Revista Internacional de Ciencia y Tecnología de la Información Geográfica
International Review of Geographical Information Science and Technology

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

DATA MODEL OF THE ONGOING GEOSCIENTIFIC MAPPING OF THE EASTERN CORDILLERA OF THE DOMINICAN REPUBLIC

FERNANDO PÉREZ CERDÁN[1*], MARÍA JESÚS MANCEBO MANCEBO[1], EUSEBIO LOPERA CABALLERO[2], ALBERTO DÍAZ DE NEIRA[1], JACQUES MONTHEL[3], PEDRO PABLO HERNÁIZ HUERTA[4] AND JESÚS GARCÍA SENZ[1]

[1] Instituto Geológico y Minero de España (IGME), Calera, 1, 28760, Tres Cantos - Madrid
[*] f.perez@igme.es
[2] Instituto Geológico y Minero de España (IGME), Plaza de España, Torre Norte, 41013 – Sevilla
[3] Bureau de Recherché Géologiques et Miniéres, 3 Avenue Claude-Guillermin BP 36009 Orléans
[4] INYPSA, General Díaz Porlier 49, 28001 - Madrid

ABSTRACT

A new data model concerning geological information, mainly geological maps, has been developed to progress towards the better understanding and effective use of geoscientific maps. This model comprises different types of geologic units, sharing geometry, geological structures, survey points, mineral occurrences and sketches and reports. The model is completed with the incorporation of the geoscientific language that controls all the terms used in the model and the domain of the properties established.

The logical model has been implemented in a pilot zone of 10 adjacent sheets to a scale of 1:50,000 in an ESRI® Geodatabase environment which has been used only as a model demonstrator. The validation of the application schema has been carried out applying geo-scientifically based rules by means of certain data query and selection.

Keywords: Geoscientific data model, geologic mapping units, geographical information systems, geoscientific language, Dominican Republic.

MODELO DE DATOS PARA LA ACTUAL CARTOGRAFÍA GEOCIENTÍFICA DE LA CORDILLERA ORIENTAL DE LA REPÚBLICA DOMINICANA

RESUMEN

Con el fin de optimizar la comprensión y explotación de la información geológica espacial (mapas geológicos) en formato digital, se ha diseñado un nuevo modelo de datos que integra diversos tipos de unidades cartográficas, estructuras geológicas, puntos de medida e información

Recibido: 7/2/2012
Aceptada versión definitiva: 28/6/2012

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

documental. El modelo se completa con la incorporación de de un lenguaje geocientífico que controla todos los términos empleados en el modelo y el dominio de ciertos atributos esenciales.

El modelo lógico se ha implementado en una zona piloto que cubre 10 hojas a escala 1:50.000 en una geodatabase de ESRI®. La validación del esquema de aplicación propuesto se ha llevado a cabo aplicando reglas geocientíficas mediante una serie de consultas y selecciones.

Palabras clave: Modelo de datos geocientífico, unidades geológicas, sistemas de información geográfica, lenguaje geocientífico, República Dominicana.

1. Introduction

The need to have a knowledge of terrain from a geo-scientific perspective, as a basic source for studies concerning natural hazards, mineral resources hydrogeological resources, fossil fuels, environmental protection and, in general, knowledge about all the activities that have the physical environment as a support, is a fact that is patent because of its major effects on the development of society. Such is the necessity to have such knowledge that the public administrations established, more than 150 years ago, what are commonly referred to as Geological Services, these being the bodies that are responsible for generating this awareness, which generally takes the form of geological and geothematic mapping.

Since 1812, when the first geological map was published, plotted by William Smith, this type of mapping has become one of the most valued scientific documents by geoscientists. The development of geographical information system (GIS) now enables the scientist to use spatial information in an integrated manner that was previously impossible. GIS allow for a new dimension of analysis in the light of the fact that remote sites are now available that can transmit the information to almost any point that is connected to the network; they open such important scenarios as the "diffusion", "exchange" and "integration" of information. However, this is only possible if the information is structured into data models that are not only consistent with the "real world", but also consistent with related concepts that are unequivocally defined. Data models are essential because they serve as interpreters between the users and the information at a time when access to the latter is, in most cases, remote. With an increasing demand for information and growing diversity of applications, data models are essential for many types of geosciences.

Throughout recent years, major efforts have been made to generate geological mapping in digital format. However, the resulting spatial databases have been greatly affected by the characteristics of the initial information, specifically, the individual sheets.

This means that there are still problems that prevent the geo-scientific information being put to the best use in digital format. The most significant of these problems are as follows: the continuity of the information, a lack of a relationship between related features, a lack of connection between the georeferenced documentation and the lack of a Geoscientific Language that provides a comprehensible and accepted terminology.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

2. Background

The European Union SYSMIN programme - whose objective is to develop the geological and mineral sector in the countries of Asia, the Caribbean and the Pacific (ACP), (Muñoz Tapia, 2006) -, was initiated in 1996. One of the projects that forms part of this programme is the one named "Geological and Geothematic Mapping of the Dominican Republic", which was developed in 2 phases, from 1997 to 2000 and from 2002 to 2004, respectively.

Three complete quadrants were made of the *Cordillera Central* at the first phase (Zone "C"), which produced 12 geological maps to a scale of 1:50,000, 3 geomorphological maps, 3 maps of mineral resources and 3 geotechnical maps to a scale of 1:100,000 (figure 1). These maps were prepared in accordance with the geothematic mapping standards developed by the *Instituto Geológico y Minero de España* (IGME) and the *Dirección General de Minería* in the Dominican Republic.

The second phase was a continuation of the previous one but towards the north-west (Zone "K"), south-east (Zone "L" West) and east (Zone "L" East). A total of 35 geological maps were prepared to a scale of 1:50,000, 11 maps of mineral resources to a scale of 1:100,000, 11 geomorphological maps to a scale of 1:100,000 and 11 maps of active processes to a scale of 1:100,000.

The mapping production of both phases was entirely digital, albeit with different degrees of development. The first phase, Zone "C", produced maps with scales of 1:50,000 and 1:100,000. The mapping units on those maps are only identified by their number, lithological description and symbology. They were mainly used to prepare mapping, and apart from the spatial datasets and the final maps in digital format, when they were plotted, no further intermediate products were obtained. However, a more complete model was defined for projects of Zone "K" and Zone "L", as a result of an analysis of the chronolithostratigraphic legends and the linear and specific geological entities of the maps, whose origins lie in the National Geologic Map of Spain at 1:50,000 scale, (MAGNA), produced by the IGME, (Pérez Cerdán *et al*., 2006). The most important properties of the geologic units that give rise to this model were obtained from an analysis of the chronolithostratigraphic column. Thus, each geologic unit is defined by its identifier, its description, its lithological components, the chronostratigraphy concerned and the symbols to be used in their representation (see figure 2). The chronostratigraphic terms and the lithological components come from basic controlled vocabularies, obtained from the legends in which standardised and extensively used terminology is utilised. The idea of "legend to scientific language" is applied to the legends (Bolduc, 2004).

Apart from mapping units, different geological structures are shown on the maps, such as contacts, faults, fold structures, stratification, schistosity, lineation and foliation.

This model enables the user to make use of the geological information at 1:50,000 scales, using a greater number of selection criteria. However there is a very wide gap between this model

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

and the whole spatial and non spatial data collected in those projects. Survey points and sampling points with all the associated information that is essential for defining the mapping units were not taken into account in the model above. Besides, related mapping units such as tectonic units, lithopermeable units or geomorphological units might improve the digital data use, bringing up new opportunities to exploit the data.

The definition of the units for the diagrams on each sheet – hydrogeological and tectonic – and the geologic map units for the mineral resources maps come from geological information to a scale of 1:50,000.

Furthermore, the structure of the geological information is focused on the 1:50,000 scale sheets, so the management of administrative units or other user defined areas requires a task of complete digital matching, with the added problem of recoding the units. At present, the relationship between one sheet and the adjacent ones is nothing more than a mere geometrical matching of the mapping units and their representation in a homogeneous way.

This analysis, together with the need to improve the management and running of digital information, allowed for a new approach to the way the information was organised, which would make the following possible:

- Guaranteeing the *continuity* of the mapping information beyond the boundaries of the sheets.
- Allocating the largest amount of *properties* possible to the mapping units.
- Reducing the number of *spatial objects*, making it possible to achieve a greater consistency between the related mapping units.
- Associating the *supplementary information* related with the respective geological features.
- Establishing a *geoscientific language* to improve the understanding of the information and the way that it is communicated.

In short, the aim was to supplement the mapping information prepared, with an associated spatial and document database containing continuous and integrated geoscientific information.

At the same time as these mapping projects were being carried out, a series of different initiatives were set in motion for the development of models of geological data mainly in the public sector. The most outstanding ones are the "Digital Geologic Map Data Model, Version 4.3" (Johnson B.R. *et al*., 1999), the variants of which were developed in the Geological Service of Canada under the title of CordLink (Brodaric *et al*., 1999), in the Geological Service of Arizona (Richard and Orr, 2001) and in the Geological Service of Kentucky (Soller *et al*., 2002); and in the North American Map Data Model, NADM Conceptual Model" (NADM Steering Committee, 2004 and Richard, 2005).

Reference must also be made to the work done by the Commission for the Management and Application of GeoScience Information, CGI which is carried out under the auspices of the

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

International Union on Geological Sciences, IUGS, and has developed the GeoScience Mark-up Language, GeoSciML. This is a data format based upon standards that, regardless of applications, provide a framework for encoding of spatial and thematic geoscientific data. Among other works, it is based upon the NADM and is currently defining a Geoscientific Language (Richard & CGI, 2007 and CGI, 2008).

All these initiatives have been analysed and taken into consideration, to a varying degree, in devising the data model that is described below, but some considerations between the NADM and the model proposed have to be done.

The North American Data Model (NADM) has been designed principally for the interchange of the geologic information represented on geologic maps. It shares with the Dominican Republic data model proposed in this paper a core with the essential geologic features: geologic units and geologic structures. However, the amount of properties associated to the geologic features, are greater in the NADM. Undoubtedly it may be considered as an advantage, but in some particular cases, the NADM turns out in a complex model, as it happens in the modelling of earth materials. Thus, model implementation in physical systems and further managing isn't an easy task.

Although the Dominican Republic geological data model presented includes fewer feature properties, it can provide the precise geological responses to the common queries that users address to geological databases. Geologic age, as well as lithology, are clearly established in the model. Furthermore, users are able to retrieve data of the different geologic units defined in a single polygon or part with their specific properties.

But the major difference with regard to the NADM is related with all the features concerning sampling, recording, and analysing field data. Unlike the NADM, the Dominican Republic geological data model comprised not only the geologic features depicted on maps, but also the survey points, samples, geochemical analyses, geochronological datings and thin section studies. Users may access those informations which back up the geological map data.

While the NADM may be considered an excellent data model for geologic maps with a remarkable scientific approach, the model proposed for the Dominican Republic is focused in the whole geological mapping process as a source of information, instead of just the maps.

3. Data model of the ongoing geoscientific mapping of the Eastern Cordillera of the Dominican Republic

3.1 General Considerations

The development of the data model for the Geoscientific Mapping of the Dominican Republic is based upon the concept of integrated mapping, which is defined as a set of spatial, thematic and document information that can be used with digital systems. Real world relations between the different objects that make it up are established implicitly or explicitly, whether they are primitive (for example, a geologic unit) or derived (for example, a report). The integrated mapping is characterised by the fact that it complies with the integration levels that are defined below.

Level I: Continuity of the geographical domain

The limits on the sheets have been dispensed with giving rise to continuous mapping, in which the only limits to the mapping units are their own extension or the general mapping domain (figure 3).

Level II: Visual integration

The information can be displayed jointly, regardless of the systems in which it is stored. This only affects the spatial information. The joint display enables the user to perceive the spatial relations between the objects (figure 4).

Level III: Spatial integration

Spatial relationships are explicitly defined linking features. So no specific tool of spatial analysis is required to find out whether a feature is linked to a mapping unit or not.

