SPANISH NATIONAL ATLAS INFORMATION SYSTEM DEVELOPMENT: A STEP FORWARD

Juan José Alonso Gamo
jjalonso@fomento.es

Mar Zamora Merchán
mmzamora@fomento.es

Noelia Pérez Mayoral
npmayoral@fomento.es

Instituto Geográfico Nacional. Ministerio de Fomento
General Ibáñez Ibero, 3
28003 Madrid (Spain)

Abstract

Spanish National Atlas Information System (SIANE) is an innovative and development project that concerns with a new definition, development and maintenance of National Atlas of Spain. In addition, it gives technology support in order to produce and publish every product or service in an efficient manner. The main objectives of National Atlas of Spain have been taken into account: to give a synthetic and global point of view of spanish geographic reality and to make a contribution to territorial knowledge, scientific and technique investigation and economic and social development. However the project has also its origin in making a more modern work and avoiding, or at least getting minimized all negative factors by which the overall work is affected. System has been conceived in a way than continuous *content* updating and publishing is possible for work maintenance, using Internet as the main media and without having to wait for all contents of certain theme to be created in order to make them be published.

1. Introduction

1.1 The National Atlas of Spain

The National Atlas of Spain (ANE) is currently in a time of change. After one century of preliminary works (Instituto Geográfico y Estadístico, 1912), and more than forty years of National Atlas (1st Edition in 1965) (Vázquez Maure, 1965, 1968; Sanz Núñez 1992; Aranaz 1993), the ANE has entered the XXIst century with renewed approaches (Alonso, Del Campo, 2005; Del Campo, Romera, 2005).

Since 2004, the Strategic Planning of the Spanish National Geographic Institute includes new lines of activity related to the National Atlas. Using achieved experience from previous ANE projects, as well as new information technology media, a new

concept of National Atlas has arisen. Nowadays, the traditional atlas definition (a collection of related maps, texts and illustrations to describe all national geographic aspects) has been adapted to current times (Sancho Comins, 1993). This means that new products and services derived form ANE contents play now an important role (web services for data queries and map creation), finding in Internet an important communication channel (see, for example, the educational website 'España a través de los mapas' (Spain in maps), made in collaboration with the Spanish Geographers Association, http://www.ign.es/espmap).

Adapting production process to new technology solves some of the main aims that traditional National Atlases wanted to get:

- **Updated data**. Content design, data capture, texts and maps creation, publication process, page design, editing, printing, etc., is high time-consuming process which makes difficult to give updated information.
- **Reaching people**. Often, National Atlases are monumental (expensive) works, just affordable for libraries.

Is in this context of change where the Spanish National Atlas Information System (SIANE) has born.

2. Spanish National Atlas Information System

2.1. SIANE definition

Very attached to new technologies, SIANE is a complex system which takes into account people, data, databases, processes, applications and hardware, and serves the main aim of making National Atlas of Spain in a dynamic and optimized way (Alonso, Del Campo 2005; Del Campo, Romera 2005). It has been conceived in a way than continuous content updating and publishing is possible for work maintenance, using Internet as the main media and without having to wait for all contents of certain theme (in which the resource is classified) to be created in order to make them be published.

2.2. SIANE today

We are going to describe current situation of the system, from the point of view of applications and the two basic user profiles. An user of production subsystem, from data capture until he or she publish the resource, and another publication subsystem user who wants to access that National Atlas resource previously created.

Specific SIANE's applications are: an Excel macro for preparing data before entering the system (if needed), a Content Management System (CMS), a Map Editor and a Web application (*SIANEWEB*). Main applications (CMS and Map Editor), have been developed in Java, and Struts Framework 1.2 + AJAX has been chosen for web. Apart

from these, there are other commercial applications used for individual resource editing of texts, images, and also reference and qualitative maps. The most representative are Adobe Suite, Microsoft Office, ArcGIS Desktop and QuarkXPress.

