GEOGRAPHICAL NAMES IN THE ERA OF BIG DATA A NECESSARY UPDATING IN THE CURRICULUM DESIGN OF THE GEOSCIENCIES

Adriana VESCOVO [1]

UDC: 81'373.21:006.32]:[528.935:004.62

SUMMARY

After a long stage of low visibility, place names are revitalized by the digital revolution and globalization. As part of the map, geographical names were also affected by the technological and paradigms changes that revolutionized not only the mapmaking process but also the management of spatial information in all its stages. In this new context, the interoperability requirements of the data intended for the Spatial Data Infrastructure and for multiple users and services drove the normalization processes. Geographical names are not alien to this need. In another aspect, new cultural paradigms associated with globalization joined to renew interest in geographical names in the last half century: the revaluation of local identities, open data policy and the collaborative model. The objective of this article is to stimulate the reflection on geographical names in this new context, their special meaning as cultural heritage and the need and benefits of their national and international standardization. And, consequently, the inclusion of geographical names in a necessary revision and update of the geoscience curricular designs.

Key words: Geographical Names Standardization; toponymy; national cultural heritage; Big Data; curriculum design; public policies.

A NEW TECHNOLOGICAL CONTEXT

No one doubts that place names are an essential element of cartography. However, geographical names (GNs) did not always receive the necessary attention and were treated with ups and downs even by national mapping

[1] **Adriana VESCOVO**, avescovo@ign.gob.ar, Instituto Geográfico Nacional (IGN), Departamento de Información Geográfica; Address: Avda. Cabildo 381, C1426AAD Ciudad de Buenos Aires, Argentina Tel. +54 11 45765576-Ext 190; adriana.vescovo2012@gmail.com, Centro Argentino de Cartografía (CAC) Member of International Cartographic Association (ICA)
Address: Avda. Cabildo 381, C1426AAD Ciudad de Buenos Aires, Argentina; Tel. +54 11 45765576Ex184

production agencies. This situation seems to be reversing as a result of digital revolution.
The so-called "geographical explosion" broke out characterized by the arrival of computers, Internet and the World Wide Web service in 1990.
As part of the map, the GNs were also affected by these exceptional technological and paradigm changes that revolutionized the mapmaking process as well as the management of spatial information in all its stages.
As a result of this new digital era, the paper map ceased to be the main objective of geospatial production, now aimed at the development of Geospatial Databases (GDB) and Spatial Data Infrastructures (SDI), which – in this new cycle of Geospatial Information (GI) - many new products, applications and services will be developed in the most diverse formats and supports.
Within this new paradigm, almost all of the features that constitute the GDB and the SDI are associated with a name. In this sense, the GNs are still present throughout the entire life cycle of the GI, from its original capture in the field to its rapid and massive expansion.
From the point of view of the mapmaking process, the digital era allows a permanent update of the GDB and all the resulting by-products, a situation that requires, however, careful attention.
In another sense, the GNs, which reached a very slow and limited diffusion through the paper support, are now immediately, easily and quickly disseminated. And, in little more than a second through the web, they become a gateway to other billions of connections and contents, which go well beyond the place name itself. This virtuous nexus of place name and virtual space has called the attention of the world of marketing and investment, giving the GNs a previously unthinkable economic value.

THE AGE OF BIG DATA

The digital revolution marked the beginning of the information age. In the field of GI, technological advances in the last half century, accelerated in recent decades, expanded and enriched the different tools, platforms and methodologies of capture and management of geographic data: global positioning systems, digital aerial photogrammetric systems, satellite remote sensing, including private and nanosatellites; the use of UAV's, the LIDAR technology; as well as the management of information through GIS, models, maps and countless applications and services.
For each new advance, the volume of data captured, processed and transmitted is multiplied exponentially. GI also joins the Global Big Data. GI Meetings, forums and congresses are aware of this power. The characteristics of Big

Data are usually summarized in the 5 V: volume, velocity, variety, veracity and value. This complexity raises other needs: precision be updated, availability… Never has society had access to such a gigantic volume of information. But its real utility is only achieved together with the greatest demand: access the source and maintain the data quality.

OPEN DATA

New paradigms of world culture add to the technological revolution. Access to information, until yesterday reserved or limited by very high costs, is currently opened as a citizen right. The open data policy marks a new treatment of information in the governmental and scientific-technical field. At the same time, massive access to the Internet, personal computers and other devices such as mobile and smart phones is spread throughout the world. The Central Intelligence Agency (CIA) estimated for 2018 an Internet access of 3.174 million users, equivalent to 43% of the world's population. According to the Mobile Economy Annual Report 2019, 5.100 million people had a mobile phone line by the end of 2018, which represented approximately 67% of the world's population. Access to technology and the right to information open a new trend that aims to offer more and more digital services through free platforms. These routes allow an increasing access to visualize, download and share geospatial information, together with tools for its treatment.

