

## **Lessons learned towards the creation of a global SDI: the example of the SALB project in the Americas**

Steeve Ebener  
World Health Organization (WHO)  
Switzerland  
[ebeners@who.int](mailto:ebeners@who.int)

Brenda Brookes  
UN Dag Hammarskjold Library  
USA  
[brookes@un.org](mailto:brookes@un.org)

Alejandra Silva  
UN Economic Commission for Latin America and the Caribbean (UN ECLAC)  
Chile  
[asilva@eclac.cl](mailto:asilva@eclac.cl)

Elisa Gagliano  
UN Dag Hammarskjold Library  
USA  
[gagliano@un.org](mailto:gagliano@un.org)

Yaniss Guigoz  
World Health Organization (WHO)  
Switzerland  
[guigozy@who.int](mailto:guigozy@who.int)

### **Abstract**

In order to be operational, any SDI needs to rely on georeferenced data. In this context, creating global level seamless and homogeneous data sets presents specific challenges.

Launched in 2001, under the umbrella of the UN Geographic Information Working Group (UNGIWG), the objective of SALB project is to provide the international community with a working platform that covers all the UN member countries for the collection, management, visualization and sharing of data/information down to the second administrative level.

When the project was launched, it was thought that the real issue to be solved in order to obtain a global seamless database of administrative boundaries maps was to improve the coordination among the data producers. Five years later, and despite concrete improvement in this area, SALB is still collecting information and data in order to complete or update this data set for all the 192 UN Member States.

Taking the Americas as an example, the objective of this paper is to share the practices and the lessons learned through this process as a possible contribution towards the creation of a global Spatial Data Infrastructure.

**Keywords:** administrative boundaries, SDI, Americas

## **Introduction**

In order to facilitate a common understanding of transnational situations, and be effective in the achievement of the Millennium Development Goals (MDGs) there is a need to be able to collect, manage, share and represent data at the sub-national level.

At the same time the constant increase of the number of deaths linked to natural or man-made disasters around the world requires the humanitarian community to be better prepared in terms of geographic information.

In both cases there is a need to have access to an administrative boundaries data set which:

- is seamless and homogenous,
- is based on information validated by the countries,
- take the continuous evolution of the administrative structures into account.

In addition to that, in order to reduce duplication of efforts and make a real difference for development we would argue that this data set should also be publicly available.

In close collaboration with the National Mapping Agency (NMA) of each United Nations Member State and under the umbrella of the UN Geographic Information Working Group (UNGIWG), the Second Administrative Level Boundaries data set project (SALB) is trying to answer these needs. It aims to do so by providing free access to a working platform that covers all the UN member countries, is standardized and validated by the National Mapping Agencies, for the collection, management, visualization and sharing of data/information down to the second administrative level.

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

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

The practices applied, as well the lessons learned through the process of creating this data set, are presented in different sections corresponding to the major elements which we think are crucial in this context.

## **Standards and protocols to insure a homogeneous and seamless data set**

Defining the vocabulary, and the data set specifications with associated standards, are the first two steps of that should form the basis of any data production chain (Ebener et al., 2006), in order to insure that the resulting data set will be homogeneous and seamless. At the same time it is also important to define the mechanisms and tools that would insure proper documentation of the resulting data set.

An important amount of time has therefore been put at the beginning of the SALB project in order to define these elements.

Definitions: In the context of SALB an administrative unit is defined as being a part of a national territory for which a unique administration is observed. An administration is an administrative entity (person, or assembly, ...) that has the responsibility of functions other than juridical, electoral or statistical (including the census).

These definition allowed us to clearly make the distinction between the administrative units and the statistical, electoral or sanitary (e.g. health districts) units that can also exist within a country.

The homogeneity and seamless aspect of the data set itself are obtained through the use of specific protocols applied at different stages of the SALB process. This process goes through the following steps (Figure 1):

