

Targeted dissemination through reverse geocoding services

Koketso Moeng (Statistics South Africa)



Introduction

The dissemination of information is a critical component of a national statistics agency's mandate. The work and effort of a statistics agency is only realised when the various stakeholders start assimilating the information into their decision-making process or when they extract further insight through further analysis of the data made available by the statistics agency.

One of the good practices highlighted in the Principles of International Statistics is "providing equal access to statistics for all users". Another critical factor is disseminating in a format that the end-user will be able to understand. These factors necessarily imply that statistics agencies should think thoroughly about the dissemination mediums and the language that they use to communicate the information.

The combination of smart devices, in large enough numbers, and the emergence and increasing use of both geo-location and reverse geo-location services on websites offer statistics agencies the possibility to disseminate data in a rich, context sensitive and immersive way like never before.

This paper will propose a new format of disseminating official statistics that can be very effective not only in its reach to information consumers, but also because it is inexpensive in relation to other dissemination mediums.

Traditional dissemination

Historically, statistics agencies have been using paper as the medium of sharing information. With the introduction and mass adoption of the Internet, most agencies have simply continued with the traditional paper medium but now have availed the content in digital format on their websites. This method of dissemination will remain the bedrock of statistical dissemination, as is the case with all other scientific fields.

At Stats SA, we include geographical elements or spatial representation of our data as part of disseminating our statistics. This is of great importance from a policy and governance perspective, because it allows those in a position to effect change to be able to monitor their policies or to implement policies based on evidence. For those who are affected by change, disseminating data with geographical elements allows them to objectively monitor promises made by those in power.

It has to be said that despite the Principles of International Statistics' standard of equal access, access to information has been historically biased towards those with the skills to decipher the statistical jargon. While people might have a genuine desire to incorporate statistics into their daily decision-making processes, if the data are such that they are presented in a way that most users will not understand, they will invariably stop using the data and look elsewhere for more digestible bits of data.

Another limitation of traditional dissemination methods is their inflexibility and lack of intelligence in how they reach the end-user. The views that were created by the

author(s) of the body of work remain static and unresponsive to the unique context of the information consumer.

Information is usually packaged based on geographic demarcations that make it difficult for the layman user to make any meaningful sense of the information. For example, the trends that a user might see in their municipal/state/regional profile might not hold true for their locale. What's needed is a context-sensitive way to deliver data.

A geographical context-sensitive dissemination strategy will provide more meaningful insights for users and allow them to navigate their landscape much more easily with the relevant data about their immediate location. People are also likely to use the data if it is readily available with few interactions with the website or service that is providing the information.

New dissemination methods

Many statistics agencies around the world are experimenting with a variety of new ways of disseminating statistics in a much more user-friendly format than the traditional reams of paper with incomprehensible graphs and statistical jargon.

The new formats include:

- Mobile applications: [Stats SA iPad application](#)
- Animated websites: [The Australian Bureau of Statistics](#)
- Animated Presentations: Statistics South Africa
- GapMinder: [World – GDP/Capita and life expectancy](#)
- Google Public Data Explorer: [Stats SA PDE Page](#)

Each of the new dissemination platforms have their own pros and cons.

What are geo-location services?

Geo-location services or, as they are commonly known, geocoding, are a combination of hardware and software services that allow one to deduce a person's geographic location.



Rudimentary versions of geolocation services have been around for many years but have recently matured and only now are they being consumed by a variety of users (knowingly so or not) during their day-to-day activities that mostly involve interacting with their mobile devices.

When one makes a call, it will usually show the recipient of the call the area from which the call was made – this is a very rudimentary form of a geolocation service.

On the Internet, websites can deduce the country and city from which the website user is coming by using the IP address of the user. Another format is by using cell phone tower triangulation to estimate the location of the mobile device.

Most smart devices are equipped with a Global Positioning System (GPS) which can also be used to locate the user's geographic location to within a few metres. Most of the mobile applications and websites use the device's GPS to determine the location.

Internet mapping solutions like Google Maps and Microsoft's Bing Maps are used by most people as navigation tools – these tools have become indispensable in today's ever-changing landscape. These solutions, however, also serve another purpose – they can be used to translate a set of GPS coordinates (longitude and latitude) to an identifiable place name like Cape Town or Cairo. This technology is called reverse geo-location services.

Google Maps and Microsoft's Bing Maps can, using the GPS coordinates, infer the name of the area a person is in and make that information available to a third-party solution like a website or a mobile application. The services essentially take non-human readable/understandable information (longitude and latitude) and convert that into a human readable address or place name. Traditional geo-location services work in reverse: the services take a person's current area and convert that into a set of GPS coordinates (these services take a human readable address or place name and give it a set of GPS coordinates).

In the last five years or so, there has been a big increase in the number of location-based services offered to consumers. Typically, the services are set up by marketers who offer discounts to consumers who are in a specific area. They would then, through their mobile application, offer consumers some discount for those who happen to be in the targeted area. This is called ring-fenced services because the discounts are usually within a specific radius of the outlet offering the service. An example of such a service is Foursquare.

