Mapcode:
A Public Location Reference Standard

Distribution courtesy of the Mapcode Foundation, The Netherlands
Version 1.042
The latest version of this document can be found on www.mapcode.com

Mapcodes

People need to find other people all the time. The tasks of locating a destination, and finding the best way to get there, become more relevant every day. Unfortunately, a large part of the world population has no address. In India alone, well over half a billion people live in houses that have no street name. Millions of man-hours are lost every day trying to locate or deliver goods to people and businesses based on business cards like this:

> **Amri Patel, *consultant***
> New Sun Technologies, District 14, Pune
> Tel. +23984982217
>
> Three miles east along the airport road, take second
> left past the golden statue, in the fourth street left
> ask final directions at baker across from white house.

Figure 1. Typical Indian Business Card

As we all know, even *having* an address is only an initial step in locating a position. Knowing someone lives on Queens' Road 123 in Brunnock still requires you to find out where Brunnock is, and where the Queen's Road is, and where number 123 is.

The mapcode system was designed as a free, brand-less, international standard that allows any location on the surface of the Earth to be represented by a short, easy to recognize and remember "code", usually consisting of between 4 and 7 letters and digits.

For example, in The Netherlands, the mapcode "49.4V" uniquely represents a particular door of a particular building in the center of Amsterdam. Any other location on Earth – a small drinking well in the Saharan desert, the back door of a bar on one of the Kiribati islands, a dwelling in a slum near Mumbai airport – has such a short code, which uniquely identifies that location and will bring you to within a few meters of it.

Key ideas underlying the mapcode system

In essence, the mapcode system defines a way to convert between the latitude and longitude of a location (such as measured through the GPS satellite system) and a short code.

We managed to make the code especially short based on six key ideas.

1. Codes only need to be accurate enough for public, every-day use. At the human scale, when you are within a few meters of a destination, you are “there”.

2. Codes can use letters as well as digits. There are enough different combinations of 9 letters and digits to uniquely identify every square meter on the surface of the Earth.

3. People live within a “country context”. Addresses seldom include a country name. Unless clearly stated otherwise, they can safely be assumed to be in “this” country. Codes in a particular territory can be designed to be much shorter. For example, every location in The Netherlands can be uniquely specified with at most 6 letters[1]. Even in a large country like India, at most 7 letters are needed.

4. A conversion need not just be based on formulae. A respectable amount of data may be used as part of an algorithm[2].

5. Codes need not all be the same length.

6. Short codes are reserved for areas where the population density is high. For example, every locations in The Netherlands can be represented by a unique 6-letter mapcode. However, the majority of the population is concentrated in a dozen urban areas totaling less than 3,000 km^2 (1,800 square miles). By reserving the 5-letter mapcodes for these areas, half the *population* of The Netherlands can be found using 5-letter codes. *On average*, mapcodes for relevant locations in The Netherlands are thus closer to 5 letters than to 6 letters. In fact, a significant percentage of the population can be found in 100 km^2 of city centers, which can be covered by 4-letter mapcodes.

[1] In most of this document, the term “letters” is usually an abbreviation for “letters and digits”.
[2] This was less obvious when the mapcode system was developed back in 2001. In today’s world, when cheap chips produce polyphonic music inside Christmas cards, a 100 kilobytes of data as part of a conversion should raise no eyebrows.

Based on these ideas, we envisaged a system that would work anywhere on Earth, and would (on *average*) provide the shortest *possible* codes to everybody. By using massive amounts of data gathered over the past 30 years by mapping company Tele Atlas (merged into the TomTom Group in 2008), a data table was constructed that

- defines mapcodes for every "context" on Earth, (each of the roughly 200 countries and nations on Earth, each of the 100 overseas dependencies, and each of 240 (semi-)autonomous states, provinces, republics and oblasts);
- defines the most densely populated areas in each of these territories to reserve the shortest mapcodes for.