1. The NMA of each country is contacted in order to obtain the update of their contact information, as well as the list of administrative units at the 1st and 2nd level as observed in January 2000 in the country. Once validated and coded the table is posted on the SALB web site.
2. The NMA's assistance is then requested again in order to obtain the historic changes observed at the 1st administrative level since 1990 and at the 2nd level since 2000. The coding scheme is then integrated in these tables before posting them on the SALB web site.
3. While the historic changes are being collected, a comparison between all the maps (digital, paper or image) that have been compiled concerning the delimitation of the administrative units and the January 2000 table is made in order to identify the map or combination of maps to be used for the rest of the process.
- 4 From this point two situations may occur:
  - 4.1 we have all the necessary information and authorization for generating a digital map. In this case we can move forward with the process and start the editing work
  - 4.2 we do not have all the information and/or authorization. In this case, further search and contact are necessary in order to obtain the missing information or the authorization using for example the memorandum of understanding specifically designed for the context of the project
5. Once we have all the necessary materials, the editing process can start in order to obtain a shape file representing the administrative structure of the country for each representative period observed since January 2000.
6. These maps are then sent for validation to the NMAs in order to ensure that they reflect the correct situation observed in the country in terms of international and administrative boundaries
7. Once validated the digital map(s) and its associated metadata are posted on the SALB web site for free non-commercial download. The metadata records are also posted on the WHO GeoNetwork metadata portal giving therefore a broader access to the SALB maps.
8. The all database is then updated on a regularly basis starting again from step 1

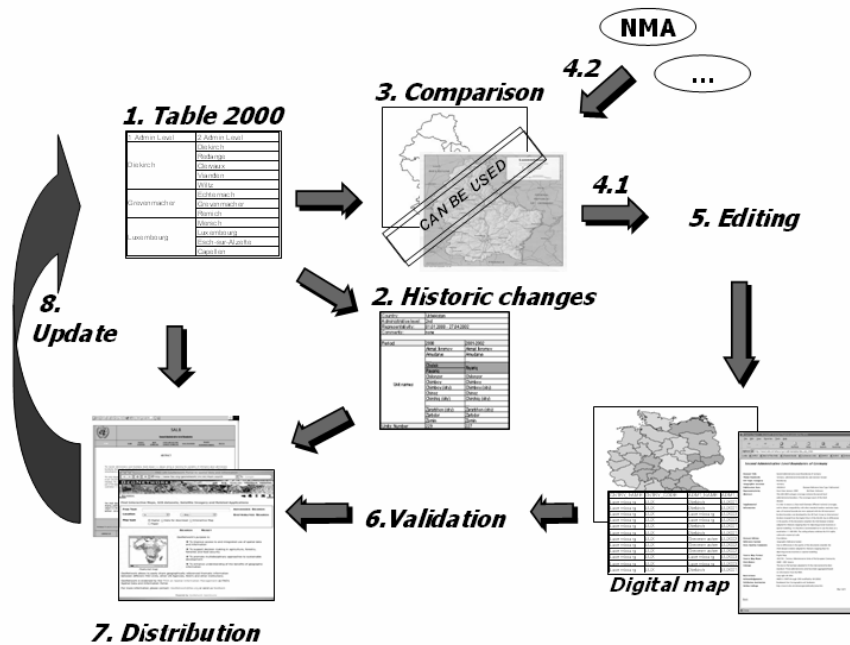


Figure 1 - The SALB process

Obtaining information and the validation of the data from the National Mapping Agencies, the question of the redistribution rights, as well as the description of the coding scheme are crucial elements in the overall SALB process. The corresponding lessons learned attached to each of these elements are reported in separate sections of this paper.

The paper or digital maps used in this process are collected from many different sources. The SALB project is therefore managing a database of paper and digital maps collected from sources located within and outside the UN. In this database the maps are grouped according to their redistribution rights in relation with the SALB data policy: differentiating the maps for which we do have the authorization of use, from the ones where this authorization is not available, either because we have a clear indication that the map cannot be distributed in the public domain, or we do not know the source of the map in question. The reason for the last category, within which most of the maps unfortunately fall, is very often linked to the frequent absence of metadata for the collected maps. It is also interesting to report here that for some countries more than 5 different sources were available (e.g. for Brazil, Colombia, Ecuador and Jamaica).

A specific protocol based on the list of administrative units provided and validated by the NMA, as well as a set of specific criteria, is then used in order to select the map, or combination of maps, to be used in the context of the SALB project. When applied to the 30 countries of the Americas for which sub-national divisions are observed, this process revealed that for 11 of them (37 %) the maps, or combination of maps, contained the situation as observed in these countries before 2000. This meant that additional information was needed in order to update these maps. Among these 11 countries are Colombia and Ecuador, which are two of the countries for which more than 5 sources were available (see above).