For example, any mineral occurrence should be associated with a geologic unit or with a geologic structure. It is not necessary to display them jointly to establish the relations (figure 5).

Level IV: Document integration

Every spatial object is linked with the documentation that is associated with it or derived from it. A survey point where a photograph has been taken is a document that is directly associated with the survey point and it has to be displayed as another attribute of that survey point (figure 6).

Level V: Conceptual integration

Features sharing same geometry even with some degree of aggregation are stored in the same spatial objects. The most immediate consequence is the absence of duplicities, which leads to the benefit that the information is consistent. Any change in the geometry affects all the associated concepts, because there is only one geometry. For example, a formation that is represented by means of a geologic unit shares the geometry with that unit (figure 7).

Level VI: Semantic integration

The concepts, classes and values of the properties, with the exception of the free descriptions, are regulated by a series of vocabularies, dictionaries and lexicons. These serve to

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

make it easier to understand the information, and it is possible to control and automate the contents and simplify the queries (figure 8).

At Levels I, II, III and IV, the integration can be done independently by geothematic disciplines, but Levels V and VI require a major degree of interdisciplinary coordination. Level VI applied to geoscientific information constitutes integrated geoscientific mapping.

3.2 Geographical and thematic scope

The study area is located on the East of La Española island (figure 1) covering an area of 6.000 square kilometres, equivalent to ten 1:50,000 scale maps. The Eastern Cordillera is the most important geographic feature, stretching from west to east. The northwest sector is dominated by Los Haitises, a karstic calcareous platform. On the southwest sector the Caribbean Coastal Plain is extended. Although the Eastern Cordillera is not very high, not exceeding 800 metres above the sea level, his relief is quite rough.

From the geologic point of view, the Eastern Cordillera is mainly composed of volcanic and volcanosedimentary rocks of Los Ranchos Formation (Lower Cretaceous) on the western sector and of a thick sedimentary and volcanic series of Las Guayabas, Río Chavón and Loma de Anglada Formations (Upper Cretaceous), on the eastern sector. Between both groups of Formations, the Hatillo Formation (Albiense), a reef limestone, is present but with some gaps. The Cretaceous rocks have been intruded by tonalitic bodies. On the southern edge, Paleogene sedimentary materials appear but without uniformity.

On the northwest sector, as well as some southern spots, (Caribbean Coastal Plain) Plio-Quaternary limestones and marls (Yanigua and Los Haitises Formations) overlay the older rocks. Geologic materials are completed with a sort of Quaternary sediments, mainly related with fluvial, karstic, marine and coastal processes.

The scope for the model that is dealt with in this work comprises all the geoscientific information compiled and prepared to make up the maps listed in table 1. Besides, the stratigraphic columns (Ardévol, 2004a), the sedimentological report that described the lithostratigraphic units (Ardévol, 2004b), the results of the chemical analyses (Escuder Viruete, 2004), the Macroforaminiferal Report – catalogue (Serra Kiel, 2004) and the dating work done (Ullrich, 2004a and Ullrich, 2004b) have also been incorporated.

On the one hand, this means incorporating the supplementary information consisting of samples, detailed columns, photographs and reports; and on the other hand, it includes the properties that are associated with the mapping units that are only represented on the diagrams. Any graphic objects whose purpose is to make up the mapping products with an institutional formality, such as scales, logos, and credits, are excluded from the model.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

The intersection between these abstract components of the model and the domain or scope of specific knowledge about it that are applied, enables the user to establish the subset of thematic components of geoscientific concepts with which to construct the logical model or conceptual schema in accordance with the International Organization of Standardization, (ISO) definition (ISO, 2008).

3.3 Conceptual domain

The data model for the geoscientific information about the Dominican Republic is shown as an abstract representation of the domain concepts that are the subject of the study, their relationships and their properties, establishing a consistent framework for designing and developing the spatial database, complying with all the requirements with respect to the necessary clarity and concision.

The representation of the geoscientific knowledge has been studied in detail by S.M. Richard, who establishes a framework of reference that comprises the Real World and Information Systems consisting of a series of agents and abstractions of different types (Richard, 2006). A more general framework of reference or abstract model is established in this work, in which a definition is given of the main reference components and that, together with the delimitation of the domain, enable the user to subsequently develop the conceptual model, the conceptual schema and the application schema.

3.4 General components of the model

The components that constitute the model are: observations, properties, geoscientific concepts, relations and semantics. There are different kinds of links between them and their final expression in the three abstraction models – conceptual model, conceptual schema and application schema – is variable, because it depends upon the nature of each one of them.

- *Observation*: This is the perception of a phenomenon. Observations are essential for defining and classifying the concepts and for identifying and delimiting their manifestations.

- *Property*: The permanent quality of an object. Properties are essential for making generalisations and for classifying concepts.

- *Geoscientific concept*: This is the abstract representation of a phenomenon or a set of phenomena observed in the physical environment and characterised by a series of common properties. The concepts are established by means of a twofold process: abstraction, to mentally separate one quality from another, and generalisation, applying to new abstractions, objects or phenomena, the common characteristics abstracted beforehand (Blázquez, 1997).

- *Relation*: A situation brought about between two objects when there is some circumstance that unites them. In the specific case of the Data Model for the Dominican Republic and following the theory applied in the data models designed at the Geological Service of Arizona, a distinction has been made between 3 types of relations (Richard and Orr, 2001): simple, hierarchical and complex.

- *Semantic*: The science of the meaning of signs and the relationships between the signs in a language and the objects to which they refer. Semantics, through geoscientific languages, enables the user to establish the domain of concepts and properties, and is essential for the diffusion and comprehension of information (SLLT, 2003; Richard, Matti and Soller, 2003).

3.5 Conceptual model

The geoscientific concepts defined are the following: Geologic Unit, Materials (Lithological Components), Geologic Structure, Geologic Process (Weathering and Metamorphism), Paleontological Trace and Survey Point. These concepts are distinguished, classified and ordered on the basis of their properties (geoscientific), which can be defined as each one of the permanent qualities of a geoscientific nature that identify and characterise the geologic concepts. The properties enable one to classify the concepts, not only by their presence, but also by the values that they take on. The values associated with the properties can be either objective or subjective, depending on whether or not there are reference patterns and measurement techniques, and on the basis of the evaluation method. There are cases where the value of a property can either be established analytically or by means of a visual estimation.

The concepts are subdivided into classes, defined as sets of objects that possess at least one common characteristic by which they can be identified as belonging to one single class. The term "class", as is the case with the term "concept", is an abstract term and it has no exact spatial representation. A Geologic Unit, for example, is a concept whereas a Lithostratigraphic Unit is a class of this concept.

Classes are split into occurrences. An occurrence is the particularisation and materialisation or manifestation of a class. The properties of occurrences that belong to one and the same class are common, but the different values of these properties are what make one occurrence different from others. Occurrences have a specific spatial domain. One occurrence in the Lithostratigraphic Unit class is, for example, the Los Ranchos Formation.

Finally, occurrences are composed of one or more parts or fractions. Parts are each one of the independent spatial objects that constitute an occurrence, (see figure 9). The term independent can have different meanings, depending on the physical model that has been adopted.

- *Geologic Unit (GU)*: A basic concept in the model, a body composed of mineral matter of natural origins, either consolidated or unconsolidated, which is differentiated from the ones surrounding it by its lithological and geochronological characteristics, or by any other property, and which is represented on a digital map to a specific scale or with a certain resolution (Adapted from the North American Code of Stratigraphy, 2003, Article 22).

- *Geologic Structure*: General layout, relative position and internal order of the rock masses in a zone or region. This is the sum of all the structural features of an area as a result of the processes of erosion, deposition, lithification, metamorphism and deformation to which they have been subjected.

Certain classes of Structures – Contacts and Faults – limit or intersect the Geologic Units giving rise to the Singular Mapping Surfaces or parts. The Fault class is the one that is most complex in the organisation as a result of its typology and its relationships with the Geologic Units, (see figure 10). Each one of the faults mapped represents one occurrence, and the occurrences are subdivided into types. Each type is determined by the displacement of the blocks and by the fault plane angle. The certainty or visibility of each type gives rise to different subtypes: seen, inferred and buried. Subtype is a term that appears in the dictionary of Structures and it is also characterised by a singular representation. The tectonics or subsequent depositing processes make faults segmented at their point of intersection with the topographical surface. Each one of the segments is one part or fraction of the occurrence concerned. Each and every one of the parts or fractions of a fault has a trace in the spatial data concerned. The occurrences are not necessarily composed of one single type.

- *Materials (Lithological Components)*: all substances of natural origins, regardless of their size or position (NADM, 2004). Land material is defined on the basis of its physical and chemical properties. In the proposed model, only solid minerals with or without crystalline structure are considered to be materials.

In the model, the materials will be considered from a lithological perspective, in such a way that the Geologic Units will be made up of one or more materials, which will be of monomineral or polymineral lithological types.

- *Geologic Process*: a natural phenomenon that acts upon a geological entity and that yields another entity as a result (NADM, 2004). Processes of alteration and metamorphism are the only such processes that fall within the scope of this model.

- *Paleontological Trace*: this is defined as any remnant, trace or remains of a plant or an animal that has been preserved by natural processes in the Earth's crust from an earlier geological era (Neuendorf *et al*., 2005).

- *Survey Point*: any appreciation that is more or less subjective, measured *in situ* or in a laboratory or an image duly located in space whose purpose is to locate, define, characterise, parameterise or record other geological concepts.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

Geological concepts are quantified or qualified by means of survey points, but on certain occasions the observations that are made at a survey point become a geological occurrence or a concept, such as the mineral occurrences.

3.6 Application schema

The specific classes derived from the geoscientific concepts are expressed in the application schema of the geological mapping of the Dominican Republic (figure 11), as well as the relationships that exist between them, their properties and the semantic domain that has been defined to date. The methodology used for the design follows that proposed in ISO Standard 19103, "Conceptual Schema Language" (ISO, 2005).

As can be seen, the application schema revolves around the Geologic Map Unit Singular Mapping Surface (part) binomial, all the other elements in the model being related to them either directly or indirectly. A description is given below of the basic classes in the Schema.

3.6.1 Geologic units

The Geologic Map Unit (GMU) is each of the different units defined in the geological mapping to a scale of 1:50,000 and that are represented in the chronolithostratigraphic legend as an independent item. They are composed of materials of mineral origin differentiated mainly by their lithological and geochronological properties. They are the main entities in the Schema, in fact most of the queries that are made in the geological mapping databases are enquiries about their lithological properties. The majority of the questions that are asked for the validation of the NADM involve searching for objects by specifying their lithological properties (Weisenfluh, 2001).

The attributes that are defined for each Geologic Map Unit are the Description, the Unit Type (Lithostratigraphic Unit or Superficial Deposit), the Geological Age and the Representation.

Every Geologic Map Unit is composed of one or more Lithological Components. These come from a vocabulary that is generated from a hierarchical classification into 5 categories drawn up for the lithological maps of the Spanish Environmental Mapping National Plan to a scale of 1:50,000.

Two additional attributes have been included in the relationship that has been established between the GMUs and the Lithological Components: Relationship and Proportion. The former shows the way in which each component is presented within the Unit, whereas the latter indicates the proportion. In both cases, the terms that may be used are controlled by vocabularies.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

The Lithostratigraphic Units, Tectonic Units, Lithopermeable Units, synthetic Geologic Units to a scale of 1:100,000 and the Surface Formations derive from the Geologic Map Units to a scale of 1:50,000, a product of the different properties that are associated with the latter.

Every Geologic Map Unit to a scale of 1:50,000 invariably forms part of a Tectonic Unit, of a Lithopermeable Unit and of a synthetic Geologic Unit to a scale of 1:100,000. However, where the Lithostratigraphic Units and Superficial Deposits are concerned, the situation is different, because at the present time, it forms part of one or the other, but they never form part of both of them at the same time.

