Design and implementation of the web Linguistic and Ethnographic Atlas of Colombia

Luz Angela Rocha S.,[a] Johnatan Bonilla,[b] Julio Bernal,[b] Catherine Duarte[c] and Alejandro Rodriguez[c]

[a] *Universidad Distrital, Grupo de Investigación NIDE, Bogotá, Colombia;* *lrocha@udistrital.edu.co*
[b] *Instituto Caro y Cuervo, Grupo de investigación Lingüística de Corpus, Bogotá, Colombia;* johnatan.bonilla@caroycuervo.gov.co,
julio.bernal@caroycuervo.gov.co
[c] *Universidad Distrital, Semillero de investigación SDI, Bogotá, Colombia;*
cduartec@correo.udistrital.edu.co, alerodriguezc@correo.udistrital.edu.co

Abstract: The Atlas Lingüístico y Etnográfico de Colombia (Linguistic and Ethnographic Atlas of Colombia), known by "ALEC" is a compilation of popular speaking Spanish of the populations of Colombia; such research was carried out for more than fifty years. The result of this work is a collection of thematic maps organized in six volumes and its supplements in analog format. In that sense was created the project entitles "Interactive ALEC" which main objective is to develop a digital and interactive web version of the ethnographic and Linguistic Atlas of Colombia (1983) and its supplements. In this way the Corpus linguistics research group belonging to the Institute Caro y Cuervo and the research group NIDE of the Universidad Distrital "Francisco José de Caldas" have been working together in the design and development of the Atlas Web, that allows the visualization and consulting of the spatial information contained in the volume III of the analog ALEC Atlas, applying concepts of Geographical Information Systems and web cartography. Therefore, the objective of this paper is to show the process of design and development of the web prototype of the ALEC as a collection of static and dynamic maps, which show spatial information, combined with multimedia content, taking into account that in addition to all maps, the total compendium includes images, illustrations, photographs, audio and text comments. Likewise, the interactive ALEC is a good example of how to use geo-technology tools nowadays, because they are essential for the dissemination of geo linguistic information through internet, achieving more access and distribution of the Atlas web.

Keywords: Linguistic Atlas web, web maps, ALEC, Colombia

1. Introduction

In order to promote the conservation, access and disclosure of the linguistic heritage and the research regarding the Spanish language in Colombia, the ALEC interactive project has the aim to develop the digital online version of the Linguistic and ethnographic Atlas of Colombia (ALEC). In the framework of this project, the research groups NIDE (Nucleus on Spatial Data) of the Universidad Distrital Francisco José de Caldas and the research group of Lingüística de Corpus (Corpus linguistics) of the Instituto Caro y Cuervo, joined efforts in order to develop the third phase, which main is objective to design, build and implement the Web Atlas for volume III of the linguistic Atlas and ethnographic of Colombia (ALEC) using free and open source software.

The Linguistic and Ethnographic Atlas of Colombia, "ALEC" is an analog Atlas published by the Institute Caro y Cuervo (ICC), Institute responsible for conducting the studies and research in literature, philology and linguistics of the Spanish language and the native languages of Colombia, focused on to make research and dissemination of the culture of reading (Instituto Caro y Cuervo, 2016). The work was directed by the researcher Luis Flórez, who due to their studies in Dialectology was the responsible of design and implement the methodology

for the elaboration of the Atlas. Also he was, who created the first questionnaires for the collection of linguistic information, which were tested in 1956 in the Department of Cundinamarca for the preparation of the first volume and the last of the six volumes was published in 1983. (Flórez, 1983).

The ALEC consists of the following features and elements which were collected during 27 years. (Flórez L. M., 1983)

- Participation of 23 researchers and 2217 informants
- The country was divided in 264 locations called "Localidades"
- Consists of six illustrated volumes, supplement and manual
- It has 1696 sheets corresponding to 1,523 maps (50 x 35 cm). (Figure 1).
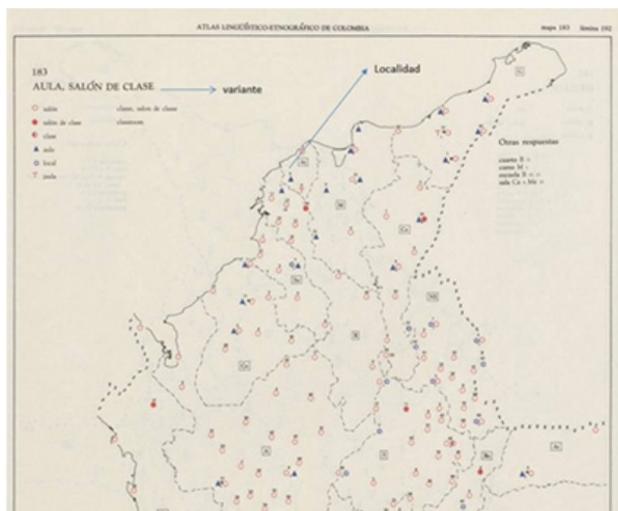- It has 1112 audio files and more than 15,000 photographs (unpublished)