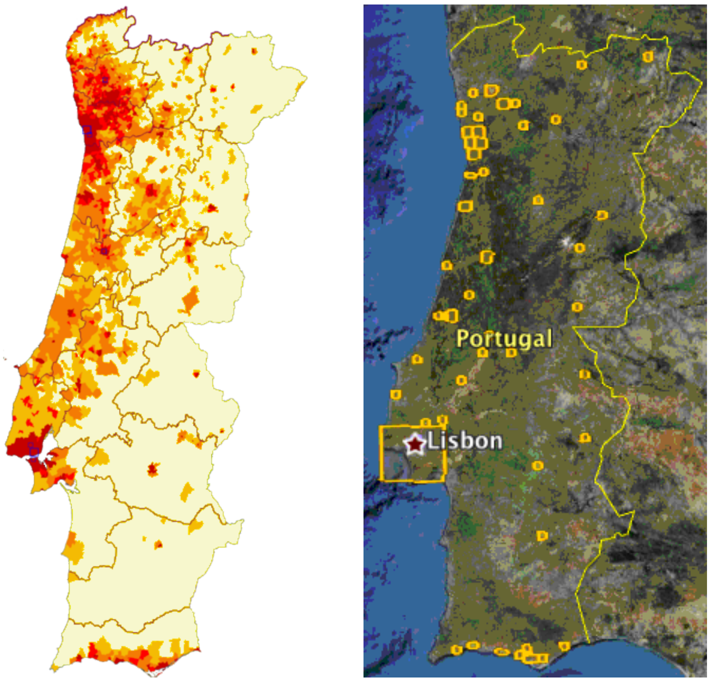

Figure 2. A population density graph (left) and the mapcode structure of Portugal (right); the orange rectangles specify areas that are covered by 5-letter mapcodes. The rest of the country (including Madeira and The Azores overseas) is covered by 6-letter mapcodes.

Furthermore, algorithms were constructed that

- efficiently and unambiguously convert between latitude/longitude coordinates and mapcodes;
- require only 32-bit integer arithmetic;

Finally, a system was developed that adheres to all the design ambitions we defined for an encoding scheme intended for free, international, public, every-day use. These are ambitions are presented in the next section.

Design ambitions for the mapcode system

We identified ten ambitions, each of which could make or break the usefulness of the whole mapcode system. As usual, some ambitions contradict, but most strengthen each other.

1. Mapcodes are short

The longer a code gets, the more awkward it becomes to use, the more difficult to remember, the more easy to make mistakes in copying or using, the less benefit it offers over more elaborate descriptions. Many other benefits depend on mapcodes being short. If length was not a problem, we might as well put our longitude and latitude on business cards and address labels.

As explained in the previous section, by reserving short codes for the most densely populated areas, mapcodes for *addresses (i.e. places that people need to visit)* are *on average* as short as theoretically possible, within the limits we set for ourselves. One such limit is to use only rectangular areas in our definitions. Another is to limit the size of the "context" data table to about 16,000 records. Probably the most relevant design choice was the choice of character set. For example, had we allowed both capital and lowercase letters, 3 times as many people could have 5-letter codes instead of 6-letter codes. Why this would have been a bad idea is explained a bit further on.

Figure 3. Manhattan is covered by 4-letter mapcodes

The mapcode system allows about 100 km^2 of a territory to be covered by 4-letter mapcodes. These are usually reserved for the capital city. About 6,000 km^2 can be covered by 5-letter mapcodes[3]. Roughly 250,000 km^2 can be covered by 6-letter mapcodes. If a territory is larger than that, the remainder requires 7 letters. A few territories on Earth (notably, Australia, Brazil, Canada, China, India, Mexico, Russia and the USA) are large enough to merit 8 letters, but these actually consist of sub-contexts, i.e. people there live in a state, republic or oblast much like people elsewhere live in a country. Within such a sub-context, mapcodes are shorter.

[3] Although mapcodes generally use 31 different symbols, there are actually 62 times as many locations that can be covered by 5-letters than by 4-letters because the dot can also be in more positions (i.e. mapcode PQ.RST indicates a different location than mapcode PQR.ST)

2. Mapcodes are easily recognized as such

In order to become widely accepted, mapcodes need to be easily recognized for what they are, just like emails, websites or phone numbers. Care was therefore taken to make mapcodes look neither like words (due to the dot in the middle, the use of capitals, the lack of vowels) nor like numbers (vowels *are* used to prevent a code from consisting exclusively of digits).
This is also crucial if we wish mapcodes to be supported in universal search boxes (e.g. in the Google search box on an internet browser). Our design is such that we believe it unlikely that anything which looks like a mapcode is not actually a mapcode.

```
www.tomtom.com        web site
pete@tomtom.com       email
R23.K9                mapcode
+631 981526           phone number
```