During system analysis stage of the project, several kinds of resources were defined with a view to have a suitable resource management for production and publication environments. In order to tell the difference among resources, the concepts *class* and *type* were chosen. They are shown in the next table with the correspondent metadata profile. All resources has geographical content in a general sense, but we established separated kind of resources for data which has an inherent coordinate reference system in its structure, and other which does not.

<table>
<tr><th>RESOURCE CLASS [/ TYPE]</th><th>METADATA PROFILE</th><th>RESOURCE CLASS [/ TYPE]</th><th>METADATA PROFILE</th></tr>
<tr><td>Series</td><td>Publication series</td><td>Object / Map</td><td rowspan="2">Geographical published resources</td></tr>
<tr><td>Publication</td><td>Publications</td><td>Object / Vertical image</td></tr>
<tr><td>Group / Chapter</td><td rowspan="3">Groups (inside a publication)</td><td>Production data / Text</td><td rowspan="8">Production resources</td></tr>
<tr><td>Group / Section</td><td>Production data / Image</td></tr>
<tr><td>Group / Subsection</td><td>Production data / Statistical data</td></tr>
<tr><td>General folder</td><td>General folders</td><td>Production data / Illustration</td></tr>
<tr><td>Object / Text</td><td rowspan="8">Published resources</td><td>Production data / Table</td></tr>
<tr><td>Object / Image</td><td>Production data / Audio</td></tr>
<tr><td>Object / Statistical chart</td><td>Production data / Animation</td></tr>
<tr><td>Object / Illustration</td><td>Production data / Video</td></tr>
<tr><td>Object / Table</td><td>Production data / Vector data</td><td rowspan="2">Geographical production resources</td></tr>
<tr><td>Object / Audio</td><td>Production data / Raster data</td></tr>
<tr><td>Object / Animation</td><td>Original data</td><td>Original data</td></tr>
<tr><td>Object / Video</td><td></td><td></td></tr>
</table>

Table 1. SIANE resource *classes* and *types*.

Content manager stores metadata information for each resource according to an own XSD schema, based on ISO 19115 rules. Only usual ANE metadata not covered by the standard have been added and this is what makes the difference from rules. There are tools for import and export to and from the standard (transformation between SIANE schema and ISO 19139 is carried out using XSLT).

In the previous table, resources called as *object* are published resources (on the Internet or printed), while those called *production data* are not visible to a Web user and are the base for generating the *objects*. An example for an object would be a publication resource which has the PDF file in its contents and for production data would be the QuarkXPress file.

In order to make a brief summary of the system main functionalities we will see the cycle life for a map, a text and an image. We will make the map example with data from National Statistics Institute, specifically *University students enrolled by sex* with data grouped by autonomous communities.

2.2.1 Excel macro for preparing data

According to given priorities in SIANE's development, we still haven't data supply agreements which would set deliver conditions such as specific data, format, periodicity, etc. Nowadays, there are almost as many formats as ANE data suppliers. However, is very common these institutions offer the Excel format in their websites as the main one or as an alternative. The problem then is the internal organization of data inside Excel file. In order to resolve this situation, at least temporarily, a simple Excel macro was designed and has two basic functions. To extract data registers that will be represented in map cartographic units, using a combination of Excel texts styles (font type, colour, size, etc.); Excel style structure is not always as suitable as we would desire but this tool let us to extract data in an automatic way in most of cases. And the other function is to correct identifiers for that data registers so that they can be recognized with no error by CMS. For example, in case we have a choropleth map of municipalities, this tool would assign the official name given by Ministry of Public Administration and the correspondent numeric code maintained by National Statistics Institute.

2.2.2 Content Management System (CMS)

Before we can continue with the example, we are going to make a brief description of how resources are organized inside the CMS. In the left side of the application interface we have the main resources browser. A kind of tree in which resources can have more than one occurrence (resource is not duplicated, only the reference to it); in practice, is a tree in a page makeup sense, because every ANE resource could appear in many publications. This is a good point when we think of resource reutilization. In the right side the main component is the content and metadata viewer. There is some other but we will see them below.