The editing part of the work is also done based on a protocol which can be downloaded from the SALB web site ([http://www3.who.int/whosis/gis/salb/maps/SALB\\_editing\\_protocol.pdf](http://www3.who.int/whosis/gis/salb/maps/SALB_editing_protocol.pdf)). The objective of

this protocol is to insure that the editing work performed is consistent among the technicians who apply it. It also facilitates obtaining an ESRI shape file format map corresponding to the specifications of the project regarding the international borders and the format of the Attribute Table, two elements that participate in making all the data sets homogeneous and seamless.

One of the major difficulties to be overcome at this stage of the process in trying to generate a seamless data set, particularly in the Americas, has been the treatment of the international borders. Despite the work done these past years by the UN Cartographic section in order to generate an international border standard, many issues such as the representation of disputed areas, remains to be solved in collaboration with the NMAs.

Having defined all of these elements at the earliest possible stage in the project has resulted in important cost saving over the past 5 years. It has also facilitated the transfer of the technical work to the SALB regional node, such as the one located at the UN Economic Commission for Latin America and the Caribbean (UN ECLAC).

Nevertheless, recent developments such as the public availability of the globally consistent mosaic of Landsat images, which can be downloaded from the University of Maryland web site, could be used to further improve the level integration of the SALB data set with other layers (Ebener et al., 2006).

### **Obtaining the reference information and the validation of the data from the countries**

In order to make sure that the information and data provided to the users correspond to the reality observed in the countries, it is crucial to identify the appropriate governmental body at the initial stage of the information gathering, as well as for the validation of the final product.

In terms of administrative boundaries, it was decided that the NMAs would play this role as this entity represents the governmental body most often in charge of this particular layer.

It is therefore crucial to have access to up-to-date contact information for them.

We also realized that providing up-to-date contact information of the NMAs to the international community could:

- increase the visibility of the NMAs and therefore offer the possibility of new markets for them,
- open a channel between the UN and the countries in case of urgent need for geographic information,

Unfortunately, we rapidly came to the conclusion that no institution was keeping up-to-date contact information for the NMAs in the Americas. Collecting this information, as well as the authorization to post it on the SALB website, therefore became one of the tasks of the project.

Reaching the NMAs for the first time and keeping the contact with them is often difficult. Most of the difficulties are in fact linked to technical, political or administrative issues.

In some of the offices faxes, telephones and computers are not always available, and even where they are available they do not always work. The regular mail is slow, and multiple reminders are sometimes required before any information is obtained. This makes for a lengthy process in obtaining the simplest information. For example, in the case of Dominica, Guyana, Paraguay the emails had to be sent to the contact's personal email; in Suriname only the Ministry, housed in a different building, had a fax machine.

The political and/or administrative situation observed in some countries is the cause of constant change in the personnel. Most of the time the new persons in the NMAs have not been informed of the SALB project and there is therefore a need to begin the explanation all over again. In the case of Venezuela for example, the contact person for the project retired. He gave the name of a new contact person who we briefed on SALB, but a new person was reassigned only one month later. Also, the perspective of changes to come in the NMA's structure (e.g. in Suriname), or the nomination of a new person who is not interested in participating in SALB can create important delays.

Language is also an issue. For example, the fact that the SALB web site is not available in Spanish is a problem where there nobody speaks English at the NMA. In addition to that, the fact that the roman alphabet, as used by the English language, is considered as a standard for the spelling of the unit names create also some issues for Spanish speaking countries which some times find it difficult to accept the lack of diacritics in the historic changes and the GIS map attribute tables.

Finally, building the historic changes of a country in the form provided by SALB is not an activity practiced by most of the NMAs. When we look at the number of changes that occurred in Mexico at the 2<sup>nd</sup> sub-national level since January 2000, we can easily understand the time required by the NMAs in order to find the necessary laws describing these changes.

In order to discuss these issues, obtain the necessary information and data, as well as the validation it is necessary to engage the NMAs in the SALB process, and contact them in Spanish if needed. In addition to that, being geographically located in their time zone is an additional advantage.

The recent analysis of the progresses made by SALB in Asia during the past 10 months (Ebener & Guigoz, 2006) confirmed that most productive way to engage the NMAs in the project is to meet their representative in person. This is possible during the major regional and global conferences or meetings generally attended by the representatives of the NMAs. It also very important to collaborate with regional or global committees/initiatives also in direct contact with the NMAs, especially if they are also working on the development of regional data sets for the Americas. In the case of the Americas the 3 main entities are:

- the International Steering Committee for Global Mapping (ISCGM),
- the Permanent Committee on Spatial Data Structure for the Americas (PC-IDEA) ,
- the Pan American Institute of Geography and History (PAIGH).