The Geologic Unit Part (GUP) constitutes each of the Parts into which the Geological Map Units to a scale of 1:50,000 are subdivided, and they are defined as each one of the close and individualised surfaces that are limited by one or more Geologic Structures of the Contact or fault type, or by a boundary of a conventional nature. Every GUP is an independent object in the spatial database and they are directly related to the Geologic Map Units to a scale of 1:50,000 from which they inherit all their properties. A Geologic Map Unit is composed of one or more Geologic Unit Parts.

Even if the Geologic Unit Parts belong to the same Geologic Map Unit, they can be distinguished when a particular singularity makes them different through the Singularity property, where certain features that are inherent to those GMUs can be specified.

The Lithostratigraphic Unit (LU) has been defined in accordance with the North American Stratigraphic Code: a body composed of mineral material of natural origin – sedimentary, extrusive igneous, metasedimentary or volcanic – that is distinguished and delimited by its lithological characteristics and its stratigraphic position (NACSN, 2003). This definition excludes intrusive igneous rocks that have their own units, referred to as lithodemics, and their inclusion has been delayed until the model is revised and updated in the future.

Following the example set by other Geological Services, such as the United States Geological Service (USGS, 2009), The British Geological Survey (BGS, 2009), the Australian Geoscientific Agency (GA, 2009) and the Geological Survey of Canada (GSC, 2009), a lexicon has been established for the Lithostratigraphic Units. This lexicon stores the information that is required to identify and describe these Units, including bibliographical references.

Lithostratigraphic Units have recursive relationships, because they are organised hierarchically in a parent-child relationship (Member-Formation).

The Tectonic Unit, the Lithopermeable Unit and the Synthetic Geologic Unit have been established as further Geologic Units.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

3.6.2 Geologic processes

The Geologic processes that are considered are Alteration, such as any change in the mineralogical composition of a rock due to the action of hydrothermal fluids or weathering processes, (derived from Neuendorf *et al*., 2005) and Metamorphism, any mineral, chemical and structural adjustment of the rocks to physical and chemical conditions that are different from the original ones (Neuendorf *et al*., 2005). They have been implemented in the application schema through "zones affected by", so two spatial geologic features have been defined: Alteration Area and Metamorphic Area.

3.6.3 Geologic structures

Contacts, Faults, Folding Structures, Stratification, Schistosity, Lineation, Foliation and Layer Traces have all been included as part of the Geological Structures concept. All of them have been directly linked with the Geologic Unit Parts.

Stratification, Schistosity, Lineation and Foliation come from the structural measurements taken in situ or from photo-interpretation (only in the case of Stratification).

3.6.4 Survey Points

The third basic group is the Survey Points, defined in a general way as points on the Earth's surface, of known spatial location, about which an observation is made, a fragment of rock is picked up or a photograph is taken.

Survey Points are related spatially to the Geologic Unit Parts, because they show the properties of these at specific points that cannot be extrapolated to the Geologic Units in a general way. They are described by means of XY coordinates and an identification code, which relates them with the types of observations that have been made about each one of them: Samples, Photographs and Mineral Occurrences.

The Sample is a fragment of rock that is used to establish parameters or properties of the environment from which it was taken or the Geologic Unit to which it belongs. A sample's expression is variable, depending on the type of study or determination that is conducted, they can be tables resulting from chemical analyses, reports or graphs. Photographs, are invariably associated with the Survey Point from which they have been taken, but if they happen to be panoramic photographs, it is necessary to establish the area that they take in. Mineral Occurrences come from data that have been compiled during the preparation for the mapping of mineral resources. The types of substances detected in the occurrences and their industrial interest are classified into two groups, depending on their type: Metallic Minerals and Fossil Fuels, or Industrial and Ornamental Minerals and Rocks.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

3.6.5 Stratigraphic Columns

Stratigraphic Columns are a special type of observation that is linear in nature, in which a specific sequence of rocks is described graphically and synthetically in situ and shown in the form of a column. Each successive unit of rock is assigned suitable geological symbols – lithology, sedimentary structures and fossil content – and its thickness is shown to scale (Fregenal *et al*., 2000).

The columns have a dual link, on the one hand they have a spatial relationship that links them to a Geological Unit Part that is determined by the location (starting point) that are shown on the column graph, on the other hand they have a thematic link, which associates them with the Lithostratigraphic Units that are identified in them.

Stratigraphic columns are associated with two types of documents: a graph with the column and a diagram of the situation, and a report in which the units are described and its sedimentary environment is interpreted. Both of them have been included in the model as entities. However, the relationship between the columns and the documents is not a simple one, because even though there is only one graph for each column, which could contain more than one lithostratigraphic unit, as has already been pointed out, each report is focused on the description of one single Lithostratigraphic Unit that includes the information that comes from all the columns in which the unit has been described.

Finally, it must be pointed out that semantics links all the elements in the model, because it takes into account everything ranging from geoscientific concepts ("high level") by giving an exact definition of the concept, to the values of the properties ("low level") through specific lists of potential terms. As a result, semantics, concepts and relationships would constitute an ontology, in view of the fact that through them they would explicitly specify the geoscientific domain terms and their relationships (Gruber, 1993). The development of ontologies constitutes one of the pillars of the interoperability of digital information (Ludäscher *et al*., 2003), and makes it easier to understand the information (Brodaric and Gahegan, 2006).

It is essential to have clear and concise definitions and descriptions in order to facilitate the understanding of the meaning of a term and to associate the same concepts or objects with a specific meaning and prevent ambiguities (It could be A or B), vagueness (It is A, but it is not known whether it is A_1, A_2 or A_3) and redundancies (It is A_1 and it is B_2).

In the description of the application schema, and on the basis of the characteristics of each class, occurrence or property, the degree of precision required and the acceptance of the terms, three components of the geoscientific language, of increasing complexity, have been used to form the semantics:

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

- Vocabulary: a set of terms concerning a concept, class, occurrence or property. They do not have a definition, because the meaning of the term is universally accepted. They are a simplification of the dictionaries.
- Dictionaries: a set of terms with their definitions with respect to concept, class, occurrence or property.
- Lexicons: a set of terms with their definitions and certain number of descriptors and properties concerning a concept or a class.

Vocabularies and dictionaries may include numerical or alphanumerical codes that serve as keys, whose basic function is to facilitate the relationships, to improve the consistency of the databases or to establish hierarchies for the terms.

3.7 Implementation Schema

The implementation schema is the result of transferring the application schema to specific software. It is also expressed through a formal graph that shows all the components that are a product of the implementation, with their nomenclature and final relationships, figure 12.

The implementation schema is closely linked to a specific technology, for an application schema there is an indeterminate number of implementation schemas. Furthermore, there can be different implementation schemas for one single application schema on the same technology. The data model and the topological functionalities that support the selected technology play a decisive role in this process.

In this particular work the implementation schema has been designed for the ESRI® Geodatabase environment. The aim has been to use this technology for testing issues. Besides the ESRI Geodatabase is used extensively in the management of geoscientific information. It should be considered, for the scope of this paper, only as a model demonstrator.

The main characteristic of the implementation schema proposed is a reduction in the number of spatial objects and feature classes, which prevents the duplication of common geometry and guarantees consistency where modifications are concerned. The geologic units represent the best example of this reduction, because all the units that are defined in the implementation schema share the same geometry, which is associated directly with the Geologic Unit Parts.

Several circumstances have given rise to an increase in the complexity of the implementation schema in relation to the application schema, the main ones being listed below:

- The properties of the classes have been allocated to different tables.
- Some classes, with a complex geometry, have been made up by more than one feature class.
- The recursive relationships between categories of one single class have been implemented through the use of intermediate tables where such relationships are dealt with.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

- The properties whose scope is specified in controlled terms have been implemented as vocabularies or dictionaries.
- The property that is referred to as "representation", has given rise to independent tables of symbols.
- The limitations to the application selected have occasionally made it necessary to double the number of geometry and tables.

The topology between the geological units and the limits that separate them – contacts and faults –, is one of the basic aspects that has been successfully dealt with in the implementation schema. Each one of the parts of the contacts and faults is associated with the Geologic Unit Parts that it separates, as a result of which they also inherit all the types of units on which they border. It is thus possible to make a selective representation of contacts on the basis of the type of unit to be plotted.

The geoscientific language materialises in the model through a series of tables that either link up with all the properties to which reference is made, when they are the ones that specify the domain of the latter, or define the properties in themselves and the concepts.

3.7.1 Geologic Units

Only one feature class is required to describe the geometry of the geologic units, the GeologicUnitPart feature class. It is linked, by one side, with the description table of the geologic units at 1:50,000 scale, GeologicUnits50 table, with the GeomorphologicObject table and the Facies table. By the other side, geologic unit parts are linked with all the geologic features, geologic structures, surveys points, mineral occurrences, that is, all features related to an specific point or polygon instead of the complete geologic unit. Geologic units at 1:50,000 scale have two groups of relationships: one group includes specific properties like age or lithology; the other one is related with the geothematic units. The link between the geologic units at 1:50,000 scale and those geothematic units is materialized by the GeothematicUnits table, a code table. This code table is the starting point to a set of relationships and tables ending in the symbology tables. Lithostratigraphic units have their own code table, LithoStratigraphicUnits table, due to the two types of units considered: Formations and Members. Both lithostratigraphic units are linked with the stratigraphic columns by an intermediate table, ColumsLithoStratigraphicUnits, where many to many relationships are resolved.

3.7.2 Landforms and surfacial deposits

As long as a geologic unit part may be related either with a landform or with a surfacial deposit or with both, a specific artefact named gemorphologic object has been created. The GeomorphologicObjets table is also an intermediate code table and allows the link between one geologic unit part with different geomorphologic features. Furthermore, the extent of a geomorphologic feature may cover more than one geologic unit part. It's important to note that the

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

feature classs AreaLandFormParts and LineLandFormParts don't include the geometry of the landforms, but some inner specific graphic elements to represent those landforms. The boundaries of the landforms and surfacial deposits remain in the GeologicUnitPart feature class.

3.7.3 Geologic Processes

A geologic process may affect either a whole geologic unit or just an area of a geologic unit part, and many to many relationships are also possible. An intermediate feature class, named ProcessParts, has been created by the intersection between the geologic unit parts and the geologic processes. Thus, the exact area affected by a geologic process of each geologic unit part is known.

3.7.4 Geologic Structures

The geometry of each type of geologic structure is represented in a single feature class, except for faults. Contacts and faults are essential features for error checking and data validation. They are split in parts as they intersect between them, so that the boundaries of each geologic unit part are precisely defined. This is the best option to guarantee the topological relationships between the geologic units at 1:50,000 scale and their boundaries. A double link has been established between the geologic unit parts and the feature class of contacts and faults, ContactsFaults. The Geodatabase limitation about joins has forced to create two tables, one for each side, with the codes of the geologic units. Besides, this twin relationship provides a better performance in selecting and displaying the boundaries of the geothematic units. When the geothematic units are displayed, the inner boundaries between parts of the same unit must be hidden, that is to say, if the unit code of the left side is the same as the code of the right side, the contact mustn't be plotted. Although faults have been broken up in parts, all of the parts that make up a single feature share a common fault code, FK_Fault attribute in ContactFaults feature class. In addition, a fault feature class, Faults, without breaking points, has been generated with the aim to improve the portrayal of the faults.

3.7.5 Survey points

Survey points geometry is stored in a single feature class, SurveyPoints. It is the link between the geologic unit parts, and samples and photographs. Each sample type has its specific table. An additional relationship has been created in order to link the mineral occurrences with the survey points.

3.7.6 Stratigraphic columns

Related with the other features through the lithostratigraphic units the stratigraphic columns geometry and attributes are included in a feature class named Columns. The link with the lithostratigraphic units has been made by means of the ColumnLithostratigraphicUnit table, which

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

manages the many to many relationship between the stratigraphic columns and the lithostratigraphic units.

3.7.7 Mineral Ocurrences and Metallotects

There are two types of mineral occurrences, metallic and industrial rocks, but both are grouped in a single feature class, MineralOccurrences. Metallotects are linked on one hand with the mineral occurrences and with the geologic unit parts on the other hand. Metallotects have been split in parts, as it has been done with the geological processes, explained earlier. However, a special feature class, MetallotecPortrayal, has been created due to the metallotect representation in maps, where the actual limit of each metallotect is replaced by a smooth enveloping line.