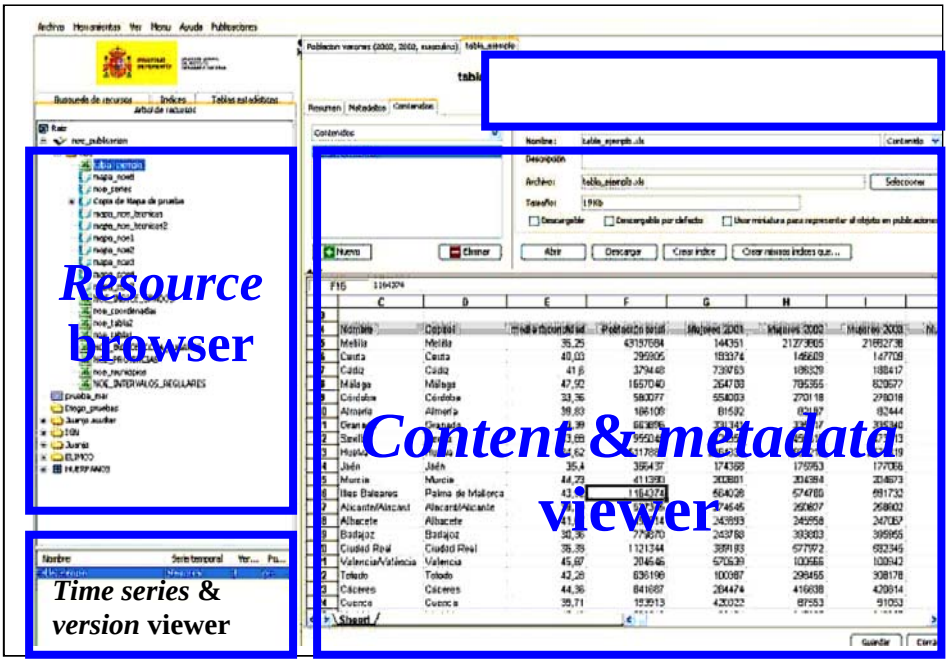

Figure 1. SIANE Content Manager

Each resource in CMS is compound of: contents and documents (any kind of file), metadata (public and private) and relationships. Public metadata is stored in XML format and private are distributed between database fields and XML elements. Content is the main data, such as a map file, a PDF file, a Word file, etc., and documents are considered non-standard metadata; for example data product specifications documents such as text or a cartographic cookbook that must be followed when a text is written or a map is designed. Parent and cross reference ISO 19115 metadata relationships are internally used by CMS, but child relationship is also considered; Parent and child are hierarchical relationships and let us to know, for example, that a publication includes some maps, texts, etc., or the other way around, that a map or an image are included inside a publication. Cross reference relationships let us to create thematic associations among different resources that also will be shown on the Internet. There is an integrated metadata editor in CMS, which helps to make collaborative cataloguing.

In the example above, once we have well formatted Excel data file, we would create into CMS a *Statistical data* resource and would upload previous file. In case we have a text, we would create a *Production data / Text* resource to store working file (for example a Word file), and then we would load the PDF version of that file in an *Object / Text* resource when the content was valid to publish. We would act the same way in case we have an image but using TIF format as the working file and JPG for publishing.

In addition, CMS has a special kind of resource called *variable*, designed to deal with all ANE statistical variables. Any variable stored in CMS has three elements: a list of

locations (Andalucía, Extremadura, etc.), a list of values (145.812 students, 76.362 students, etc.) and metadata. There are two types of locations: identified locations (which are linked to database georreferenced geometries) or point locations with X and Y coordinates. The difference between them is that, in the first situation, user only needs to know about the name or identifier of cartographic units that the variable value is associated with (for example name of municipalities, provinces, countries, etc.), while in the second, coordinates are needed, apart from variable values (and it is only valid for quantitative maps with point symbols). CMS has a wizard tool which let's to extract variables data from Excel sheets, in order to make them usable in SIANE. Note that in the example, we have six variables: number of university male students enrolled in academic year 1999-2000, number of university female students enrolled in academic year 1999-2000 and the same for 2003-2004 and 2007-2008 academic years.

