just flipping bits.



Worldwide

No discontinuities. Even at the 180th meridian.

-16.9074,+179.9999 : WAIYEVO-YNWFJ

-16.9074,+180.0000 : WAIYEVO-YNWSK

-16.9074,-180.0000 : WAIYEVO-YNWSK

-16.9074,-179.9999 : WAIYEVO-YNWSJ



Extendible/Customizable

Use custom wordlists without altering the code.

Performance numbers

Reverse geocode about 10 million points a day on a minimal instance with 1 core and 1 GB of RAM.

Scale up to 200 million points a day on a server with 24 cpu cores.



Characteristics

CPU intensive (mostly number crunching.)



Data

12 million place names from Geonames.org, Openstreetmap, geonames.nga.mil and GADM.

Customizations / ToDo

Custom Geoname lists

Greater Levenshtein/Phonetic distance for better error correction.

Adaptive Hilbert polygons

Variable size polygons for remote areas

Other languages

Translated geonames based on geonames.org

Demo

Questions?