3.8 Data model validation

Four types of processes are used to validate the application schema and, as a consequence, the conceptual schema:

- Geoscientific consistency rules.
- Samples concerning feature query and selection.
- Making sure that the degrees of integration are fulfilled.
- Accessing the information through the use of terms that are controlled in natural technical language.

3.8.1 Geoscientific consistency rules

The relationships that exist between classes and the geoscientific domain occurrences are no random. They are often regulated by a series of principles that are based on the nature of the classes themselves and that are of a general nature, regardless of the spatial scope of application. They are translated into a set of rules and restrictions whose verification through the implementation schema must be possible, which means that the application schema and the model in general, are validated from a geoscientific perspective. Controlling this rules and restrictions also amounts to a major breakthrough in the validation of the digital information.

The following rules, among others, have been used to validate the model:

- A sedimentary contact invariably separates two different geologic map units.
- A schistosity measurement cannot be found in a Quaternary geologic map unit.
- There can be no entities belonging to the layer traces class on geologic map units defined as plutonic rocks.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

3.8.2 Testing samples

Testing samples might be considered as user cases, defined as a set of interactions between the user and a system aimed at obtaining a specific result (Booch *et al*., 1999; INSPIRE, 2008).

The samples used for the validation of the model consist of a series of queries and selections of different complexities but that can be considered as being essential in the analysis and operation of a geoscientific information system. They account for a small subset of all the potential user needs of data querying and selection.

figure 13 shows the results of the sample "Selection of Geological Units Parts and Geologic Map Units by content in oxides". The selected values were Si02 < 53% and Fe2O3 > 9.6%, which allow for the partial identification of tholeitic basalts.

3.8.3 Verification of the integration levels

The verification of the integration levels defined was the third process carried out to validate the model.

Level I: Continuous map without limits between sheets (figure 14).

A continuous map is generated upon 10 sheets, with no boundaries or limits between adjacent sheets. A new legend with 85 units was created from the individual legends (205 units in total) that are on the 10 sheets. Later the spatial and conceptual matching were done.

Level II: Superimposing different classes (figure 15).

All the information is georeferenced, so the simultaneous display is immediate.

Georeferenced data points, originally delivered on paper documents or databases, like sample points or geochemical analysis, have been converted into spatial objects either combining digitization and Optical Character Recognition (OCR) or using ESRI specific tools to transform X/Y data into feature classes.

Level III: Relationships between the different classes (figure 16).

The occurrences in the classes that have spatial links have been listed explicitly to make it easy to refer to the information. For example, every Geologic Unit Part is associated with the survey points, mineral occurrences and structural measurements that it includes. This relationship is clearly two-directional.

Level IV: Document integration (figure 17).

All the alphanumerical information or information that can be georeferenced has been associated either directly with its position in the space, as is the case with the lithostratigraphic columns, or indirectly through the class that gives rise to it, as is the case with the photographs, which have been linked to the survey points.

The path of the picture files, for example, is included as an attribute in the table associated with the photograph. This attribute acts as a hyperlink between the point where the photograph has been taken and its image.

Level V: Integration of classes (figure 18).

The resulting classes that have the same spatial definition share the original geometry. One geometry can be used to represent different classes with the symbols that are inherent to the occurrences that make up each one of the classes.

Level VI: Semantic Integration (figure 19).

The information can be queried and selected using controlled terms that are expressed in the natural technical language that constitute the dictionaries and vocabularies.

4. Conclusions

All the concepts that form part of different geologic and geothematic maps, including their relationships and properties, have been integrated into a spatial database derived from the proposed model. The model, as a whole, is a solid data model, because it is based upon criteria derived from geoscientific information, that is to say, geologic, structural, spatial and temporal relationships.

A geoscientific language has been established that facilitates data input, because the domain of its properties is duly defined. An understanding of all the components of the model is also facilitated, because they are accurately defined in the language, which makes it easy to exchange and operate them.

Maximum spatial consistency has been achieved. There are no redundancies in the geometry of the geologic units, so it is possible to obtain different views on the basis of a limited set of spatial objects.

The relationships that are established between all the entities enable the user to make complex selections and queries, tackling a basic aspect such as the geoscientific consistency of the information, which makes it easy to purify the information by applying a series of rules arising from the geologic scope of the zone covered by the model.

By integrating the sampling points and stratigraphic columns and their associated documentation, mineral occurrences, geologic structures and geologic units, it has been possible to create a system that provides data, information and a knowledge base that enable the user to operate the geoscientific information in a better way.

Finally, it must be pointed out that the model proposed is one that is open to the incorporation of new classes and even concepts, as a result of the enlargement of the conceptual domain or the spatial scope of application.

References

Arctur, D. and Zeiler, M. (2004): *Designing Geodatabases*. California (USA), ESRI Press, 393 pp.

Ardévol, L. (2004): *Geologic Map of the Dominican Republic to a scale of 1:50,000. Stratigraphic Columns*. (*)

Ardévol, L. (2004): *Geologic Map of the Dominican Republic to a scale of 1:50,000. Sedimentological Report* (*).

BGS, British Geological Survey (2009): *Lexicon of Named Rock Units*. Keyworth, Nottingham (UK), [Access 09-01-2009]. Available at: www.bgs.a.uk/Lexicon/home/html

Blázquez, F. (1997): *Glosario de las Ciencias Humanas*. Estella, Navarra (Spain), Verbo Divino.

Bolduc, A. M. *et al*. (2004): "Science Language, Parsing and Querying: The Surfacial Side of Things", in Soller, D. (Ed.): *Digital mapping Techniques 2004*, USGS Open-file Report 2004, 1451, Reston, Virginia (USA), pp. 74 - 84.

Booch, G. *et al*. (1999): *The Unified Language of Modelling*. Addison Wesley, Iberoamericana, Madrid (Spain).

Brodaric, B. *et al*. (1999): *Cordlink Digital Library. Geologic Map Data Model, Version 5.2*. National Resources of Canada, Ottawa, Ontario (Canada). [Access: 30-05-2012]. Available at:www.ndam-geo.org/dmdt/CORDLink_Variant_Description.html

Brodaric, B. and Gahegan, M. (2006): "Representing Geoscientific Knowledge in cyberinfrastructure: Some Challenges, approaches and implementations", in Sinha, A. K. (Ed.): *Geoinformatics Data to Knowledge: Geological Society of America.* Special Paper 397, Boulder, Colorado (USA), Geological Society of America, pp. 1- 20.

CGI (2008): *GeoSciML-Data_CookBook_V2.pdf*. Reston Virginia (USA). Commission for the Management and Application of Geoscience Information. [Access: 30-05-2012]. Available at: www.geosciml.org/geosciml/2.0/cookbook/

Díaz de Neira, A. (2004a): *Geomorphologic Map of the Sheet to a scale of 1:50,000 Number 6372-I (Miches). Sheet and Report* (*).

Díaz de Neira, A. (2004b): *Geomorphologic Map of the Sheet to a scale of 1:100,000 Number 6272 (Monte Plata). Sheet and Report* (*).

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

Díaz de Neira, A. (2004c): *Geomorphologic Map of the Sheet to a scale of 1:100,000 Number 6372 (El Seibo). Sheet and Report* (*).

Díaz de Neira. A. (2004d): *Geomorphologic Map of the Sheet to a scale of 1:100,000 Number 6472 (Las Lisas). Sheet and Report* (*).

Escuder Viruete, J. (2004): *Petrology and Geochemistry of Igneous and Metamorphic Rocks in Project L* (*).

Fregenal, M. *et al*. (Coordinators) (2000): *Earth Sciences Dictionary*. Madrid (Spain), Oxford-Complutense, ISBN 84-89784-77-9.

GA (2009): *Australian Stratigraphic Units Database*. Geoscience Australia, Canberra (Australia). [Access: 09-06-2009]. Available at: http//www.ga.gov.au/oracle/stratname/index.html

García Senz, J. (2004): *Geologic Map of the Sheet to a scale of 1:50,000 Number 6372-III (Hato Mayor del Rey). Sheet and Report* (*).

García Senz, J. (2004): *Geologic Map of the Sheet to a scale of 1:50,000 Number 6472-III (Rincón Chavón). Sheet and Report* (*).

Gruber T. R. (1993): "A Translation Approach to Portable Ontology Specifications", *Knowledge Acquisition*, 5, 3, pp. 199-220.

GCS (2009): *Lexicon of Canadian Geological Names*. Geological Survey of Canada (Natural Resources of Canada), Ottawa, Ontario (Canada). [Access 09-06-2009]. Available at: cgkn1.cgkn.net/weblex/weblex_e.pl.

Hernáiz Huerta, P. P. (2004): *Geological Map of the Sheet to a scale of 1:50,000 Number 6272-III (Monte Plata). Sheet and Report* (*).

Hernáiz Huerta, P. P. and Díaz de Neira, A. (2004): *Geological Map of the Sheet to a scale of 1:50,000 Number 6272-I (Antón Sánchez). Sheet and Report* (*).

IGME (1997): *Computing Procedure for the National Thematic - Environmental Mapping Plan*. Unpublished Document. Instituto Geológico y Minero de España, Madrid (Spain).

INSPIRE (2008): Methodology for the Development of Data Specifications. Joint Research Center, Ispra (Italy). [Access: 30-05-2012]. Available at: inspire/jrc.ec.europe.eu.

ISO (2002): *Geographic Information-Reference Model. ISO 19101:2002(E)*. International Organization for Standarization, Geneve (Switzerland).

ISO (2005): *Geographic Information-Conceptual Schema Language. ISO/TS 19103:2005(E)*. International Organization for Standarization, Geneve (Switzerland).

ISO (2008): *Geographic Information-Terminology. ISO/TS 19104:2008(E)*. International Organization for Standarization, Geneve (Switzerland).

Johnson B. R. *et al*. (1999): *Digital Geologic Map Data Model, Version 4.3*. United States Geolgical Survey, Reston, Virginia (USA). [Access: 09-06-2009]. Available at: geology.usgs.gov/dm/.

Lopera, E., Locutora, J. and Florido, P. (2004): *Map of Mineral Resources to a scale of 1:100,000 Number 6272 (Monte Plata). Sheet and Report* (*).

Lopera, E., Locutora, J. and Florido, P. (2004): *Map of Mineral Resources to a scale of 1:100,000 Number 6372 - 6472 (El Seibo-Las Lisas). Sheet and Report* (*).

Ludsächer, B. *et al*. (2003): "Toward a Cyberinfrastructure for the geosciences. A Prototype for Geologic Map Integration via Domain Ontologies", *Digital Mapping Techniques 2003*, USGS Open-File Report 2003-167. Reston, Virginia (USA), pp. 223 - 229.

Monthel, J. (2004a): *Geological Map of the Sheet to a scale of 1:50,000 Number 6372-II (El Seibo). Sheet and Report* (*).

Monthel, J. (2004b): *Geological Map of the Sheet to a scale of 1:50,000 Number 6372-IV (El Valle). Sheet and Report* (*).

Monthel, J. (2004c): *Geological Map of the Sheet to a scale of 1:50,000 Number 6472-IV (Las Lisas). Sheet and Report* (*).

Monthel, J. and Capdeville, J. P. (2004): *Geological Map of the Sheet to a scale of 1:50,000 Number 6272-II (Bayaguana). Sheet and Report* (*).

Monthel, J., Nicol, N., Fondeur, L. and Genna, A. (2004): *Geological Map of the Sheet to a scale of 1:50,000 Number 6272-IV (Sabana Grande de Boyá). Sheet and Report* (*).

Muñoz Tapia, S. (2006): *Progress Report on the Geothematic Studies in the Dominican Republic*. Dirección Nacional de Minería, Santo Domingo (República Dominicana). [Access: 09-06-2009]. Available at: dgm.gov.do/sdgeologia/avancelistado.html.

Neuendorf, K. K. E. *et al*. (2005): *Glossary of Geology. 5th Edition*. American Geological Institute (AGI), Alexandria, Virginia (USA).