3. Mapcodes can be made a "natural" part of an address

If we wish mapcodes to become a standard part of an address (e.g. on a business card), it is important they are not easily confused for something else (like a company name, house number, street name or city name). It is therefore recommended they are printed before or after the country or state name (or behind the city name if the country name is left out):

SomeCompany Ltd.	*SomeCompany Inc.*
Oosterdoksstraat 114	*Luxury Ave 135*
Amsterdam ***49.4V***	*Concord, MA,* ***RC2.3R***

Note that mapcodes are usually *more* precise and flexible than an address, and can very much enhance the address information. A very specific location within the confines of an address can be specified via the mapcode, such as the main entrance, a particular side entrance, or the entrance to the parking area.

Also note that Mapcodes are independent of the availability or correctness of local map data. They exist before buildings are built, are not affected by any change to the names of streets, cities or countries, and will provide the correct position on a navigation system no matter how erroneous or incomplete its map or address information is.[4]

4. Mapcodes are easy to verbally communicate

When you need to tell someone your mapcode (e.g. over the phone, so they can enter it into a navigation system and visit you), it is important that the mapcode is easy to communicate verbally. This means that the code must be short (see

[4] Although these benefits are shared with other international coordinate systems (like the latitude/longitude system), they are not shared by postcodes, which encode addresses rather than coordinates.

section 1) and easy to pronounce, which in turn requires the constituent symbols to be simple. For this reason, mapcodes consist only of letters and digits, and use no complicated characters such as ampersands or exclamation marks. Nor are letters case sensitive, obviating the need to spell out things like “capital A lowercase B”.[5]
The separating “dot” is nice and short in most languages, and is communicated so often in email and web addresses nowadays that it feels natural to include in the spelling of a code.

5. Mapcodes are unambiguous

A single mistake in a mapcode can yield a completely different location. It is therefore important to make mapcodes as unambiguous as possible. Fortunately, both the alphabet and the numeric system were designed to be unambiguous, each symbol shaped as distinct as possible from the other symbols. The exception is in the overlap between the two sets, where the letters I and O are easily confused with the digits 1 and 0. The letters I and O are therefore *never* used in the mapcode system. If it looks like an I or O, we recommend it is interpreted as a 1 or 0.

Similar care was taken for mapcodes in other alphabets (discussed further on). Even inter-alphabet ambiguities were taken into account. For example, the roman H, the Cyrillic H (pronounced “N”) and the Greek H (pronounced “E”) are indistinguishable by shape. A business card stating the Moscow address 4H.93 could thus be ambiguous. However, the alphabets were encoded in such a way that such similar-looking symbols result in exactly the same mapcode locations.

6. Mapcodes are *sufficiently* precise

You can reach any destination on Earth with nothing but a mapcode to guide you. Since we want mapcodes to be short (for all the reasons stated elsewhere), we want mapcodes to be no more precise than is necessary for practical purposes. Eventually, we decided that if you reach a point *at most* 5 meters from a certain location, you are “there”. In fact, where an address will at best lead you to the front of a *building*, a mapcode will lead you to a particular door, or the entrance to the parking lot. But even if you were to find more than one front door within that distance, it is unlikely to pose any practical problem.

All that being said, there *is* in fact a “high precision” extension for mapcodes: by adding a hyphen and a letter to a mapcode, a location is represented with an accuracy of less than a meter. Adding a second letter yields an accuracy of 16 centimeters. Although this is precise enough to, say, indicate the exact location of

[5] Codes would seem shorter by distinguishing between uppercase and lowercase letters, as in “m4.TbX”. Indeed, 6 letters would probably be enough to cover the whole Earth. Pronunciation would be such a hassle, however, and feel so unnatural, that such a system would probably not catch on: “small m four dot capital T small b capital X”.

a drainage pipe, it really defeats the whole purpose of the mapcode system. For that kind of precision, normal coordinates would probably suit just as well.

It should also be noted that positions can not easily be *determined* more accurately than to within a few meters. For example, without sophisticated ground-based aids, the positions measured through the GPS satellite system are themselves only accurate to within a few meters.