A step in the direction of this collaboration has been made during the 8<sup>th</sup> UN Regional Cartographic Conference for the Americas (UNRCC-A) with the acceptance of the resolution reported in Annex I.

In order to solve the language and time zone issue, SALB can count on the support provided by two entities located in the Americas: UN Dag Hammarskjold Library and the UN Economic Commission for Latin America and the Caribbean (UN ECLAC).

### **Choosing a proper coding scheme**

The administrative structure of a country evolves continuously. Any information system that wants to manage data collected over time and from different sources therefore needs to integrate this evolution. When it comes to data visualization, the same information system also needs to have access to a map, in digital format, for each period of representation observed in the data base.

The SALB data set is trying to answer this need through the provision of a specific coding scheme that allows easy identification of the administrative level the user is looking at and to follow the administrative structure changes through time. 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://www3.who.int/whosis/gis/salb/salb\\_coding.htm](http://www3.who.int/whosis/gis/salb/salb_coding.htm).

Putting a code only on validated list of administrative units also present the advantage of providing the necessary support for linking the SALB codes with other coding schemes in the form of key corresponding tables. Discussions are currently taking place with several UN agencies in order to therefore include their own codes directly in the validated maps providing therefore the possibility for the users to directly map statistical data coming from different sources. In order to be effective, this new version of the SALB shape files will also have to contain the country specific codes.

### **The question of the redistribution rights**

As indicated in the introduction we strongly believe that, in order to make a difference, the data set to be created needs to be publicly available. SALB has therefore being working on being able to publicly redistribute the data that has been validated by the NMAs.

Several elements have been put in place in order to make this possible. Firstly, a specific data policy has been established as follows:

*"The SALB dataset (codes and digital maps) is copyrighted. The owner of the data agrees on the use, reproduction, distribution, display, publication and dissemination at no cost to third parties of the SALB dataset, in any manner and in any form whatsoever, subject to the copyright and acknowledgement mentioned in the metadata. It is prohibited to change the contents of the database without consent of the owner. The SALB dataset is for non-commercial purposes only."*

The project then makes sure that all the necessary authorization has been obtained from the data owner or provider before posting it on the SALB web site. Regarding the maps, a Memorandum of Understanding has been prepared in collaboration with the UN Legal Counsel in order to be able to use maps coming from institutions willing to keep the copyright on the original data, but allowing the UN to modify it if necessary, as well as

sharing it with the public for non-commercial use according to the project policy regarding data redistribution. Despite this document, national laws sometimes act against the release of the cartographic information, especially when it is supposed to allow an external agency to modify an official national map.

In addition, the users of the SALB data set are requested to mention the copyright and acknowledgement agreed with the data owner and which is also mentioned on the SALB web site, as well as in the Metadata record of each map.

Finally, the data providers are listed on a specific page of the SALB website acknowledging their contribution.

Even if this practice slows down the entire process, it ensures that the data owners, providers and users can feel comfortable when contributing to or using the SALB data set.

### **The time and resources it takes to create such data set**

During the 8<sup>th</sup> UNRCC-A conference at the beginning of July 2005, the following information and data were reported as accessible from the SALB project web site (Ebener et al., 2005) for the 35 Members of the United Nations in the Americas:

- No NMA contact information,
- January 2000 list of administrative units for 29 countries (not including 5 without administrative subdivisions),
- No historic changes table up to January 2005,
- January 2000 map validated for 7 countries, the map for 19 countries being still under validation at that time.

The current state of progress of the project for the same set of countries is reported in Table 1. Cells in grey correspond to the progress made since the 8<sup>th</sup> UNRCC-A conference.