Fig. 1. Example of a Map sheet of ALEC (Source: Atlas lingüístico -Etnográfico de Colombia. ALEC, 1983)

Due the fact the ALEC is a mandatory reference Book for consultation regarding the Spanish of Colombia, the ICC determined that it should be more widespread, because sometimes for their users and researchers is complicated have access to it, not only by the size of the Atlas but also by the number of volumes that constitute this publication, so important for the linguistic in Colombia. For this reason it was decided to make a web Atlas of the ALEC, allowing:

– take advantage of the tools offered by geo-technologies for the online publishing of maps of the ALEC in the form of interactive dynamic web maps
– use multimedia tools for the dissemination of the ALEC and its components unpublished such as audio files, photographs and videos
– dissemination of the interactive ALEC online, to be consulted both by users experts as those that do not have knowledge of the theme
– minimize the cost in the publication
– facilitate a possible updating

According to Hoch & Hayes (2010) the inclusion of computer technology for the storage of data on linguistic research and production of linguistic atlas begins in the seventies. The first publication on the use of Spatial Da-ta Bases and modeling, was made by Alan Richard Thomas in 1980 in the work "Areal Analysis of Dialect Data by Computer: to Welsh Example". In this work, Thomas (1980) presented an early example of the possibilities of the use of GIS to measure the spatial correlation in digital data regarding their linguistic research. By assigning numerical values to the use of the Welsh words in different regions on a map, he managed the storage and display of multiple linguistic attributes, but with the limitations of the era such as the administration of large amounts of text or difficult allocation of special characters (Hoch & Hayes, 2010). Years later, inspired by the works of Thomas (1980), Pederson designed a digital spatial database and subsequent viewing of the Linguistic Atlas of the Gulf

States (Pederson, McDaniel, Baley, Basset, & Dent, 1986), according to Hock & Hayes (2010) Pederson (1986) work was an improvement on the use of GIS tools, since many of their single develop- ments would be available in a graphic interface until almost a decade later. On the other side in the Latin Amer- ican context the first linguistic atlas based on Spatial Database was the linguistic and ethnography Atlas of the province of Santander which Alvar (1991) described the database developed as a very useful product to facilitate the creation of maps on demand and the preparation of indexes used in the interpretation of linguistic Atlas.

From these first efforts until now, have been published linguistic Atlas based in spatial databases and GIS technologies in two different ways: the first one is a linguistic Atlas designed since the beginning using the spatial database modeling and GIS tools, such as the Linguistic Atlas of the Middle and South Atlantic States, of Schneider and Kretzschmar (1989). In the same way online, it is possible to get: the Linguistic Atlas (www.lap.uga.edu), Ethnologic (www.ethnologue.com), Modern Language Association Language Map (MLA, 2009), the Atlas linguistic and ethnographic of Castilla La Mancha of Garcia Mouton and Moreno Fernández (2003) and The World Atlas of Language Structures Online (Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013). The second way is relating with the processing of analog Atlas and to conversion to digital presentation, such as the Sprach und Sachatlas Italiens und der Südschweiz of Jaberg and Jud, and the project which aim was to elaborate in digital form the material of the Atlas Lingüístico de la Peninsula Ibérica (ALPI). GIS technologies are also used in the ALPI to generate an interactive application on the Internet (alpi.csic.es), through a web portal with web services based on geographical location, according to the standards of the Open Geospatial Consortium (OGC) and combined with searches at the thematic database (Heap, 2003).

2. Methodology

The design and implementation of the Atlas web of the ALEC is framed in the "Interactive ALEC project" which was conceived in three phases as is shown in figure 2:

– phase 1: which aimed to incorporate all the information of the ALEC in Excel files including the metadata in the six volumes and the Handbook information
– phase 2: this phase is aimed at the design and implementation of the Geographic Information System of the ALEC, which is the basis of the creation of the web Atlas of ALEC
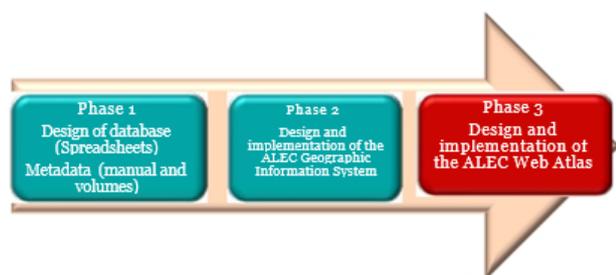– phase 3: is the phase which develops and implements Atlas web of the ALEC