7. Mapcodes are easy to remember

As with phone numbers, most people will at most bother to remember their *own* mapcode. However, as with phone numbers, mapcodes sometimes need to be communicated. A mapcode like "CX.5R" may be short enough to remember even when glanced on a billboard or picked up during a conversation. At least long enough to jot down later.

The separating dot helps. People find it easier to interpret and remember a code (or number) broken up in smaller groupings.

If also helps that mapcodes are short. Note that they are especially short in the centers of the largest cities – which is exactly where many venues are located that people want or need to visit.

But even in small villages, where mapcodes are sometimes a bit longer, it is often the case that *all* locations in the village have the same prefix (i.e. the same letters before the dot). For example, when every location in a village start with the letters M3C, most people only need to make the effort of remembering the *last* three letters of a mapcode to identify a location within their village.

8. Map Codes allow access to locations that have no address

Over two billion people are estimated to live without street name or house number. They are forced to communicate locations in the form of directions ("Go to the town square, go 6 streets to the north, take 4th sidestreet on the left, Samir lives next to the yellow house"). Such directions are in fact often relative to landmarks which themselves have no address, and are seldom related to the direction from which visitors are coming.

Every single dwelling on Earth, however, has a Map Code, easily acquired, and easily communicated and remembered.

9. Mapcodes are international

From the very start, we aimed for a worldwide system, not restricted to a specific country, culture or even alphabet. We invested a lot of effort in defining tables for every place on Earth, from Antarctica to Siberia and from Amsterdam to Timbuktu.

All 540 countries, dependencies, overseas territories, autonomous states, krai and oblasts are covered. Areas claimed by two countries are simply included in both contexts.

We also designed a character substitution system that allows mapcodes to be represented in different alphabets (as explained further on).

Finally, on top of any and all national mapcodes a location may have, it also always has a 9-letter, context-less map code.

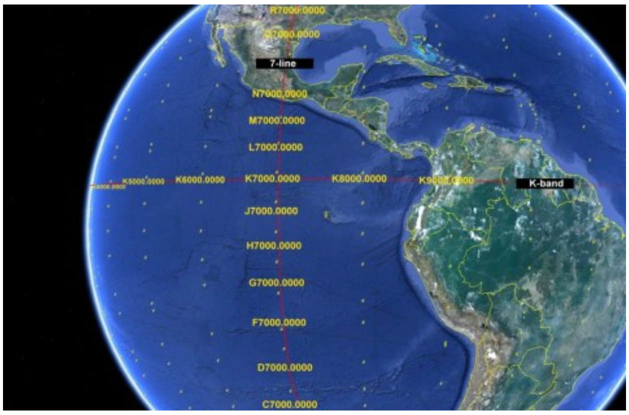

Figure 4. every location on Earth has a 9-letter context-less mapcode

10. Mapcodes are universally supported

We want the mapcode system to be a standard. In order for that to happen, on top of being truly international as described in the previous paragraphs, we believe

- mapcodes should be free; even the smallest license fee would prohibit adoption
- mapcodes should be non-proprietary, brand-less, and unrestricted; no government, for example, would adopt a system that is owned by a company, or requires a copyright or even a "courtesy" notice, not to mention a trademark symbol

Furthermore, in order for mapcodes to become widely adopted, it must be easy for people to determine their own mapcode, and to convert between mapcodes and coordinates. At the very least, this requires

- a free website, in as many languages as possible, where people can convert between mapcodes and coordinates
- free tools that allow bulk conversion between coordinates and mapcodes
- wide support by map websites (Google Earth, Bing Maps etc.), navigation brands (TomTom, Garmin, etc.) and map companies (NavTeq, Tele Atlas etc.)
- free distribution of source code and specification, to stimulate adoption and use in all kinds of applications.

Design history

The mapcode system was specifically designed and developed to set a free, easy, international standard for representing locations on the surface of the Earth. There were several systems for encoding locations already, but few were designed for **free, public, everyday use.** Those that were are almost without exception restricted to one particular country.