Any variable stored into CMS may be used for generating other variables through simple mathematics operations. In addition, they can be transformed using database relationships; for example, if we have a variable referred to spanish municipalities, we can make some operations like sum, mean, etc., in order to get another variable referred to spanish provinces. In this case, municipalities to provinces relationship (many to one) must exist in database.

Some tools let store different contents and metadata for the same resource when there are data updates (temporal series) or corrections to resources that have been published previously (versions); (see figure 1). In addition, maps are updated automatically when updated data comes into the system, but manual revision is required. A metadata based search tool let us to find resources in a comfortable way. And finally, we have to mention the copy tools that makes easy to generate new resources from other; a really helpful tool, for example, in metadata creation.

2.2.3 Map Editor

Once a variable have been stored, may be used to draw a thematic layer using Map Editor application (it is integrated in CMS). First step will be base map selection. Due to the need to make also printed maps, the base map list not only let us to chose among different geographic areas (Spain, Europe, World, etc.), projections (Lambert conformal conic designed for Spain, Lambert equal area for Europe, Winkel III for World maps, etc.) or scales but also among different layout dimensions. Base map base list is really a list of ArcGIS Server services. To manage a short services list, the application let us enable or disable ArcGIS layers, so, one service may be used in practice for several kinds of base maps (for example, to change from one province map to another which represents autonomous communities, or simply to change drawing styles). Default map base also includes a text layer with additional map information such as title, scale or data source.

Once we have selected the base map we are guided to choose cartographic technique, cartographic units for which we want to draw the map (in the example the list of locations will be autonomous communities), and finally, graphic properties that we want to vary with variable values and the variables.

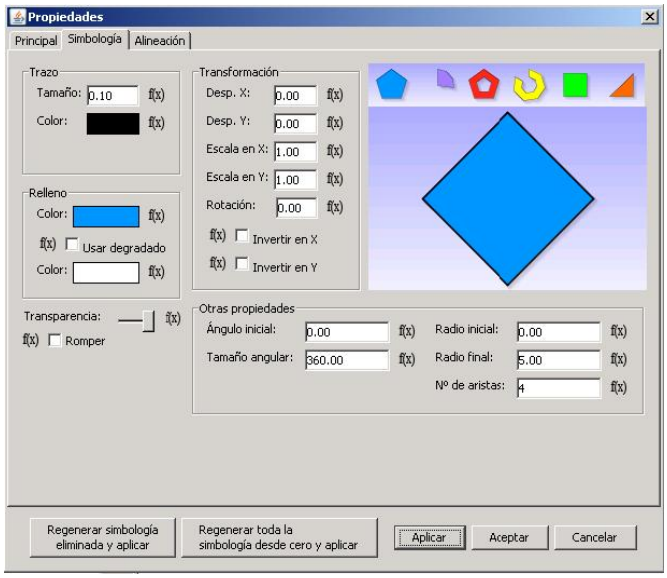

Figure 2. Symbology and variable assignment

After making some symbology adjustments, there are tools to automatically draw the legend, some other to create symbols and texts layers, and export tools to get PDF and SVG formats.

Map editor has been made as a complementary tool for ArcGIS Desktop for quantitative maps (qualitative maps are made in ArcGIS Desktop); however, when designed, we wanted the application was more flexible than ArcGIS in this kind of symbology; for this reason and looking for homogeneity, in the end, we decided that the tool would draw all quantitative symbols. Symbology Java API is an own design. Today is possible to draw proportional symbols, cake charts, line graphs, bar charts and pyramid graphs.

When all necessary metadata is stored, next step would be to publish the resource (map file, JPG image file or PDF file). In that moment, information is replicated to another publication database.

2.2.4 Web application

National Atlas of Spain on the Internet is called *SIANEWEB* and can be found in main Spanish National Atlas website in http://www.ign.es/ign/es/IGN/ane.jsp, inside SIANE project.

