Standardizing the name of political divisions:
The example of the SALB project

S. Ebener

WHO, 20 av. Appia,
1211 Geneva 27,
Switzerland,
Tel.: +41 22 791.47.44
Fax.: +41 22 791.48.06
Email: ebeners@who.int

Abstract

In most, if not in all countries, a large number of administrative decisions are taken at the provincial if not at the district level. This therefore certainly explains the large quantity of information and data that are available for these levels.

The fact that not only the government but also a large number of organizations, such as the United Nations, collects, manages and shares sub national level data calls for finding ways to insure the compatibility (interoperability) between these different sources of information in order to support decision making from the sub national to the global level.

If standardizing the names of these divisions seems to be a natural solution, a closer look at this actually reveals other elements that needs to be addressed at the same time if we want to reach this objective.

The present paper therefore introduces these elements before describing the process followed and the products obtained in the context of the Second Administrative Level Boundaries data set project (SALB) as a contribution towards the better use of sub national data which is produced and used by the UN and the international community.

Introduction

As reported by Orth (1990) and later in the manual for the standardization of geographical names developed by the Department of Economic and Social Affairs (2006) in the context of the activities of the United Nations Group of Experts on Geographical Names (UNGEGN) report that civil/political divisions, also called administrative divisions, are one of the categories of geographical entities that are normally given proper names.

Many organizations therefore believes that standardizing the name attached to each of these geographic entities is sufficient in order to insure interoperability between databases coming from different sources.

It is true that standardizing the name of the divisions facilitate the linkage between different datasets by reducing the potential difference of spelling that could be observed for a same unit.

Nevertheless, there is one particular issue which can not be managed through the standardization of the division's names. This concerns the fact that political divisions such as provinces or districts are geographic entities which might evolve through time in order to reflect the changes that do occur in the administrative and/or political structure of a country.

The importance of these changes on data collection, management and/or sharing can be easily understood by looking at an extract of the historic changes that occurred at the 1st sub national level in Uganda between January 1990 and November 2000 (Table 1).

01.01.1990-14.03.1991	15.03.1991-20.03.1997	21.03.1997-27.11.2000
Moyo	Moyo	Adjumani
		Moyo
Apac	Apac	Apac
Arua	Arua	Arua
Bundibugyo	Bundibugyo	Bundibugyo
Bushenyi	Bushenyi	Bushenyi
	Ntungamo	Ntungamo
Mbarara	Mbarara	Mbarara
Tororo	Tororo	Tororo
		Busia
	Pallisa	Pallisa
...	...	...

Table 1 - Extract from the 1st sub national level historic changes table provided by the Lands and Surveys Department of Uganda

Three different district level structures have existed over this particular period of time. If we look for example at the division of Tororo (Table 1) we can observed that:
- This division has been split into two divisions in March 1991, one of the new division keeping the same name Tororo, as before.
- The new division of Tororo got split again in March 1997 with one of the new divisions keeping the same name as before.

In other words, the division of Tororo, and indirectly any information that would be attached to it, as observed in 2000 can not be compared to the one presenting the same name in 1990 as its "geography" has changed over time.

This example demonstrates that standardizing the names of administrative divisions might not be enough in order to insure interoperability between different data sources as having the same name does not necessarily guarantee that you are talking about the same division. The time component therefore needs to be embedded in the different databases in question in order to avoid this potential problem.

An other matter concern the use of the Roman script (also referred to as Latin script), adopted as a base for international use by the United Nations, and the Group of Experts (Department of Economic and Social Affairs,2006). The fact that this script allows for the use of diacritics might, still today, generate some problems if the data management system does not handle them like it is for example the case in some Geographic Information Systems (GIS).

To summarise, in order to be able to insure interoperability between different sub national databases it is needed to develop a kind of platform which provide users with standardized names for these division through the use of a specific script that would easily pass from one system to an other and a way to capture the changes that occurs through time.

Some value would also be added to this platform in case it would also provide access to maps, in digital format, allowing the visualization of information through time.

Finally, establishing consistent written forms of political division names and their applications throughout the world depending heavily on the official use of names within each country it is crucial that all the processes leading to the creation and distribution of this platform be developed in close collaboration with recognized governmental entities in individual countries.

The example of the SALB project

The Second Administrative Level Boundaries data set project (SALB), launched in 2001 under the umbrella of the UN Geographic Information Working Group (UNGIWG), is a collaborative effort between the National Mapping Agency (NMA) of each UN Member State, the Asian Institute of Technology (AIT), the UN Economic Commission for Africa (UNECA), the UN Economic Commission for Latin America and the Caribbean (UN ECLAC), the UN Economic and Social Commission for Asia and the Pacific (UNESCAP), the UN Map Library and the World Health Organization (WHO) through its headquarters and two of its Regional Offices.

This collaborative effort has the objective of providing free access to, for non-commercial use, a working platform for the collection, management, visualization and sharing of data/information down to the second administrative level that covers all the UN member countries, is standardized and validated by the concerned NMA.

This working platform is accessible in the form of a data set which contains:

- the contact information of the NMAs,
- historic changes tables, stored in Microsoft Excel format files, presenting the evolution of the administrative structure in terms of geography and names at the 1st sub-national level since January 1990 and at the 2nd sub national level since January 2000,
- GIS format map for each period of representation observed since January 2000.