Stats SA's experience in reverse geocoding is with Google Maps Geocoding API. This paper will focus on that geocoding engine

Why should I care about geo-location services?

Most of the data made available by statistics agencies include geospatial attributes. In the case of Stats SA, along with many other statistical agencies, we use enumeration areas (EAs) as the lowest data collection block or area. The data, by default, carry a geography attribute which can be mapped to the existing geography hierarchy.

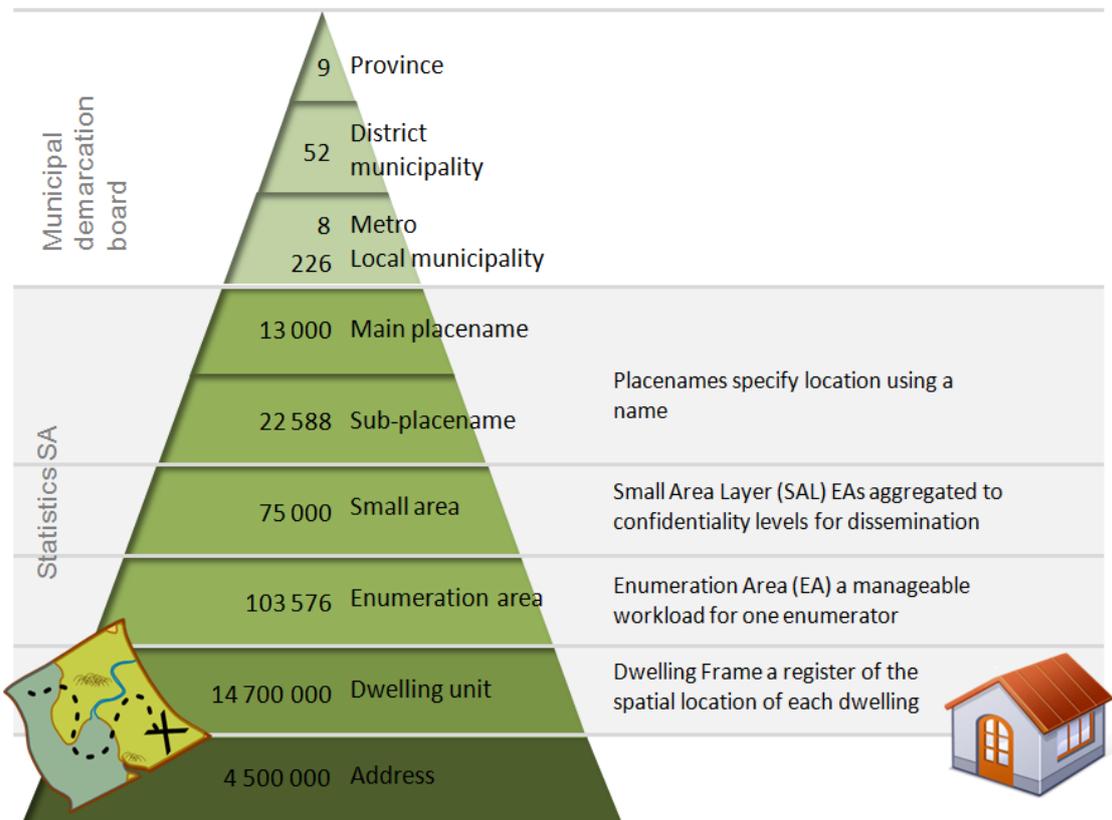


Figure 1: South Africa's geography layers

South Africa's geography hierarchy is as follows:

- Provincial level
- District municipalities
- Metro / Local municipality
- Main place name
- Sub-place name
- Small area
- Enumeration area
- Dwelling Unit
- Address

When one combines the accuracy of smart devices with the possibilities offered by reverse geocoding services and the spatially enabled data, the stage is set for agencies to disseminate information in a very targeted and rich way.

Geolocation services make dissemination interesting in that they cut through the many administrative layers that the user might not be interested in and direct the user to the data about their immediate surroundings – in many ways, it can be likened to augmented reality.

The ease of use and precision of the solution and the view or perspective that one gains from localising data in this fashion are unparalleled in the statistics field; it is a big step in digital dissemination and the digital revolution.

There are many uses for this kind of dissemination method:

- A family searching for a house in a foreign neighbourhood can use the service to obtain demographic data about the area.
- An entrepreneur might use the service to get a quick glimpse of the local economic data in an area that he is considering for business.
- Politicians can access the information when addressing constituents.
- The tool can be used for disaster management response teams to familiarise themselves with the area.
- City workers can use the tool to assist them as they plan and deliver services.

How does it work?