It consists of an application in which all ANE contents and metadata are published, making the maps to be the main resource. Navigation is based on, individual resources (each map, image, text, etc.). Access to these resources is organized through ANE thematic structure. An advance search tool, based on metadata, lets that general users gets more accurate list of resources, and IGN users to define new possible publications by searching through keywords (for example a childhood Atlas). Every resource is related with the other thematically or causally (by cross reference relations we saw before). This let us to present a resource network which helps us to make a more coherent, explicative and didactic atlas.

This website is designed to public administrations staff, private professionals, researchers, teachers, students and general users interested in geographic issues or with spanish general knowledge.

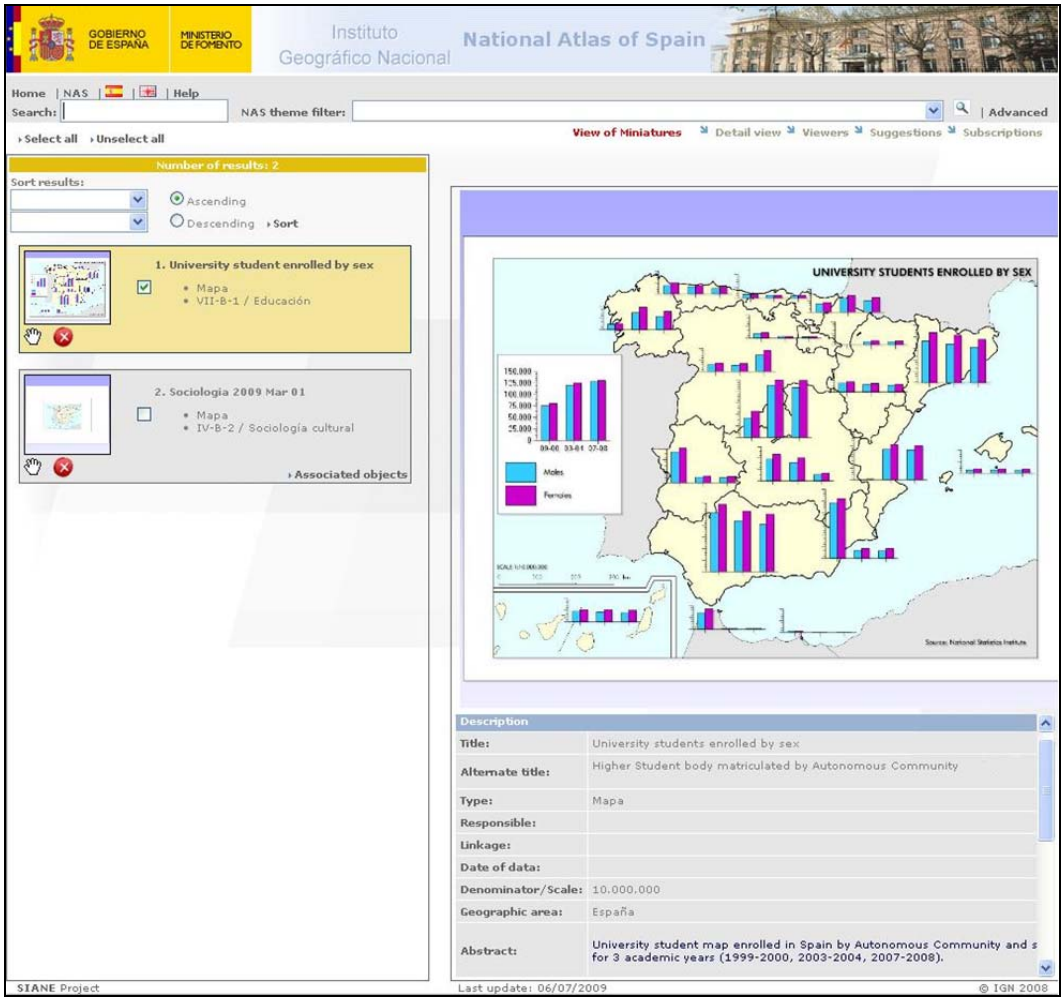

Figure 3. *SIANEWEB*

SIANE is a continuously updated system, which has to be able to manage both printed and digital publications. Up to now, main effort has been dedicated to web publication, however, we have to take into account that this dual publication should lead us to make two kind of maps, and thus, to increment financial resources. As we would like to maintain printed publications, map creation is print oriented but web presented taking advantage of digital tools; for example a dynamic zoom which let us maintain symbols size while map base increases in size.