In this dataset the standardization of the name of each political divisions is insured through the use of the Roman script (English alphabet), the spelling of each name being validated by the NMA of the concerned country.

The time component is itself being captured through the provision of a specific code (identifier) attached to each political division once its existence and spelling has been validated by the NMA. The code evolves through time in order to reflect the changes that do occur in the administrative structure and allows for an easy identification of the administrative level the user is looking at. The description of this coding scheme, as well as the protocol that is used in order to code the changes that occurs through time in the administrative structure is accessible from the SALB web site at: http://www.who.int/whosis/database/gis/salb/salb_coding.htm.

As an example, Table 2 contains an extract of the historic changes for Uganda, similar to the one presented in Table 1, but this time with the mention of the codes which are used in the context of the SALB project.

01.01.1990-14.03.1991		15.03.1991-20.03.1997		21.03.1997-27.11.2000	
Moyo	UGA997	Moyo	UGA997	Adjumani	UGA001
				Moyo	UGA033
Apac	UGA002	Apac	UGA002	Apac	UGA002
Arua	UGA003	Arua	UGA003	Arua	UGA003
Bundibugyo	UGA005	Bundibugyo	UGA005	Bundibugyo	UGA005
Bushenyi	UGA987	Bushenyi	UGA006	Bushenyi	UGA006
		Ntungamo	UGA039	Ntungamo	UGA039
Mbarara	UGA991	Mbarara	UGA031	Mbarara	UGA031
Tororo	UGA993	Tororo	UGA999	Tororo	UGA045
				Busia	UGA007
		Pallisa	UGA040	Pallisa	UGA040
...	...	...	...	...	...

Table 2 - Extract from the 1st sub national level historic changes table provided by the Lands and Surveys Department of Uganda with the mention of the SALB codes

The standardized names and codes are then been integrated into the attribute table of each GIS format map that is prepared in collaboration with and validated by the NMA before being posted on the SALB project web site.

In order to allow users to have access to other official spelling and codes used in the country, key correspondance tables, built in collaboration with the NMA, are often added to the excel file containing the historic changes tables.

What is presented here is of course a very summarized version of the all process which is followed in order to obtain the final information and data. The description of the full process can be found either on the project overview page of the SALB web site (http://www.who.int/whosis/database/gis/salb/salb_PO.htm) or in Ebener et al. (2006).

Current state of progress of the project

As of July 2007, the following information and data that can be freely downloaded from the SALB project web site (http://www.who.int/whosis/database/gis/salb/salb_home.htm) for non-commercial use:
- The contact information of the NMA for 146 countries,
- January 2000 tables for 168 countries,
- Historic changes table for 131 countries, the information going at least until 2005 for 99 of them,
- January 2000 maps for 47 countries, 49 other countries remaining under validation,
- Map presenting the situation observed after 2005 for 31 countries.

The state of progress for each country can also be visualized under the form of maps that can be downloaded from the "Project Overview" page on the SALB web site.

Discussion and Conclusion

The standardized information and data provided through the SALB project can be used for many other purposes than just producing thematic maps. Among the principal ones we can mention:
- contacting the National Mapping Agencies in countries especially when there is a need for data and expertise during an emergency
- integrating the geographic and time component within Information Systems managing sub national data and allowing therefore the interoperability between all the sub national level data sets produced and/or used for decision making
- serving as a reference to link point features (*e.g.,* populated place) to the above political structure
- if not existing, provide one of the layer of importance when developing a Spatial Data Infrastructure (SDI)

In addition to that, the fact that SALB is developed in close collaboration with the National Mapping Agencies in countries and these agencies being directly involved in, if not leading, the development of National Spatial Data Infrastructure in countries (NSDI) the channel that this project has established with countries can serve as an entry point for this purpose.

For the moment, UNGEGN, through the activities of its Working Groups, is mainly looking at issues such as country names, toponymic data files and gazetteers, toponymic terminology or romanization systems.

If these issues are directly or indirectly linked to the standardization of political divisions names, the Group of Experts is for the moment not proposing a process that does systematically and consistently provides the UN and the international community with standardized names and unique identifier for each of the divisions observed in countries.

For these different reasons the Group of Experts might be interested in looking at the SALB process as a potential complement to its activities when it comes to standardization of the political division names.

Acknowledgements

References

Department of Economic and Social Affairs (2006): " Manual for the standardization of geographical names, United Nations group of Experts on Geographical Names", Department of Economic and Social Affairs, Statistics Division, New York, 2006, document n° ST/ESA/STAT/SER.M/88

Ebener S., Brookes B., Silva A., Gagliano E., Guigoz Y. (2006) "Lessons learned towards the creation of a global SDI: the example of the SALB project in the Americas", 9th Global Spatial Data Infrastructure (GSDI) Conference, Santiago de Chile, 3-11 November 2006.

Orth D. J. (1990):"Organization and functions of a National Geographical Names Standardization Programme: A Manual." *World Cartography*, Volume XXI. New York, United Nations, pp. 11-40.