North American Commission on Stratigraphic Nomenclature (NACSN) (2003): "North American Stratigraphic Code", *The American Association of petroleum Geologists Bulletin*, 67, 5. (May, 1983), Update 2003. Tulsa, Oklahoma (USA), American Association of Petroleum Geologists, pp. 841- 875.

North American Geologic Map Data Model Steering Committee (NADM) (2004): *NADM Conceptual Model1.0. A Conceptual Model for Geologic Information*. USGS Open-File Report 2004-1334, Reston, Virginia (USA). [Access: 09-06-2009]. Available at pubs.usgs.gov/of/2004/1334.

Pérez Cerdán *et al*. (2006): *Modelo de datos de la cartografía geológica MAGNA en formato digital*. Internal Report. Instituto Geológico y Minero de España, Madrid (Spain).

Richard, S. M., Matti, J. and Soller, D. (2003): "Geoscience Terminology Development for National Geologic Map Database", *Digital Mapping Techniques 2003*. USGS Open-File Report 2003-167. Reston, Virginia (USA), pp. 157-166.

Richard, S. M. (2006): "Geoscience Concept Models", in Sinha A. K. (Ed.): *Geoinformatics. Data to Knowledge: Geological Society of America.* Special Paper, 397. Boulder, Colorado (USA), Geological Society of America, pp. 81 - 107.

Richard, S. M. and Orr, T. M. (2001): "Data Structure for the Arizona Geological Survey Geologic Information System: Basic Geologic Map Data", *Digital Mapping Techniques 2001*. USGS Open-File Report 2001-223. Reston, Virginia (USA), pp. 167-188.

Richard, S. M. and CGI Data Model and Testbed Working Group (2007): "GeoSciML - A GML Application for Geoscience Information Interchange", *Digital Mapping Techniques 2007*. USGS Open-File Report 2007-1285. Reston, Virginia (USA), pp. 47-57.

Serra-Kiel, J. (2004): *Geological Map of the Dominican Republic to a scale of 1:50,000. Report/catalogue. Selection of Macroforaminiferals from the Geological Sheets to a scale of 1:50,000 for the K and L Projects* (*).

Soller, D. R. *et al*. (2002): *The Central Kentucky prototype: An Object-Oriented Geologic Map Data Model for the National Geologic Map Database*. USGS Open-File Report 02-202. Reston, Virginia (USA).

SLLT (2003): "Science Language for Geologic Map Databases in North America: A Progress Report", *Digital Mapping Techniques 2003*. USGS Open-File Report 2003-167. Reston, Virginia (USA), pp. 167-188.

Richard, S. M., Soller, D. R. and Graigue, J. A. (2005): *NGMBD Geologic Map Feature Class Model*. United States Geological Survey, Reston, Virginia (USA). [Access: 09-06-2009]. Available at support.esri.com/index.cfm?fa=downloads.dataModels.filteredGateway&dmid=30.

Ullrich, T. D. (2004): *Absolute Dating Using the $^{40}Ar/^{39}Ar$ Method of Sampling Igneous and Metamorphic Rocks from the geological sheets to a scale of 1:50,000 from the K and L Projects* (*).

Ullrich, T. D. (2004): *Absolute Dating Using the U/Pb Method of Sampling Igneous and Metamorphic Rocks from the geological sheets to a scale of 1:50,000 from the K and L Projects* (*).

USGS (2009): *Geologic Names Lexicon. Geolex*. United States Geological Survey, Reston, Virginia (USA). [Access: 09-06-2009]. Available at: ngmdb.usgs.gov/Geolex/

Weisenfluh, G. A. (2001): "Map Units Descriptions and the North American Data Model", *Digital Mapping Techniques 2001*. USGS Open-File Report 200-223. Reston, Virginia (USA), pp. 79-86.

(*) SYMSMIN 7 ACP DO 024 Program for the Geothematic Mapping of the Dominican Republic. Project L - East. Consortium IGME-BRGM-INYPSA. Dirección General de Minería, Santo Domingo (República Dominicana).

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

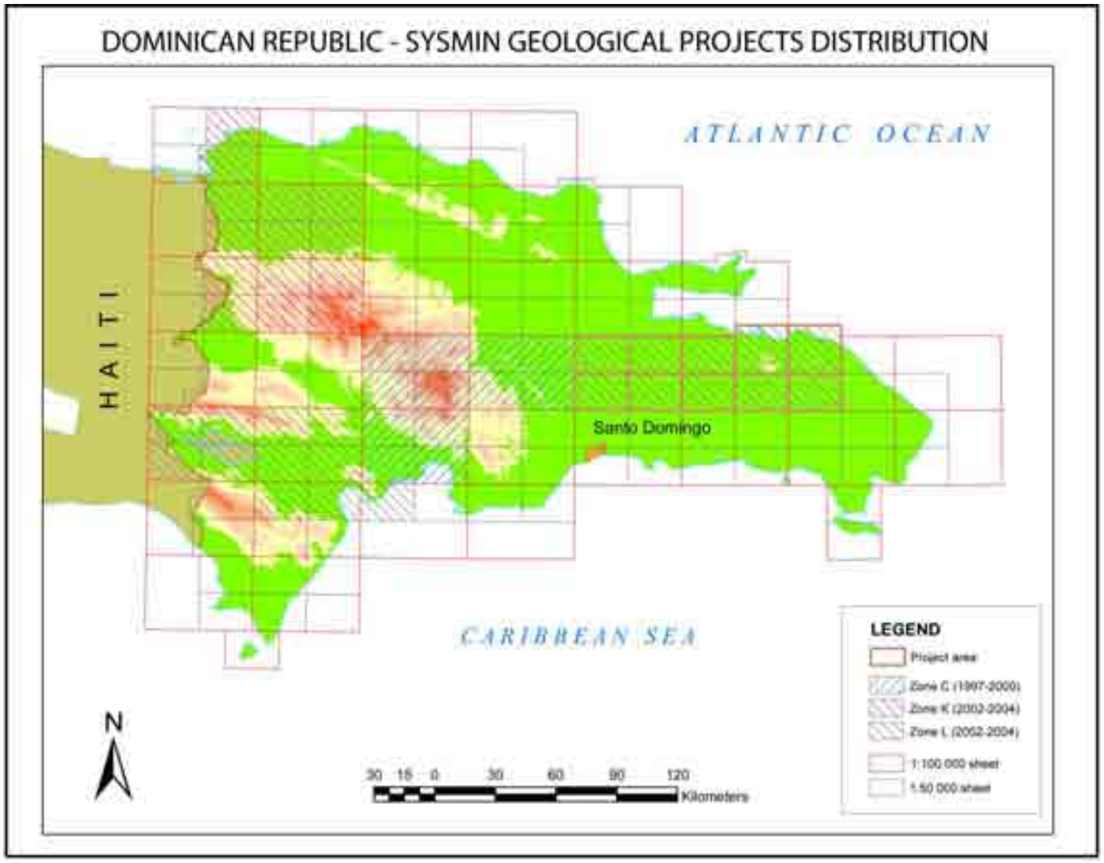
TABLES

Table 1. List of maps used for the data model design and implementation

Map type	Sheet	Scale	References
Geologic	6272-I	1:50.000	Hernáiz Huerta y Díaz de Neira, 2004
Geologic	6272-II	1:50.000	Monthel and Capdelville, 2004
Geologic	6272-III	1:50.000	Hernáiz Huerta, 2004
Geologic	6272-IV	1:50.000	Monthel *et al*., 2004
Geologic	6372-I	1:50.000	Díaz de Neira, 2004a
Geologic	6372-II	1:50.000	Monthel, 2004a
Geologic	6372-III	1:50.000	García Senz, 2004a
Geologic	6372-IV	1:50.000	Monthel, 2004b
Geologic	6472-III	1:50.000	García Senz, 2004b
Geologic	6472-IV	1:50.000	Monthel, 2004c
Mineral resources	6272	1:100.000	Lopera, Locutura and Florido, 2004a
Mineral resources	6372/6472	1:100.000	Lopera, Locutura and Florido, 2004b
Geomorphological	6272	1:100.000	Díaz de Neira, 2004b
Geomorphological	6272	1:100.000	Díaz de Neira, 2004c
Geomorphological	6272	1:100.000	Díaz de Neira, 2004d

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

FIGURES



Fig

figure 1. SYSMIN Geological projects distribution.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

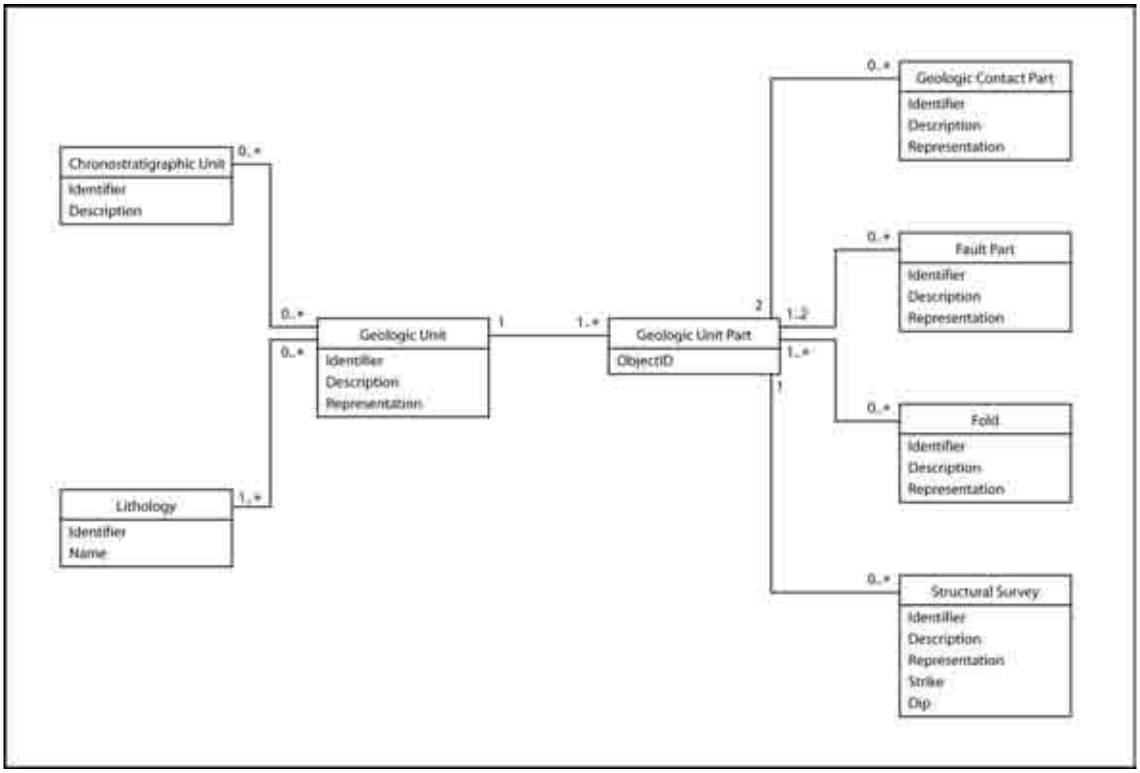


Figure 2. Original Data Model developed for Projects "K" and "L".

