

# National Statisticians Embracing Location-Based Information

By Dr. Paul Cheung, National University of Singapore

In August 2014, the official statisticians of 75 countries joined their geoinformation counterparts in a United Nations (UN) Global Forum to find common ground for integrating geospatial and statistical information.

This is the culmination of a process initiated in 2011 when the United Nations approved a new intergovernmental body to deal with Global Geospatial Information Management (UN-GGIM) and organized the first-ever discussion among the statistical and geoinformation communities on the potential integration of data processes and protocols. In 2013, the United Nations Statistical Commission, in its 44th session, officially endorsed a program of work to develop a statistical-spatial framework as “a global standard for the integration of statistical and geospatial information.”

In many countries, the process of working together across disciplines and organizations to integrate multiple sources of information to meet user needs has gained momentum and recognition. The urgent need to introduce global standards will add impetus to this momentum.

## A Paradigm Shift

National Statistical Organizations (NSOs) are the governmental institutions responsible for the collection and dissemination of economic, social, and environmental statistics. NSOs are not new to location-based information. The enumeration districts (EDs), a key deliverable produced from NSOs, supply the backbone for organizing population, housing, agriculture, and economic censuses.

EDs form the basis of statistical maps and delineate polygons used for data input. They are designed for a practical purpose. EDs divide enumeration work into manageable groups of units. However, ED boundary lines do not necessarily coincide with the administrative boundaries, so census maps very often have their own

with distinct boundaries and characteristics. Most NSOs use basemaps obtained from their national mapping agency and modify them as necessary to create EDs. Organizationally, a small unit in an NSO will update EDs and other geospatial features.

Establishing global standards for integrating official statistics with geospatial information builds on the existing contributions of the statistical community through establishing a hierarchy of EDs. This bold move will unlock potential contributions by the statistical community and the vast pool of data it compiles and manages.

In the past, NSOs have tended to focus on the compilation and release of national-level data, rather than location-based information or small area geography. International data dissemination protocols have yet to address best practices in the release of local-level location information. The statistical community recognizes that there is much to be done and has reviewed a range of important topics, from institutional arrangements and regulatory compliance to technical data issues.

## Global Standards

Two issues have emerged as priority areas for further discussion. The first issue concerns the size of the smallest statistical area in the hierarchy of statistical units for public release. Censuses are increasingly collecting information based on x,y coordinates, implying the geocoding of individual households or establishments. While this is extremely useful for backroom compilation and analysis, it poses a problem when used as the basis of data aggregation and analysis because of confidentiality concerns.

Safeguarding the confidentiality of information is a key element of the official statistical system. The Fundamental Principles of Official Statistics, first drafted in 1994 and recently endorsed by the United Nations General Assembly,

states clearly that “individual data collected by statistical agencies for statistical compilation are to be strictly confidential and used exclusively for statistical purposes.” This international standard has a profound constraining impact on how NSOs manage their data and how they define the smallest statistical area. It is obvious that data based on x,y coordinates will never be publicly released as the basis of data aggregation.

The official statistical community understands there is an urgent need to do something and to link “socioeconomic and spatial information to improve the relevance of the evidence, on the basis of which decisions will be made,” as stated by the UN Statistical Commission. It further recognizes the merits of employing consistent practices in defining a common set of hierarchical geographic boundaries, which coincide with the hierarchy of statistical areas.

At present, official statisticians appear to be comfortable with defining the smallest geographic area as comprising about 300 to 400 people. This will be a significant improvement from current ED-based statistical maps.

To create this hierarchy of statistical areas, there is still much work to be done to improve the geocoding framework of many countries. A consistent system of house addresses—the simplest form of geocoding—is still lacking in many countries.

How should a country design an efficient geocoding system for public use? What should be the key features of such a system? Can we make them comparable across countries? These questions are critical for the development of comparable and consistent systems of statistical areas across countries.

The second technical issue to emerge addresses the relative merits of grid versus administrative boundaries as the basis for aggregating statistical areas. The grid system builds on the Nomenclature of Territorial Units for Statistics (NUTS) protocol endorsed by the European Commission, which

uses a grid to divide Europe into uniform areas, disregarding administrative boundaries.

However, the administrative boundaries system is more commonly used. The two systems could be developed simultaneously for meeting different needs. A grid system could ultimately evolve into a global grid that divides the world into uniform areas with measurable socioeconomic characteristics.

## A New Information Era

The world of information is changing rapidly. NSOs are aware that they need to move with the times to meet the needs of users. A new information architecture—comprising official statistics, geoinformation, and the unstructured big data—will be the new order of the day. Official statisticians will no doubt play an important role in this new world as they modernize their practices and bring a wealth of information to bear in support of evidence-based decision making. There are important technical and governance issues to be resolved. Will NSOs take up this challenge and capture center stage in this new information architecture?

For more information, contact Dr. Paul Cheung at [paul.cheung@nus.edu.sg](mailto:paul.cheung@nus.edu.sg).

## About the Author

Professor Paul Cheung, a national of Singapore, is professor of social policy and analytics at the National University of Singapore. He also serves as an adviser to many governments. In 2013, he returned to Singapore after serving for nine years as the director of the United Nations Statistics Division (UNSD). In that position, he facilitated the development of the global statistical system and coordinated the work of the UNSD. In 2011, his initiative established an intergovernmental platform to address issues on Global Geospatial Information Management and was endorsed by the UN. This global multilateral mechanism addresses critical issues on geospatial information, and a series of high-level meetings have been held. Prior to his appointment at the UN, he served as chief statistician of the Government of Singapore (1991–2004) and is currently the chair of International Steering Committee on Global Mapping, an intergovernmental body with secretariat in Japan. He has received many national and professional awards.