The mapcode system was developed in 2001 (based on original designs from 1998) by Pieter Geelen and Harold Goddijn, founders of the Dutch company TomTom that invented the modern-day car navigation systems. The mapcode system was revived and worked out in 2006 when TomTom decided to expand their operations to countries like India and China. It was difficult to sell navigation devices in these countries, among other things because a significant part of the population lives without address, lacking not just house numbers but often even street names, which often made it hard or even impossible to enter a destination. It was thought that the introduction of a simple and ubiquitous system for representing locations (and hence destinations) would in due time make navigation devices as welcome there as anywhere else.

Alphabets

Mapcodes can be easily translated between several different alphabets. The default system is based on the roman alphabet. Mapcodes in other alphabets are generated by character substitution. The substitutions were chosen carefully:

- **mapcodes must be easy to read and write.**
  Preference for characters that have few "strokes", and no accents.
- **mapcodes must be easy to read aloud and communicate.**
  Preference for characters that are short and easy to pronounce (in the sense that the roman letter "B" is shorter and easier to pronounce than the roman letter "W"). Furthermore, characters were chosen that sound as different as possible from all other chosen characters.
- **mapcodes must be easy to type.**
  Preference for characters that do not require Shift or other special keys to access on a local computer keyboard.
- **Mapcode characters must be easy to recognize and distinguish.**
  Preference for characters that look as simple as possible, and look as different as possible from all other chosen characters.
- **Mapcodes must not look like normal language.**
  Preference for consonants. If possible, all vowels are excluded.
- **Mapcodes must be unambiguous.**
  The Latin H can not easily be distinguished from the Greek H (pronounced "eta") or the Cyrillic H (pronounced "n"). Thus, the mapcode HH.HH might be misinterpreted unless all these same-looking characters are mapped to the same value.

High-precision mapcodes

Although it slightly defeats the purpose of mapcode concept, an extension is available that allows locations to be defined at much higher accuracy than is possible with normal mapcodes.

To understand mapcode accuracy, you need to understand that a specific mapcode *defines* a specific location X, but *represents* all the locations within a certain distance meters of X. For example, the Dutch mapcode "49.4V" decodes into coordinate 52.376514, 4.908542. This does not only mean that coordinate 52.376514,4.908542 can be represented by mapcode 49.4V, but that all coordinates within roughly 5 meters of that coordinate can also be represented by mapcode 49.4V. To be precise, any latitude between 52.376**470** and 52.376**559**, in combination with any longitude between 4.908**471** and 4.908**616**, yields mapcode 49.4V. The rectangle defined by these ranges is called the *encoding zone* of the mapcode.

The high-precision extension for mapcodes is based on the simple idea of appending letters at the end of a mapcode (using a hyphen as separator). The first letter divides the encoding zone into 30 subzones[6] through a 5x6 grid. A second letter divides each subzone into 6x5 even smaller zones. Thus

Netherlands 49.4V
defines all locations within roughly 5 meters of coordinate 52.376514, 4.908542.
Netherlands 49.4V-K
defines all locations within roughly 1 meter of coordinate 52.376506, 4.908545.
Netherlands 49.4V-K2
defines all locations within roughly 16 centimeters of coordinate 52.376512, 4.908540.

Although this idea could easily be extended to more letters and higher accuracies, our mapcode algorithms today do not support that. The reason is that the algorithms were designed around 32-bit integer arithmetic, both for the sake of efficiency and to assure that calculations do not easily suffer from floating-point errors when ported between different programming languages. Internally, coordinates are therefore represented in millionths of degrees (e.g. the latitude 52.376506 is represented by the 32-bit integer 52,376,506). The smallest difference that can affect the outcome of any calculation is therefore 0.000001 degrees, or about 11 centimeters.

Maintenance and Politics

Our world is always changing. New countries are sometimes born (like South Sudan in 2011). Furthermore, the *current* world situation isn't always completely clear, what with disputes about the borders or even the very existence of certain countries.

[6] there are actually 31 letters available for normal mapcodes; the letter Z is excluded for the high-precision extension.

The Mapcode Foundation reserves the <u>exclusive</u> right to make changes to the mapcode data table and algorithms, even though they are distributed freely. This is done for the simple reason that *any* change to the data table or algorithm would result in one or both of the following:

- after a change, certain mapcodes would result in *different coordinates* than before
- after a change, certain locations on Earth would get mapcodes on newer systems that can not be recognized by older systems (or vice versa)

The first issue should be prevented at all costs. The second may be necessitated by changes to the world order, i.e. new countries coming into being. But unless changes are made by a single central authority, the first issue will quickly come to be.