COLLABORATIVE MODEL

The old “top-down” model in which mapping agencies or government institutions were the only generators of geographic information has been diluted as a result of the changes in the digital era. The massive access to geographic information and a more fluid and open communication between user and generator, between authority and citizen, have resulted in the collaborative model. This new trend requires an adaptation by both sectors of the information circuit: specially, an open mind of the bodies responsible for generating the GI, and trained personnel and regulations that include criteria for analyzing the quality of the data, prior to its inclusion in the information flow. An active user has proved to be an important node in the generation of information, both in the elaboration of mapmaking process and in the provision of geo-referenced data for the management of humanitarian aid or the risk of natural disasters such as hurricanes, earthquakes, volcanism, floods or Tsunamis, among others. The GNs also accompany features here, and are especially important in these circumstances. In this sense, there is an

increasing interest as well as the bibliography referring to the analysis of the collaborative model and its problems.

THE MAP IN THE ERA OF BIG DATA

The digital era and the emergence of GDBs, GIS and other multiple digital applications and services questioned the current value of the map. However, the analysis of reality confirms the opposite.
Access to digital and mobile devices has increased their use. Eighty percent of people who have smart devices use some type of map for their own location, geo-referenced searches or to reach their destination. Since most of the information we handle has spatial location, the map is currently used in everyday life like never before. It remains mandatory - in paper support- in sea and air transport, and proved to be essential and practical in any other transport or sporting activity. This, not to mention the infinity of sectors of the economic and governmental activity in which the map increased its use as a result of new technological tools or the development of special applications.
Examples of this are its application in smart agriculture, in the delivery of documents or merchandise through the use of UAV's, control of the supply of drinking water or public lighting, parking, waste management and numerous other urban services (sanitary, educational, tourist, etc.) as well as in innumerable applications developed with destination to the user from tools and digital devices.
But, beyond this everyday use, maps play a unique and irreplaceable role: they allow us to identify features in their environment and analyze spatial relationships.
Summing up, like no other tool, maps make possible a visual synthesis of the essential information in the era of Big Data.

COPERNICUS PROGRAM: A CASE STUDY

Considered the best cartographic system in the world, the *Copernicus Program* of the European Union, as a product resulting from this special moment that the GI is going through, can be taken as a case study. In that sense, it means the choice of remote sensing as the most efficient and transparent tool for accessing a wide range of data on the Earth condition. *Copernicus* captures and delivers, almost in real time, huge amounts of global data from satellites and terrestrial, aerial and maritime measurement systems for the understanding and sustainable management of the planet. According to its latest 2019 report (Copernicus Market Report 2019), the volume of

downloads was 28 TB through its Data and Information Access Services (DIAS), with a growth of 133% compared to 2016.

About the 5 V of *Copernicus Program* Big Data, The Visionary Paper From Copernicus Big Data to Extreme Earth Analytics (Koubarakis et al., 2019) offers the following information:

"Volume: The repository of Sentinel products managed by the European Space Agency (ESA) has so far published more than 5 million products, and it has more than 100 thousand users who have downloaded more than 50 PB of data since the start of the operations of the system. This volume will increase in the following years, as new Sentinel satellites are launched.

Velocity: Copernicus data has to be delivered and processed in a short time frame to allow the provision of 24/7 information to users requiring fast responses. By the end of 2016, 6 TB of data were generated and 100 TB of data were disseminated every day from the Sentinel product repository. These rates will increase in forthcoming years as new Sentinel satellites are launched.

Variety: The Sentinel satellites have different types of sensors (e.g., radar and optical) and different levels of processing (from raw data to advanced products). Moreover, datasets used for geospatial applications can be not only satellite data but also aerial imagery, in-situ data and other collateral information (e.g., public government data). This wealth of data is processed by Earth Observation actors to extract information and knowledge. This information and knowledge is also big and similar big data challenges apply. For example, 1PB of Sentinel data may consist of about 750.000 datasets which, when processed, about 450TB of content information and knowledge (e.g., classes of objects detected) can be generated.

Veracity: Decision-making and operations require reliable sources. Thus, assessing the quality of the data is important for the whole information extraction chain.