2.3.- System requirements

Minimum hardware requirements are a PC with, at least, 512 MB of memory, Windows XP OS (only for CMS), an Internet connection (DSL is recommended) and Internet browser (Internet Explorer 6 or Mozilla Firefox 3). Web application can be explored only with the browser, but if we want to obtain full functionality, some plugins are needed: Adobe Flash Player 10, Adobe Acrobat Reader 9 and Java Runtime Environment 5 or 6.

2.4.- SIANE future

This is the first time ANE contents are on the Internet, so, social impact must be measured and analyzed. Up to know, and before this occurs, these are the main lines of future system improvements:

- A tool for massive resource creation and content uploading and downloading.
- Individual statistical chart creation not associated to map cartographic units.
- Undo functionality in Map Editor.
- Download system integration with *National Centre of Geographic Information*.
- Direct colour range selection, in Map Editor, according to the number of legend intervals.
- Increase the number of metadata which is automatically stored on XML files, based on database field contents or user actions.
- User permissions based on resource access (up to now, application functionalities are only considered for permissions).

3. Conclusions

One of the main objectives of SIANE is to make easier and faster the National Atlas content updating. Process subsystem will let to minimize the time period between data is received and resource is published. In addition, ANE content spreading will be improved and costs reduced.

Content language management is also considered, but is a task that has to be dealt with metadata.

Qualitative cartography will be also published using WMS services as a Spanish Data Infrastructure node (IDEE, http://www.idee.es).

ANE metadata will be also published through Directory of Geographic Information (DIGA), IGN documental base which lets to get descriptions, make graphic and alphanumeric queries from all geographic products produced in IGN (http://www.ign.es/ign/es/IGN/diga.jsp).

This project is part of another called *XXI century ANE*, a National Atlas which tries to serve as a relevant contribution to territorial planning, economic and social development, research and educational activities and geographic knowledge of the country.

4. References

- Instituto Geográfico y Estadístico (1912): *Reseña Geográfica y Estadística de España*, Dirección General del Instituto Geográfico y Estadístico, Ministerio de Instrucción Pública y Bellas Artes, Madrid.

- Alonso, J.J.; Del Campo García, A. (2005): "Technical Aspects of The Object Generation Process in the Sistema de Información del Atlas Nacional De España (SIANE)", *Mapping Approaches into a Changing World*, XII International Cartographic Conference (ICC 2005), A Coruña 2005, Cd-rom.

- Aranaz del Río, F. (1993): "El Atlas Nacional de España", *Comunicaciones a la Reunión de la Comisión de Atlas Nacionales y Regionales de la ACI*, Madrid, mayo de 1992, Ministerio de Obras Públicas, Transportes y Medio Ambiente, Madrid, pp. 33-43.

- Del Campo, A.; Romera, C. (2005): "ANEXXI: The XXI Century National Atlas of Spain", *Mapping Approaches into a Changing World*, XII International Cartographic Conference (ICC 2005), A Coruña 2005, Cd-rom.

- Sanz Núñez, A. (1992): "El Atlas Nacional de España", Anales de Geografía de la Universidad Complutense, nº 11, Madrid, pp. 225-230.

- Vázquez Maure, F. (1965): "Informe sobre el Atlas Nacional de España", *III Coloquio de Geografía*, Salamanca, pp. 165-171.

- Vázquez Maure, F. (1968): "La nuevas láminas del Atlas Nacional de España", *Aportación Española al XXI Congreso Geográfico Internacional*, Nueva Delhi, pp. 27-29.