

**Meeting of the UN-GGIM Expert Group on the integration of statistical and
geospatial information – Paris 25 and 26 April 2016**

Update proposal for the statistical-geospatial terminology database

Date: 30/03/2016

Version: 0.1

Authors: PETRI

Revised by:

Approved by:

Public:

Reference Number:

Document History

Version	Date	Comment	Modified Pages
0.1	30/03/16	Document created by PETRI	All

1 PURPOSE OF THE DOCUMENT

This document complements a document on the same topic that was presented to the Expert Group at its meeting in May 2015 in Lisbon¹.

The purpose of this document is to better define the forthcoming Statistical Geospatial Framework (SGF) and two of its key concepts: geocoding and georeferencing and related terms.

2 SOURCES

The main source for the terms of this paper is the draft the Statistical-Geospatial-Framework² and relevant information from earlier UN census documentation. A complete list of data sources can be found in Annex 1.

3 BACKGROUND

After the meeting in Lisbon in 2015, the EFGS steering committee decided to transfer the terms defined in the first paper to the terminology database on the EFGS website where they can be reviewed and possibly modified. This solution is now available.

However during the drafting of the Statistical Geospatial Framework the need to add new terms or fine-tune existing definitions has come up.

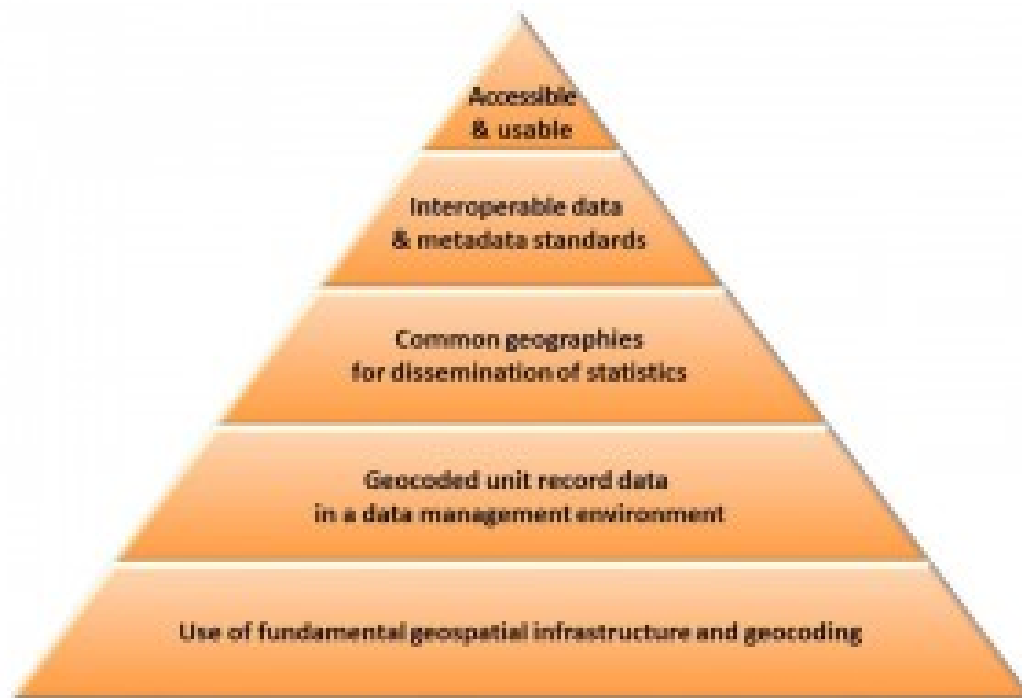
4 NEW OR UPDATED TERMS

4.1 Statistical Geospatial Framework (SGF)

The Global Statistical Geospatial Framework is a high-level, generic framework that consists of five principles that are considered essential for integrating geospatial and statistical information for the purposes of official statistics.