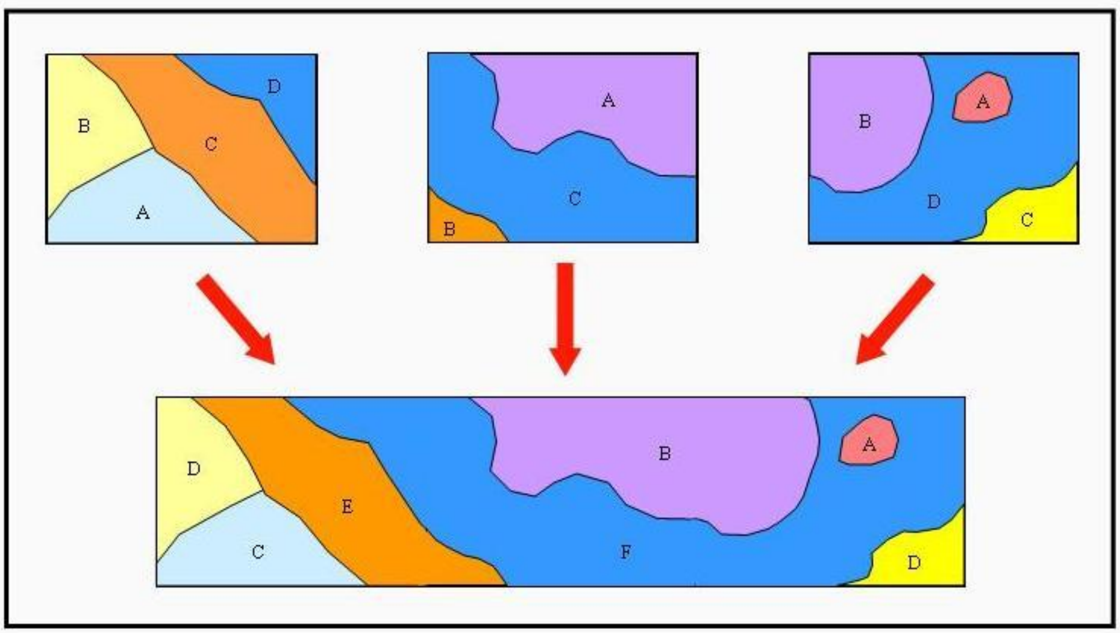


Figure 3. Continuity of the geographical domain.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

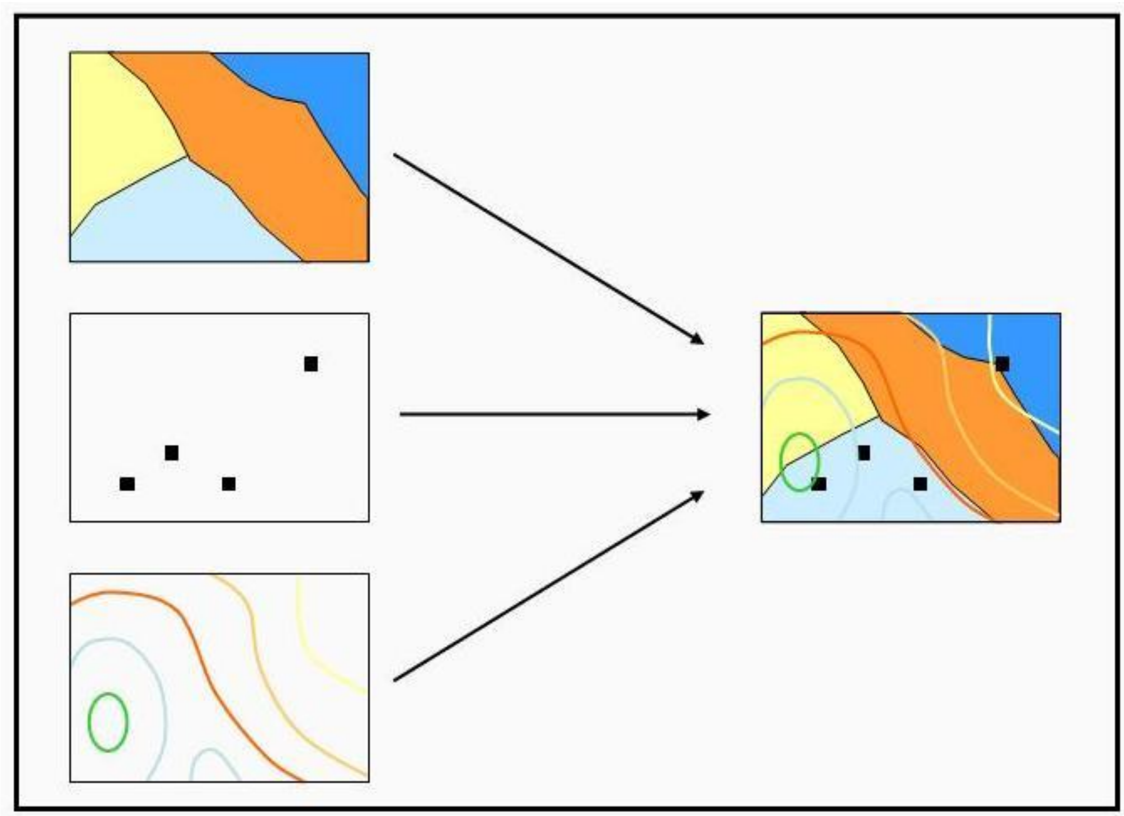


Figure 4. Visual integration.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

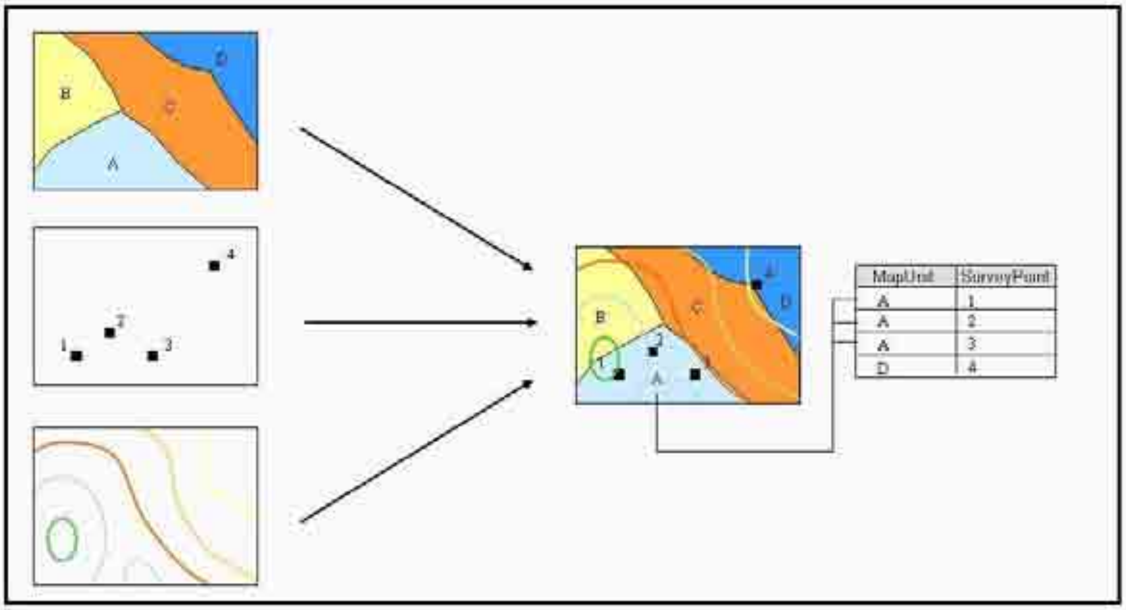


Figure 5. Spatial integration.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

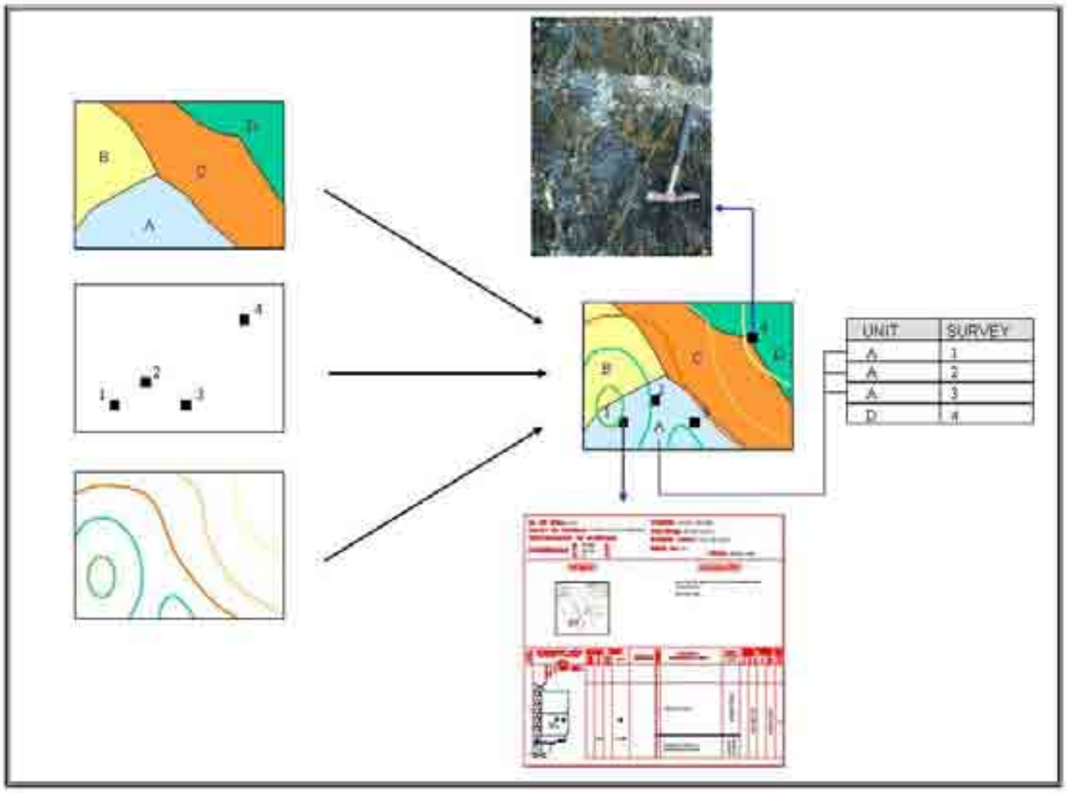


Figure 6. Document integration.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

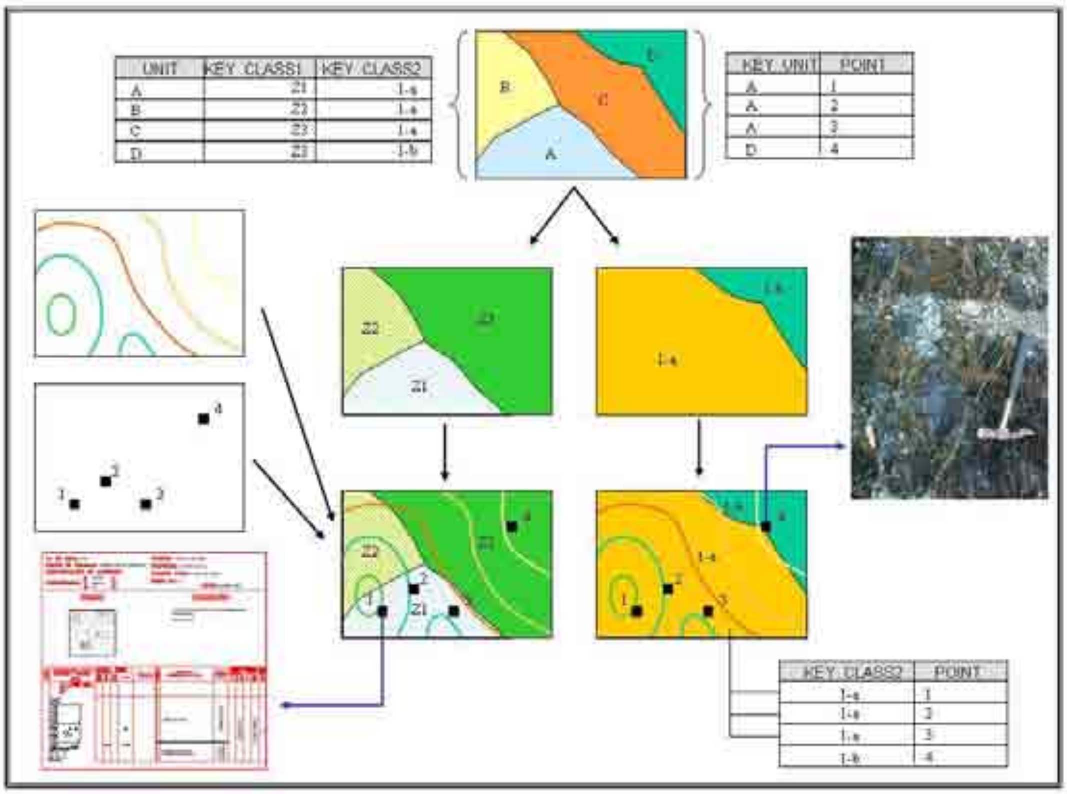


Figure 7. Conceptual integration.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