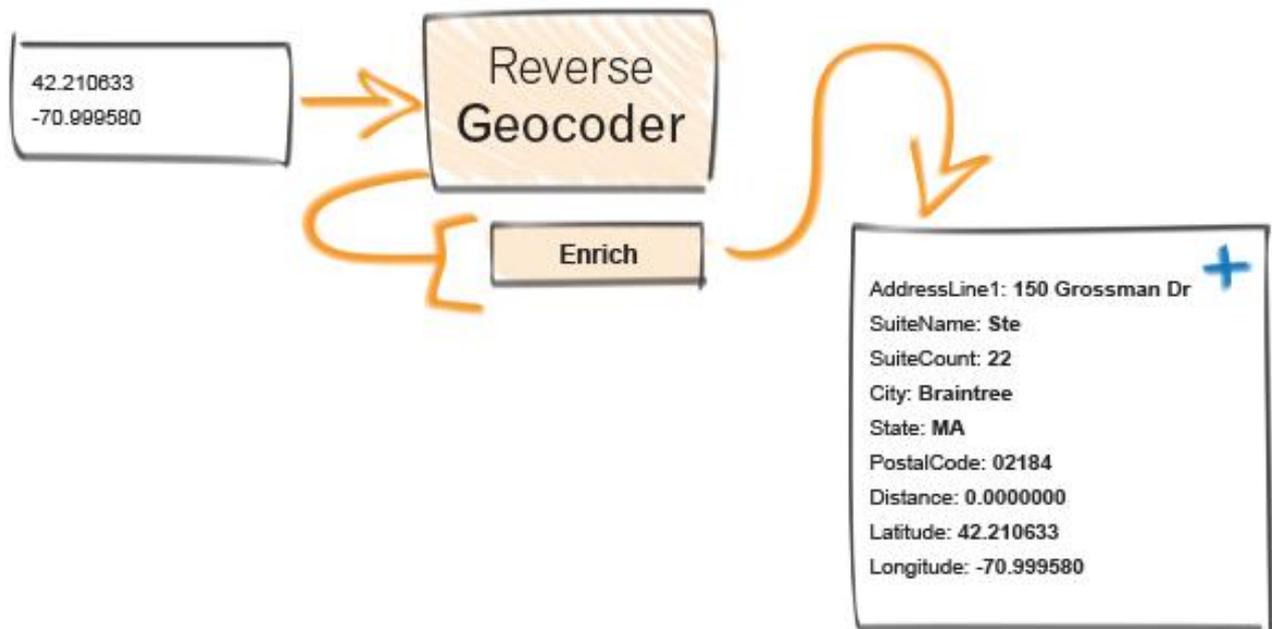


Figure 2: A simple depiction of how reverse geocoding works

Depending on which platform the services are accessed, via mobile application or through a browser on a mobile device, the interaction might be slightly different, but in general, the following steps will be followed:

1. A user will visit the agency's website.
2. The website will then prompt the user to share his/her GPS



- coordinates with the agency's website.
3. The website will then submit the coordinates to the geocoding engine.
 4. The geocoding engine will in turn, return a set of data (province/state, municipality and place name) derived from the submitted coordinates.
 5. The website will then query the agency's internal database for the indicators that are linked to the given place name.
 6. The user is presented with a selection of the available indicators for the given place name.

Considerations for reverse geocoding

The Google Geocoding API is available in two versions.

- **The free API:** This version is fully functional, but Google has placed the following limitations on it:
 - 2 500 requests per 24-hour period
 - 5 requests per second
- **The paid version:** Often referred to as Google Maps API for Work has the following limitations:
 - 100 000 requests per 24-hour period
 - 10 requests per second

To use either version of the API, the mobile application or website should use an API key that is issued by Google. This allows Google to monitor the mobile application or website's usage of their services. This also allows Google to know who to contact if there should ever be a need to.

Communication strategy

Geocoding works on most modern browsers, but only after the end-user has explicitly allowed the website or mobile application to access their location services (*users opt-in*). Due to rising privacy concerns worldwide, it is important to educate users on how the service works.

In our opinion, this is more so for an NSO, as users might think or suspect that you are collecting their location data or snooping on them through the service. It is important to stress that the service only works if they give consent and that the service takes a snapshot of their location and will not poll their devices to determine their location after they leave the website or exit the mobile application.

It is also important to communicate on which devices and operating systems the solution will work. Some browsers such as Chrome on Android devices will only ask you once if you want to share your location with a website, while when browsing on IOS, the user will be prompted every time the mobile application or website tries to access their location services. The following browsers and operating systems/devices support geocoding services.

Browsers

Internet Explorer 9.0+
Firefox 3.5+
Safari 5.0+
Chrome 5.0+
Opera 10.6+

Operating systems/devices

iPhone 3.0+
iPad
Android 2.0+

Conclusion

The rate at which mobile devices have been adopted worldwide has been faster than any other technology that preceded it.

Their pervasiveness, versatility and processing power make them the ideal platform for disseminating statistics to the various users and makes them a seamless fit into the Principles of International Statistics.

An increasing number of people are realising the value of statistics and are eager to get hold of the right data at the right time. Statistics agencies need to find the ideal solutions that will cater for the users' needs and also make access to the information as easy as possible - Reverse geocoding is a step in the right direction towards this goal.