Value: The extraction of information from the Copernicus data has direct economic benefits for Europe. Several economic studies have concluded that the Copernicus programme has the potential to significantly impact job creation, innovation and growth. The Copernicus Market report of 2016 estimates that the overall investment in Copernicus will reach EUR 7.4 billion in the years 2008-2020, while the cumulative economic value generated by it in the same period will be around EUR 13.5 billion, and it will support 28.030 job years in the Earth Observation sector. (Koubarakis et a.l, 2019)

As an example of open data policy, the Program guarantees full, open and free access to *Copernicus* data and information, and is the European contribution to the Global Network of Earth Observation Systems (GEOSS). *Copernicus* also offers an example of a collaborative paradigm through its Thematic Exploitation Platforms (TEPs). Its multiple end users include public administration, senior students, researchers, commercial companies, entrepreneurs, NGOs and citizens, both European and worldwide.
For all the above, the *Copernicus Program* could be incorporated as curricular content in Geography, Cartography and associated careers, not only for its use but also for its conceptual and contextual analysis of current trends in the field of geosciences.

A NEW CULTURAL CONTEXT

Beyond the impact of technological changes, geographical names have also been affected by a new cultural context characterized by the revaluation of local culture in response to decolonization processes and the phenomenon of globalization. This trend was reflected in the recovery of local languages and their inclusion as official languages.
In the same direction, the United Nations Educational, Scientific and Cultural Organization (UNESCO) approved in October 2003 the Convention for the Safeguarding of the Intangible Cultural Heritage. And last January, it inaugurated the International Year of Indigenous Languages 2019 (IY2019), as a reserve of culture, knowledge, values and identity. In Europe, Decision 2017/864 established the Declaration of the European Year of Cultural Heritage 2018.
This trend was very widely reflected in the GNs, with the recovery of place names in the indigenous languages, both from countries and cities, streets, squares or other features. Just to mention the best known examples worldwide - among thousands of cases - can Sri Lanka, Myanmar, Beijing, or Kolkata be cited. And more recently the recovery of the name of the highest mountain in the United States of America, Mount Mc Kinley, in Alaska, which in 2015 was renamed by President Barak Obama as Denali, “the highest”, original name in the language of the indigenous local community.

STANDARDIZE GEOGRAPHICAL NAMES

In the current context in which the GI is developed, geographical names offer two fields of analysis: as basic and fundamental data of the IDE and as cultural heritage. The digital revolution promoted the standardization processes of the

entire GI in order to allow interoperability of data and the development of applications and services. Geographical names are no stranger to this requirement.

Although the initiatives to establish a single written form for each geographical name date back from the end of the 19th century, the United Nations became the permanent international discussion space on this subject since 1948. The First United Nations Conference on the Standardization of Geographical Names was held in 1967, culminating in the creation of the United Nations Group of Experts on Geographical Names (UNGEGN), an organization that will dedicate half a century of existence to this arduous task. The goal of UNGEGN is to create a unique record of all official place names of the earth. But standardizing the GNs implies achieving a unified record of the writing of all world toponyms in a global scenario that includes about 7.000 languages, among which a large part of them corresponds to languages without writing, to which different types of alphabets are added. Papua New Guinea is the country with the greatest linguistic diversity: in an area of less than half a million square kilometers (somewhat less than Spain) some seven million people speak more than 800 different languages. Greenlandic is today the only official language of the Danish island territory. Although it is spoken by only about 54.000 people - equivalent to half of the amount of inhabitants of the city of Winterthur - this language has three dialects.

In the midst of this "*Babel Big Data"*, the ambitious goal of the UNGEGN can only be achieved by means of the national standardization of the GNs, the process in which each country decides its own place names following standardized principles, policies and procedures. By having standardized Geographical Names National Gazetteers will be the source of basic and official information on the construction of a global SDI.

For more than 50 years, UNGEGN pursued its objective through the development of 30 sessions and 11 conferences, and their respective Resolutions and documents resulting from the Meetings of the Working Groups, the advisory work of its Task Groups, Training courses, publication of manuals and brochures, as well as lists of geographical names and reference information of countries, digital files and compilation of documents included in its website. The UNGEGN has a structure of 24 Geographical-Linguistic Divisions, nine Working Groups and two Task Groups. It brings together more than 400 members from more than 100 countries, which includes geographers, cartographers, linguists, planners and specialists in different geosciences. The UNGEGN is one of the four bodies composed of government experts from the Economic and Social Council (ECOSOC), one of the six main bodies of the United Nations.