Country	Contact information update	List January 2000	historic changes 1st level	historic changes 2nd level	Process regarding the map, August 2006
Antigua and Barbuda	05.07.2005	no admin divisions	no admin divisions	no admin divisions	no admin divisions
Argentina	30.08.2005	Yes	1990-2002	2000-2002	Jan 2000 - Sep 2002 map under validation
Bahamas	-	Yes	1990-2006	2000-2006	Comparison map-table to be done
Barbados	28.07.2005	no admin divisions	no admin divisions	no admin divisions	no admin divisions
Belize	26.07.2005	Yes	1990-2005	2000-2005	Jan 2000 - Jul 2005 map validated
Bolivia	-	Yes	1990-2003	2000-2003	Jan 2000 - Oct 2003 map validated
Brazil	-	Yes	1990-2003	2000-2003	Map ready for editing
Canada	-	Yes	1990-2002	under process	Jan 2000 map under validation
Chile	-	Yes	1990-2003	2000-2003	Problem for obtaining a map
Colombia	25.08.2005	under process	under process	under process	Jan 2000 map under validation
Costa Rica	-	Yes	1990-2004	2000-2004	Jan 2000 - Apr 2004 map under validation
Cuba	-	Yes	1990-2003	2000-2003	Problem for obtaining a map
Dominica	09.08.2005	no admin divisions	no admin divisions	no admin divisions	no admin divisions
Dominican Republic	-	Yes	under process	under process	Problem for obtaining a map
Ecuador	13.12.2005	Yes	1990-2005	2000-2005	Missing information
El Salvador	-	Yes	1990-2002	2000-2002	Jan 2000 - Aug 2002 map under validation
Grenada	28.07.2005	no admin divisions	no admin divisions	no admin divisions	no admin divisions
Guatemala	-	Yes	1990-2004	2000-2004	Jan 2000 - Jan 2004 map under validation
Guyana	26.08.2005	Yes	1990-2005	2000-2005	Jan 2000 - Aug 2005 map under validation
Haiti	-	Yes	1990-2003	2000-2003	Jan 2000 - Jan 2003 map under validation
Honduras	14.02.2006	Yes	1990-2006	2000-2006	Jan 2000 - Apr 2004 map under validation

Jamaica	-	Yes	1990-2002	no 2nd admin level	Jan 2000 - Nov 2002 map validated
Mexico	15.03.2006	Yes	1990-2006	2000-2006	Problem for obtaining a map
Nicaragua	07.02.2006	Yes	1990-2006	2000-2006	Jan 2000 - Feb 2000 map under validation
Panama	-	Yes	under process	under process	Jan 2000 map under validation
Paraguay	21.02.2006	Yes	1990-2003	under process	Missing information
Peru	19.12.2005	Yes	1990-2003	2000-2003	Jan 2000 - Mar 2003 map under validation
Saint Kitts and Nevis	29.03.2006	Yes	1990-2006	no 2nd admin level	Jan 2000 - Mar 2006 map validated
Saint Lucia	28.07.2005	Yes	1990-2005	no 2nd admin level	Jan 2000 - May 2005 map validated
Saint Vincent and the Grenadines	28.07.2005	no admin divisions	no admin divisions	no admin divisions	no admin divisions
Suriname	-	Yes	1990-2002	2000-2002	Jan 2000 - Jul 2002 map under validation
Trinidad and Tobago	-	Yes	1990-2004	no 2nd admin level	Jan 2000 - Nov 2002 map validated
United States of America	21.03.2006	Yes	1990-2006	2000-2006	Jan 2000 - Dec 2002 map validated
Uruguay	19.12.2005	Yes	1990-2005	no 2nd admin level	Jan 2000 - Sep 2002 map under validation
Venezuela	13.12.2005	Yes	1990-2003	2000-2003	Missing information

Table 1 - State of progress regarding the tables and maps produced in the context of the SALB project for the Americas

This progress translates into the following information and data that can now be downloaded from the SALB web site:

- NMA contact information for 20 countries (+20 compared to July 2005)
- January 2000 list of administrative units for 29 countries (no changes),
- Historic changes up to January 2005 for 10 countries (+10),
- January 2000 map validated for 7 countries (no changes), the map for 21 countries (+2) being still under validation at that time.
- Map up to January 2005 validated for 2 countries (+2)

The countries where significant progress has been made are: Belize, Bahamas, Brazil, Ecuador, Guyana, Honduras, Mexico, Nicaragua, Saint Lucia and the United States of America.

As we can see, the progress made over one year remains limited. The main reason for this is that we are still in the process of making the SALB project reflect the present which is very time consuming. It is estimated that once this first step will have been reached, the yearly update of the NMA contact information, of the historic changes as well as the map will not take more than 6 months for the 35 countries located in the Americas.

In terms of cost, it is estimated that US\$60,000 was necessary in bring the project to the current status for the Americas, and that a further US\$25,000 will be needed in order to make the SALB data set reflect the present situation.

Nevertheless, it is also estimated that once this is done, the yearly update of the database for the Americas would only cost US\$15,000.

## **Conclusion**

The experience and lessons shared in this paper demonstrate that it is possible to put a collaborative process in place for the generation of a globally consistent georeferenced database of the administrative boundaries.