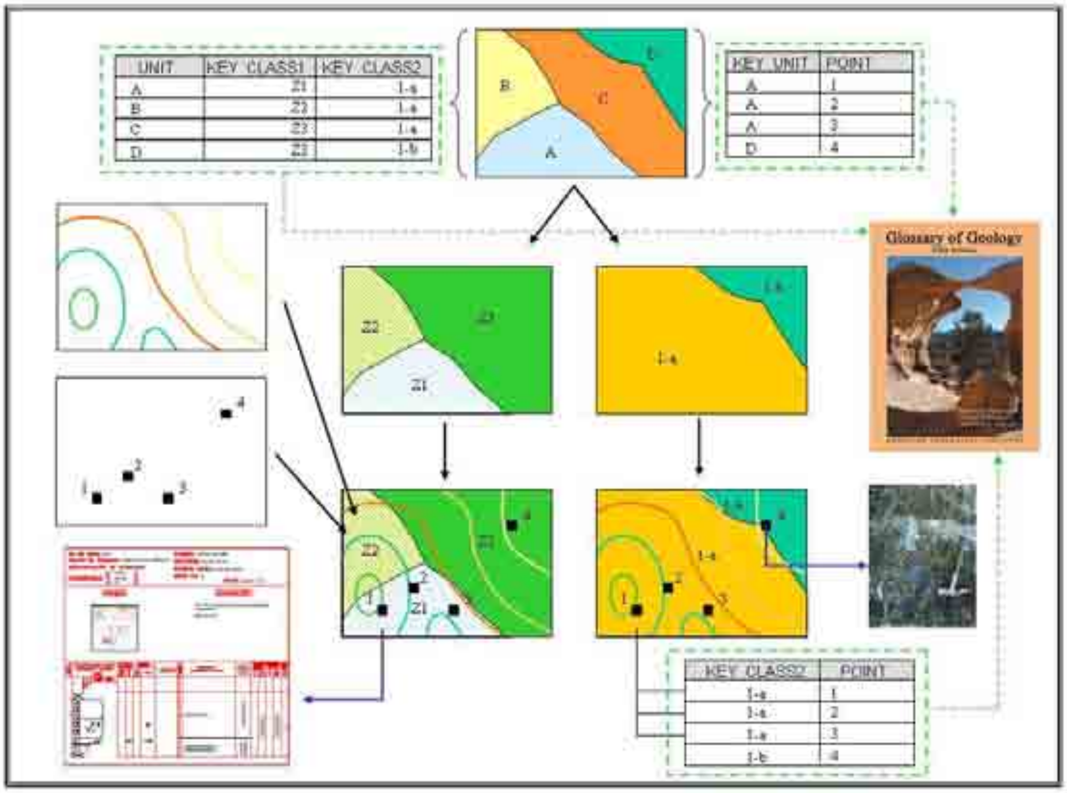


Figure 8. Semantic integration.

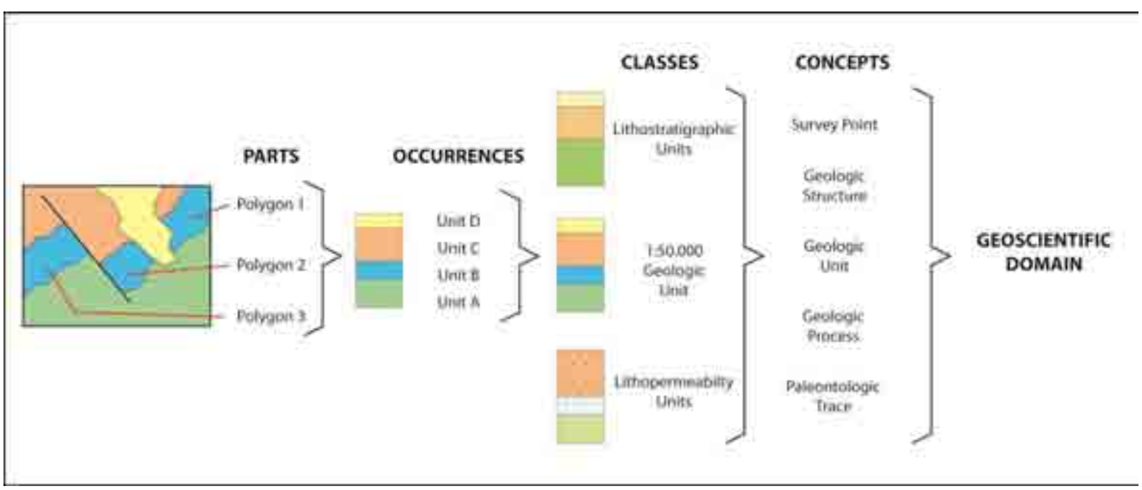


Figure 9. Division of the "Geologic Unit" concept into classes, occurrences and parts.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

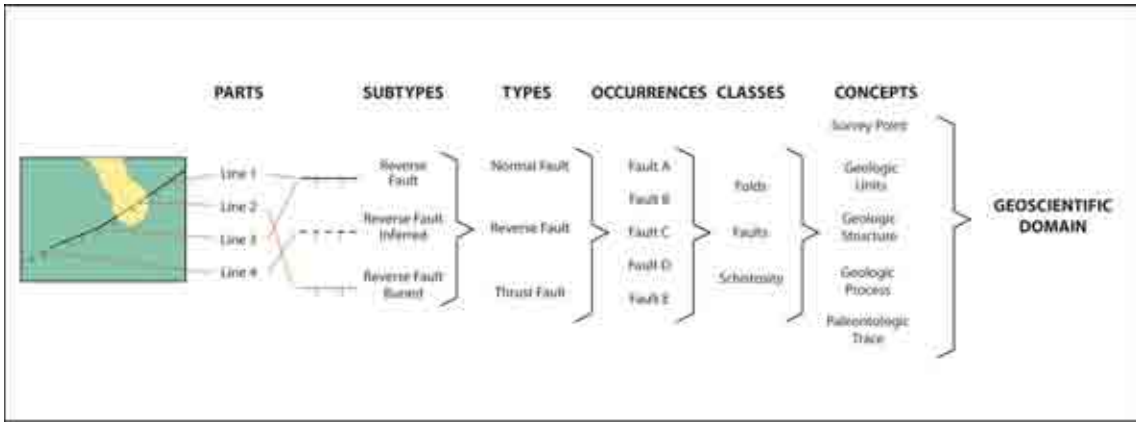


Figure 10. Division of classes of "Faults" into types, subtypes and parts.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

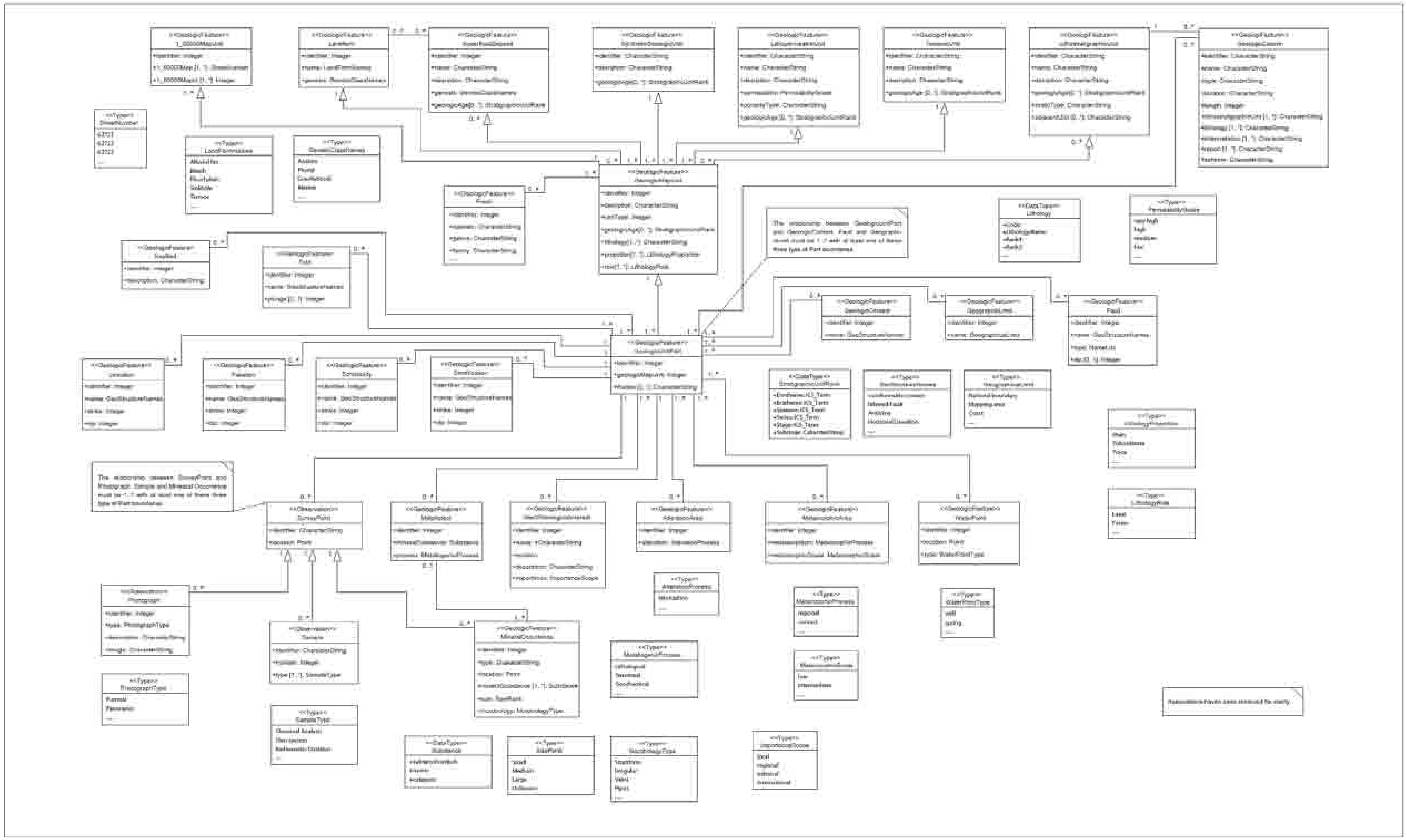


Figure 11. Application schema. (See attached figure file).