NEED AND BENEFITS OF GEOGRAPHICAL NAMES STANDARDIZATION

The hard work of UNGEGN stems from the conviction of the advantages, benefits and the need to standardize the Geographical Names, an objective that is renewed in the current technological context in which the GI is developed.
At the national level, these benefits are associated with the improvement of efficiency in communications and resource management that generates a unique database of GNs in the management of public policies and private sector.
These advantages are clearly visible in the cartographic activity and in the IDE of the different levels of public administration; in statistical work; disaster management; defense, security and land, sea and air communications; humanitarian aid; food safety; urban, territorial and environmental planning; tourism, industry, and commerce; academic and scientific activities; strategic planning; and development cooperation, among many others.
The establishment of an authority and of a recognized process of validation of geographical names at the national level prevents overlapping and confusing management of processes, consequent saving of human resources, capital and time, as well as human and structural risks and losses arising from inconsistencies
The official and uniform use of the writing of GNs also facilitates the communication and coordination of regional and international projects.

GEOGRAPHICAL NAMES IN THE 2030 AGENDA

In this regard, in order to serve the current global objectives of the United Nations more efficiently, the UNGEGN was dissolved and recreated according to ECOSOC Resolution 2018/2 of November 2017.
The "new" UNGEGN maintains its initial objectives and structure, updating its work methodology and dissemination formats of its results.
Under this new regulation, and organized by the United Nations Statistics Division of ECOSOC, UNGEGN celebrated its 1st. Session at the UN headquarters in New York, between April 29 and May 3, 2019.
Undoubtedly, GI is an essential input in all United Nations programs. In this regard, the 2030 Agenda and the achievement of the Sustainable Development Goals summarize the main goals of this international organization to be met in the respective states, and cannot be achieved without accessible and quality GI. In this regard, the joint work of UNGEGN and the United Nations Committee of Experts on Global Geospatial Information Management (UN-

GGIM) will be essential. The importance of GNs in the 2030 Agenda was addressed in a special presentation and in different exhibitions during this 1st. Session of the "new" UNGEGN, as well as in other previous meetings, accompanying the growing interest of the subject throughout the world.

Figure1: The 1st. Session of the new stage of UNGEGN (New York, 2019)

The 1st. Session was attended by 264 participants. It included representatives from 70 Member States, an Observer State, and representatives of the International Cartographic Association (ICA) and the International Geographic Union (IGU).

Figure 2: Geographical Names and the 2030 Agenda. (New York, 2019)

The link between the Geographical Names and the 2030 Agenda had a prominent place during the 1st.Session of the new UNGEGN.

GEOGRAPHICAL NAMES AT PRESENT

Geographical Names continue incorporating new themes into their traditional discussions: migrations, gender issues, risk management, trademarks, independence movements are just a few examples. There are still changes in the names of countries (North Macedonia, Czechia, Eswatini); wind farms are named; and streets and squares take the names of women or recent political events; original names are recovered in Estonia; the name of Pablo Neruda for Santiago de Chile Airport, the emergence of globalized maps on official digital sites, and the choice of programs to manage GDB are discussed ... Geographical names accompany the historical and cultural evolution in its broadest sense.

Figure 3: Geographical Names in road signs.
Some examples in Egypt, Ethiopia, Canada and Croatia are shown on the pictures. UNGEGN participates in the International Organization for Standardization (ISO) in the international representation by alphabetic codes. In this last picture: HR for Croatia

Meanwhile, the traditional problems of the GNs are maintained, such as those related to solving writing problems in GNs of native languages; the use of exonyms or endonyms in the Atlases; GNs in cross-border geographical objects, the use of GNs in legal documents or translations; their inclusion in scientific and technical publications; GNs of features under territorial dispute; the application of new technologies in the registration of the GNs during field works; or decisions in the normalized management of the GNs: proposals of new names, changes of GNs, use of controversial names…
Or there are other specific studies: studies on the inclusion of names of flora or fauna species in local toponymy allow recreating the evolution of environments and land use; analyses of toponymy in native languages help manage disaster risk; GNs studies of foreign languages explain the historical evolution of territorial occupation and migrations; social perception assessments of Geographical Names of tourist areas define promotion plans... Closely linked to the human spirit, Geographical Names are full of meanings.