It also demonstrates that it is possible to make the data set publicly accessible, a key element in ensuring wide usage and consequently possible improvement of data integration among the users.

The return on investment once the process has reached the phase of the yearly update is important (reduction of the duplication of efforts, of the costs,...), and therefore justifies the use of resources and effort which were necessary at the beginning of the process. This type of data set will improve data analysis capacities, since the international community will be using information provided by the country itself. It will also increase the interoperability of the data collected at the sub-national level by different UN agencies, as it allows the data for each country to be integrated into the regional picture in a seamless way.

This project also demonstrates that the participation of the National Mapping Agencies is crucial in order to insure the validity of the information and data provided to the users, and that this participation is possible.

The questions that we can therefore legitimately ask are why did we not start this process earlier, and also why is it not already in place for other layers?

The most important part of the answer is the difficulty that institutions, such as the UN agencies, have in committing resources at the beginning of this type of process. The SALB project would never have taken off if the World Health Organization and the

United Nations Fund for International Partnerships (UNFIPS) had not provided the necessary resources at the inception. It was only after the project has reached a certain level of maturity, that it was possible to find other sources of funding within and outside the UN.

In addition to that, the notion of data preparedness did only appear within the UN these past years. Before that, the system tended to be mostly responsive, looking for data only at the time of a humanitarian crisis. This situation has now changed and many UN agencies, such as WHO or OCHA, are now integrating a preparedness component in their activity.

It is therefore hoped that this project will serve as an example for the development of other globally consistent and publicly available layers to form the data set necessary for the proper functioning of a global SDI at the service of the population. As there are challenges that we will have to face in the future, this process should start as soon as possible and must include a coordination system to ensure the compatibility between these layers.

### **Acknowledgements**

The authors of this paper would like to take the opportunity of this paper to first thank all the National Mapping Agencies who have already provided the SALB project with the necessary information and data. This project would not be possible without their contribution.

They would also like to thank the institutions and individuals who have contributed to the progresses of SALB in the Americas this past year. Hoping that nobody has been forgotten these thanks go to:

- USAID and UNHCR for the funding provided to the project,
- the WHO Office of the Legal Counsel for their help regarding the representation of disputed areas and borders in question and other related issues
- the International Steering Committee for Global Mapping (ISCGM) for its support in contacting countries,

### **References**

Ebener S., Brookes B., Guigoz Y, El Morjani Z., Chang S, 2005, The SALB project: Current Status and new challenge for the Americas, 8th UN Regional Cartographic Conference for the Americas, New York, June 27-July 1, 2005

Ebener S., Guigoz Y., 2006, Data preparedness in Asia: Where we are with SALB 10 months after the last PCGIAP Executive Board Meeting, 17<sup>th</sup> UNRCC-AP Conference, Bangkok, Thailand, September 2006.

Ebener S., Stokes C., Lance K., Terborgh C., 2006, Moving towards GSDI: Improving data integration at the global level", 9<sup>th</sup> Global Spatial Data Infrastructure (GSDI) Conference, Santiago de Chile, November 2006.

University of Maryland Landsat images download page,  
<http://glcfapp.umiacs.umd.edu:8080/esdi/index.jsp>

WHO GeoNetwork metadata portal, <http://www.fao.org/geonetwork/srv/en/main.home>

## **Annex 1 - Resolution accepted during the 8<sup>th</sup> UNRCC-A Conference, New York July 2005**

### **Resolution VI -Global mapping and Second Administrative Level Boundaries projects**

#### **The Conference,**

**Recognizing** that the availability of fundamental framework data set, such as national administrative divisions, are crucial for the analysis and management of socio-economic phenomena,

1. **Welcomes** the efforts of the Global Mapping project and those of the UN Geographic Information Working Group (UNGIWG) Second Administrative Level Boundaries project (SALB) in the generation of global seamless data sets of importance for the Americas;
2. **Recommends** that Member States, particularly through their National Mapping Agencies (NMAs) and notably with the help of the Permanent Committee on Spatial Data Infrastructure for the Americas (PC-IDEA) and the Pan American Institute of Geography and History (PAIGH), support and participate actively in the SALB and Global Mapping projects by providing the necessary data and information they require;
3. **Further recommends** member states in the Americas take full advantage of participating in the Global Mapping project for capacity building to help establish National and Regional Spatial Data Infrastructures in the region."