Fortunately, the mapcode system is a practical rather than a political system. Mapcode territories are "rectangular" and never perfectly match the actual borders of a country or state anyway. The mapcodes in a country are intended first and foremost to benefit the people who live there. They can pick their own context. Areas claimed by more than one country were therefore simply included within the mapcode borders of all claimants.

Note that usually, the country context is just a "background context", seldom explicitly mentioned in an address. On navigation systems, people usually specify their country only once, after which every search and address they type is assumed to be within the borders of that country. The decision to change the name of a country (like Burma to Myanmar in 1989, or Zaire to Congo in 1997) would thus have little effect on *local* addresses or mapcodes. Nor will people locally, for many years to come, be surprised or confused when some older system, business card or letterhead refers to the old country name.

Something very similar holds even for countries that secede from an existing country. Take the secession of Kosovo from Serbia. Any location and any address in Kosovo already existed within Serbia. We would thus make Kosovo an "alias" for Serbia in the mapcode table. The mapcode **0FD.CT** will then get you to the Library in Prishtina. Not the *system*, but the *person* who provides you with that mapcode will tell you whether the context is "Serbia" or "Kosovo" (if at all – when you are both in Prishtina, the context probably needs no mention). Either will get you to the library entrance.

Appendix

Disputed territories

The following table does not presume to be complete, but lists some important areas (some disputed, some recently changed in status) in relation to the release of the mapcode table.

With only a few exception (e.g. Antarctic claims), a territory appears in the data table if it had a 3166 ISO code in April 2013, or when it is a state (province, oblast…) in Australia, Brazil, Canada, China, India, Mexico, Russia, or the USA.

Name	Status	Relation to mapcode table
Abkhazia	Claimed by Georgia; considers itself an independent state; also called Apsny	Not in mapcode table. No known ISO 3166 code. All locations are covered by Georgia (GEO)
South Ossetia	Claimed by Georgia; self-declared independent republic	Not in mapcode table. No known ISO 3166 code. All locations are covered by Georgia (GEO)
Kosovo	Currently part of Serbia	Not in mapcode table. No known ISO 3166 code. All locations are covered by Serbia (SRB)
Nagorno-Karabakh	Claimed by Azerbaijan; de facto independent but unrecognized state	Not in mapcode table. No known ISO 3166 code. All locations are covered by Azerbaijan (AZE)
Somaliland	Claimed by Somalia; unrecognised self-declared de facto sovereign state	Not in mapcode table. No known ISO 3166 code. All locations are covered by Somalia (SOM)
Transnistria	Claimed by Moldova; self-declared independent territory	Not in table. No known ISO 3166 code. All locations are covered by Moldova (MDA).
Palestine	Disputed by Israel	In mapcode table as territory “PSE”. Note: mapcodes in PSE also exist in ISR.
Western Sahara (Sahrawi)	Claimed by Morocco	In mapcode table as territory “ESH”. Note: mapcodes in ESH also exist in MAR.
Taiwan	Claimed as a province by the People’s Republic of China	In mapcode table as territory “TWN”. Note: ISO 3166-2 abbreviation CN-TW (representing a Chinese state) functions as an alias for TWN in the mapcode alias list
Tibet	Autonomous region within the People’s Republic of China	In mapcode table as Chinese state “CN-XZ“ (called “Tibet” or “Xizang”).

Border conflicts

As explained in section 8, areas claimed by more than one territory (of which there are quite a few on Earth) are usually simply included within the mapcode rectangles of both territories.

Claims on islands, overseas dependences

Many nations have or claim overseas dependencies. Some are completely undisputed (Hawaii is an undisputed state of the USA, the Cocos Islands an undisputed territory of Australia), some are peacefully autonomous (Greenland is an autonomous country within the Kingdom of Denmark, New Caledonia a special collectivity of France), some are disputed (Cyprus considers itself a sovereign republic, Turkey considers part of it as a Turkish republic).

For addresses on any island, the country context is usually *very much* a background context. And fortunately, all inhabited overseas dependencies have their own ISO 3166 abbreviation, and thus their own (often very short) mapcodes, making the exact relation with any far-away claimant moot.