GEOGRAPHICAL NAMES AND CURRICULAR UPDATE

The evolution of Geographical Names is part of the development of GI. GNs are an essential part of mapping and the entire management of the GI. As discussed in other MMM Geo Information articles, the handling of the multiple technological tools that currently integrate the cycle of mapping production and GI requires specific training. As an essential element of cartography and basic and fundamental data of the GDB, the GNs have their own problems. Its standardization is indispensable to allow the interoperability of GI. The quality of the data is the main challenge in the era of Big Data. As an integral part of most of the Features of a GDB, the management of the GNs requires more than technical expertise. The normalization of GNs also requires human resources formed in concepts, contexts and trends: officials and operators who are aware of the importance, the need and the benefits of standardization of national GNs and of maintaining the quality of the data throughout the entire production cycle. The standardization of the GNs also requires human resources formed in principles, policies and procedures. Together with the technical training that offers the necessary skills for the management of technological tools, it is necessary to train in conceptual knowledge. At a higher level of training, the inclusion of content associated with the importance of GNs in public policies will also be necessary. Contents will be adapted to the different degrees levels. The enormous changes produced in the field of GI make it necessary to update the curricular designs of all careers associated with cartography, geography and geosciences. Geographical Names should be incorporated into all of them.

The case studies can be a very suitable support to bring the concepts to real practical examples.

CONCLUSIONS

Geographical Information has always been strategic knowledge. As part of this information, geographical names are an essential element of public policies, as georeferenced data and as identity provider. As intangible cultural heritage, geographical names contain in themselves a social and economic value, as well as rich information regarding the context of their origin and their changes. As a link between generations, they create a network of belonging and continuity in the midst of a global, fragile and mobile world. Maps, in the past, and databases at present, can be important reserves of that living memory that technology allows to disseminate as never before. In the era of Big Data, access to information seems to be secured through the availability of technological tools and the open data paradigm. In this same current context, information and knowledge are a collective creation. In a world with a volume of data and geographical information as never before, quality is one of the greatest challenges of geosciences and one of the greatest responsibilities of the present time society. The use of standardized geographical names ensures a reliable quality geographical information flow. The academic world must train capable human resources to respond to these challenges.

REFERENCES

1. Buzai, G. (1999) Geografía Glob@l. El paradigma geotecnológico y el espacio interdisciplinario en la interpretación del mundo del siglo XXI. Buenos Aires, Lugar Editorial. pp216.
2. GSMA (2019) Mobil Economy 2019. Mobile Economy Annual Report. https://www.gsma.com/r/mobileeconomy/ [accessed 05-10-2019]
3. ICOS, (2010) XXIV Congreso Internacional de ICOS sobre Ciencias Onomásticas (2010), Universidad de Barcelona. http://blog-idee.blogspot.com/2010/09/xxiv-congreso-internacional-de-icos.html [accesseed 16.08.2019]
4. Koubarakis, M. et al.(2019) From Copernicus Big Data to Extreme Earth Analytics.Visionary Paper. Published in Proceedings of the 22nd International Conference on Extending Database Technology (EDBT), March 26-29, 2019, ISBN 978-3-89318-081-3 onOpenProceedings.org.

https://openproceedings.org/2019/conf/edbt/EDBT19_paper_321.pdf [accesseed 05.10.2019]

5. PwC (2019)Copernicus Market Report 2019. https://www.copernicus.eu/sites/default/files/2019-02/PwC_Copernicus_Market_Report_2019_PDF_version.pdf [accesseed 05.10.2019]
6. Ruiz Almar, E. (2014) *La explosión geográfica: las tecnologías de la Información Geográfica y su impacto en la Geografía y la Cartografía*. Revista do Departamento de Geografia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, volumen especial Cartogeo, pp. 286-301 ISSN 2236-2878 DOI:10.7154/RDG.2014.0114.0011Brasil. http://www.revistas.usp.br/rdg/article/view/85556/88345 [accessed 04.06.2019]
7. Ruiz Almar, E. (2010). Consideraciones acerca de la explosión geográfica: geografía colaborativa e información geográfica voluntaria acreditada, GeoFocus- Revista Internacional de Ciencia y Tecnología de la Información Geográfica- (Artículos), nº 10, pp.280-298. ISSN: 1578-5157. http://www.geofocus.org/index.php/geofocus/article/view/ [accesseed 24.06.2019]
8. UNESCO, Convention for the Safeguarding of the Intangible Cultural Heritage https://ich.unesco.org/en/convention [accesseed 04.06.2019]
9. UNESCO, 2019 International Year of Indigenous Languages https://en.iyil2019.org/about/ [accesseed 04.06.2019]
10. UNGEGN, United Nations Group of Experts on Geographical Names https://unstats.un.org/unsd/ungegn/ [accesseed 15.08.2019]