¹ <http://ggim.un.org/docs/meetings/2nd%20UN-EG-ISGI/UN-GGIM%20EG%20Lisbon%20meeting%20session%204%20background%20paper%20terminology.pdf>

² <http://unstats.un.org/unsd/statcom/47th-session/documents/BG-2016-31-proposal-for-a-global-statistical-geospatial-framework-E.pdf>



The SGF has the goal to spatially enable statistics along the full statistical production process. It connects spatial information that describes our physical man-made and natural environment, and statistics that describe their socio-economic and environmental attributes.

The SGF does not contain statistical or geospatial data. Its purpose is to improve the integration of statistical and spatial data, support the generation of new information through a number of processing steps, and improve the dissemination and use of geospatial statistics. From a user perspective, the framework should allow the user to discover, access, integrate, analyse and visualise statistical information for any spatial area of interest.

To fulfil the above requirements the statistical-spatial framework requires the geocoding of statistics at unit record level using authoritative geospatial information. It mandates a system of consistent geographies and grid systems for statistical dissemination. It also calls for the technical integration of statistical and geospatial information in IT based production systems. This requires the existence of technical standards for data exchange, structural and semantical interoperability, dissemination, and metadata.

Moreover, it also needs to address organisational issues such as a data policy, data maintenance, data quality and data protection to regulate access conditions to data and dissemination rights of information.

4.2 Geocoding

Geocoding is the process of transforming a description of location or location information (such as an address, name of a place, or coordinates) to a location on the earth's surface. In other words geocoding is a way to ensure data know where they are.

For the purposes of the Global Statistical Geospatial Framework, geocoding is generally defined as the process of geospatially enabling statistical unit records so that they can be used in geospatial analysis.

More specifically, geocoding is the process of linking unreferenced location information, often in the form of a text string (e.g. an address or address id), that is associated with a statistical unit, to a geocode. Alternatively, the geocode can be directly incorporated into the statistical unit record.

The condition for geocoding are high quality physical address, property or building identifiers, or other location descriptions, in order to assign accurate coordinates and/or a small geographic area to each statistical unit.

4.3 Geocode

A **geocode** is the spatial representation of descriptive location information such as an address string. Geocodes are, preferably, fine scale geospatially referenced objects that are stored as a geometry data type, such as: location coordinates (i.e. x, y, z coordinates) and/or small area geographies (e.g. mesh blocks, block faces or similar small building block geographies). Larger geographic units, such as enumeration geographies, can be used as geocodes where finer scale geospatial units are not available.

4.4 Location information

Location information can include addresses, property or building identifiers, as well as other location descriptions, such as enumeration geographies and other standardised (e.g. what3words reference) and non-standardised (e.g. village names) textual descriptions of a location.

4.5 Statistical units

Statistical units can include persons, households, businesses, buildings or parcels/units of land.

4.6 Georeferencing

Georeferencing is a set of broad processes that includes **geocoding**. Georeferencing, or geospatial referencing, is the process of referencing data against a known geospatial coordinate system, by matching to known points of reference in the coordinate system (e.g. image rectification to survey points or addresses linked to parcel centroids), so that the data can be viewed, processed, queried and analysed with other geographic data.

5 NEXT STEPS

After discussion in the Expert Group in its meeting in April 2016 and approval, the definitions should be added to the terminology section of the web site of the European Forum for Geography and Statistics (EFGS)³ for further discussion and fine-tuning.

In addition to discussing this paper the Expert Group should establish in its meeting in Paris an editorial board composed of experts from both communities (statistics and geospatial). The editorial board should walk through all the terms in the EFGS website and review their definitions and also establish cross-references.

The editorial board will remain in charge of revisions of existing terms and addition of new terms. The webmaster of the EFGS website can then modify the database.

6 ANNEX 1 – INFORMATION SOURCES

- (1) Statistical Geospatial Framework <http://unstats.un.org/unsd/statcom/47th-session/documents/BG-2016-31-proposal-for-a-global-statistical-geospatial-framework-E.pdf>
- (2) Wikipedia <https://en.wikipedia.org/wiki/Geocoding>
- (3) EFGS terminology database <http://www.efgs.info/information-base/introduction/terminology/>

³ <http://www.efgs.info/information-base/introduction/terminology/>