Fig. 2. Phases of the Interactive ALEC project. (Source: Authors)

2.1 Phase III: Atlas web of ALEC

The Atlas web is considered as a compendium of static and dynamic web maps that present geographic information in an interactive way for different users with access to the network, combined with multimedia content related to topics of interest (Carías, 2010). The web maps of ALEC are graphics models, characterized and distributed through the network using a browser web (Kraak, 2004), which conserve the primary elements of the analog maps of the ALEC by with the main characteristics of the dynamic maps as follows: interaction tools with the user; change of the scale of display of spatial information (zoom in, zoom out); displacement inside the map (panning); consulting information from a spatial database; showing and hiding layers of information; modify the color or appearance of the data; select spatial data according the location and other attributes of the data set; modify the dis- placement and scale of display (Fly) and consult information using hyperlinks.

The methodology used for the design and implementation of the web ALEC was based on the following activities:

Step 1: Construction of the web Atlas

– diagnostic about the information of volume III of ALEC
– definition of the architecture of the Atlas web
– definition and creation of the structure of the web Atlas
– development of the geographic visor

Step 2: definition and conceptualization of web maps

– map template definition
– definition and design of the interactive dynamic web maps
– design and creation of the symbol catalog

Step 3: implementation

– data entry and publication of web maps
– validation and adjustment of the web Atlas
– Elaboration of documents

3. Results

The main products results of the project is the Atlas web implemented in a geographic visor that was constructed using the following software tools: PostgresSQL, PostGIS, GeoServer, Html, JavaScript and Grails. Figure 3 presents an example of the prototype of web ALEC.

In addition as a result of the project all the information generated was documented, for instance: features of the web maps, software application developed for the implementation of the web Atlas, methodology used for the design, creation and implementation of the web Atlas and the results of the validation of the prototype of the web Atlas.
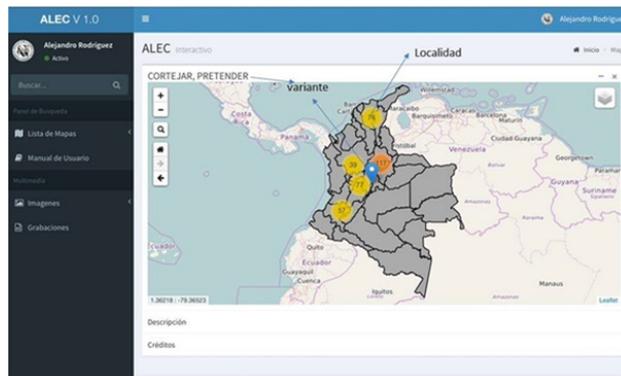

Fig.3. Atlas web of ALEC. (Source: Authors)

4. Conclusions

The ALEC interactive web Atlas is a tool that allows approaching all type of population, of different ages and generations, that makes it use and access more effective, since it is done through Internet, for which it is disseminating an easier and simple way.

On the other hand, the ALEC web Atlas, fulfills its pedagogical purpose, that is to becoming a tool for disseminating knowledge in a practical, clear and simple way.

In the same way, the use of free software for the publication of Atlas Web allows generating, a very low cost, a product that meets the characteristics and expectations of all types of users, through the net.

Finally the ALEC web Atlas shows that it is possible to use geo-technological tools for the generation of interactive linguistic maps, which is a technical progress in the field of linguistic geography.

5. Acknowledgements

6. References

Alvar, M. (1991). Geografía lingüística. Editorial Paraninfo.

Carías, L. S. (Diciembre de 2010). Atlas multimedia prototipo del Centro Histórico del Distrito Central de Honduras. (U. N. Honduras, Ed.) Revista Postgrados UNAH , 1 (4).

Flórez, L. (1983). Manual del ALEC. Bogotá.

Flórez, L. M. (1983). Atlas lingüístico -Etnográfico de Colombia. ALEC. Bogotá: Instituto Caro y Cuervo.

Heap, D. (2003). Las encuestas del ALPI a los cuarenta años de la publicación de su primer tomo. En: Actas del XXIII Congreso Internacional de Lingüística y Filología Románica (pp. 201-10). Max Niemeyer Verlag.

Hoch, S., & Hayes, J. J. (2010). Geolinguistics: The incorporation of geographic information systems and science. The Geographical Bulletin, 51(1), 23.

Instituto Caro y Cuervo. (2016). Instituto Caro y Cuervo. Obtenido de www.caroycuervo.gov.co

Kraak, J. M. (2004). The role of the map in a Web-GIS environment. Journal of geographical Systems, 83-93.

Pederson, L., McDaniel, S., Baley, G., Basset, M., & Dent, B. (1986). Linguistic atlas of the Gulf States. University of Georgia Press.