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

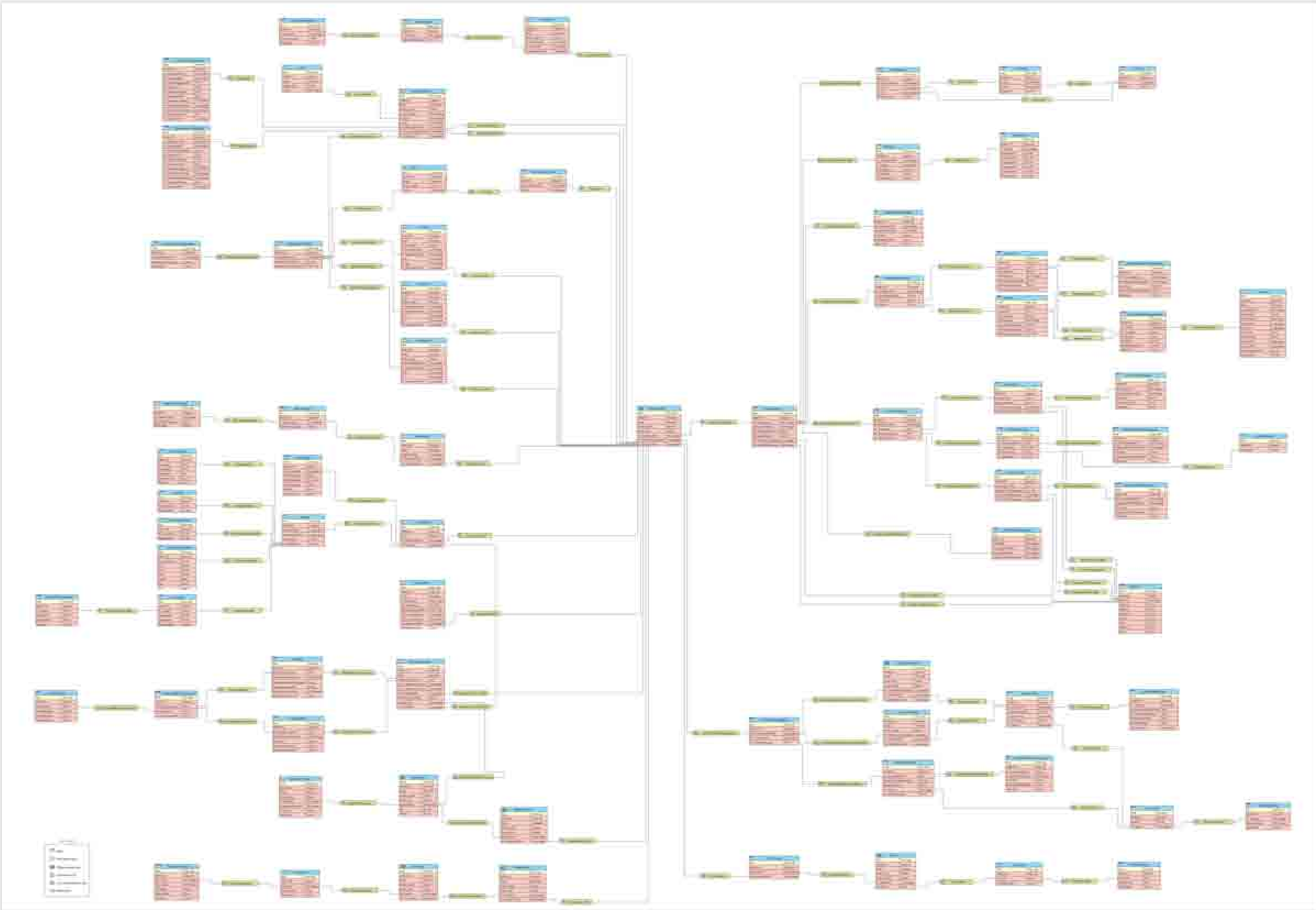


Figure 12. Implementation schema (See attached figure file).

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

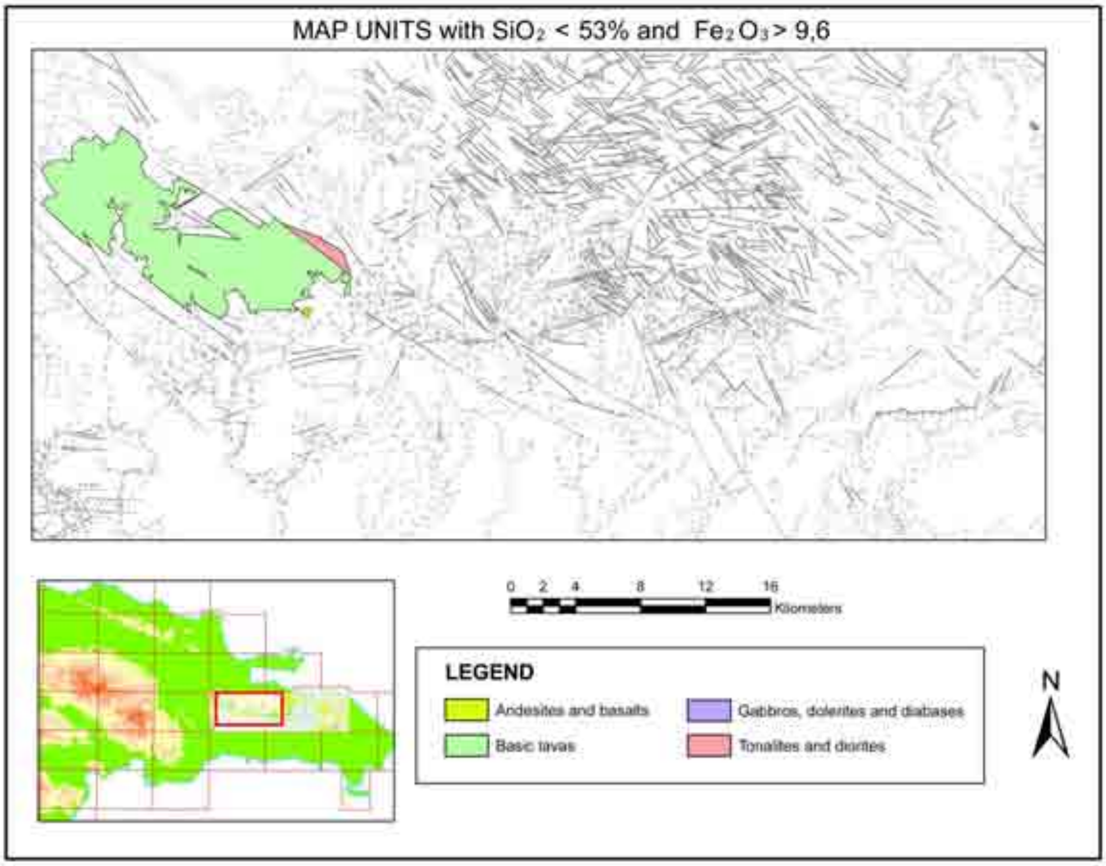


Figure 13. Geologic Units where a chemical analysis has been conducted with the oxide values referred to above.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

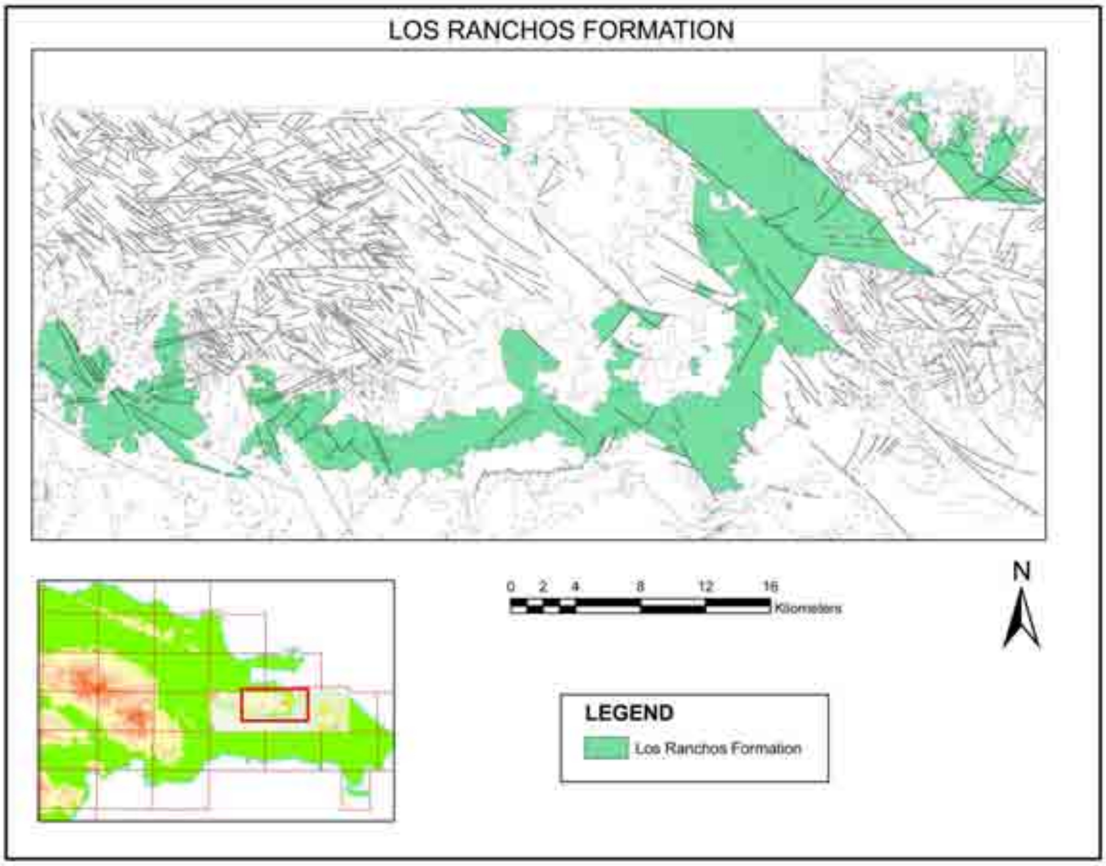


Figure 14. Geological Unit Parts that constitute the Los Ranchos Formation with a distribution of sheets to a scale of 1:50,000. It can be seen that the extension exceeds the limits of the sheets.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

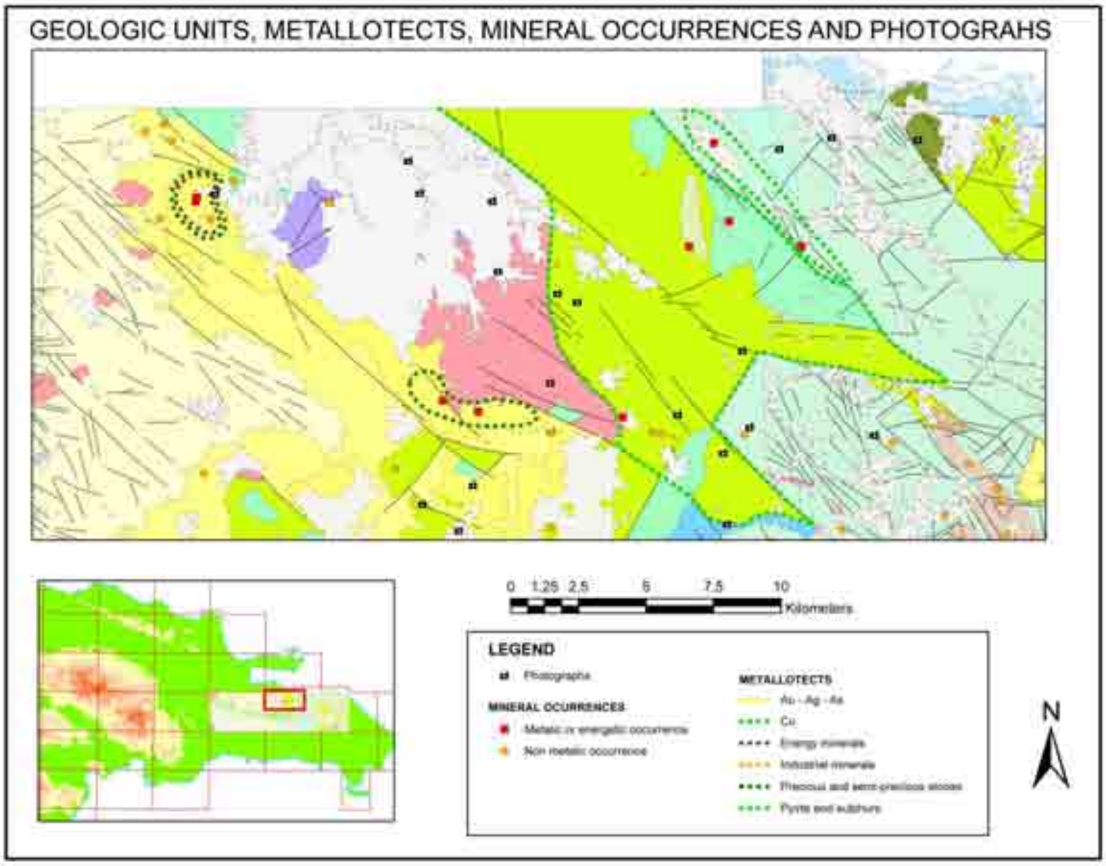


Figure 15. Map with Geologic Map Units (and Geologic Units Parts), Metallotects, Mineral Ocurrences and Photographs.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

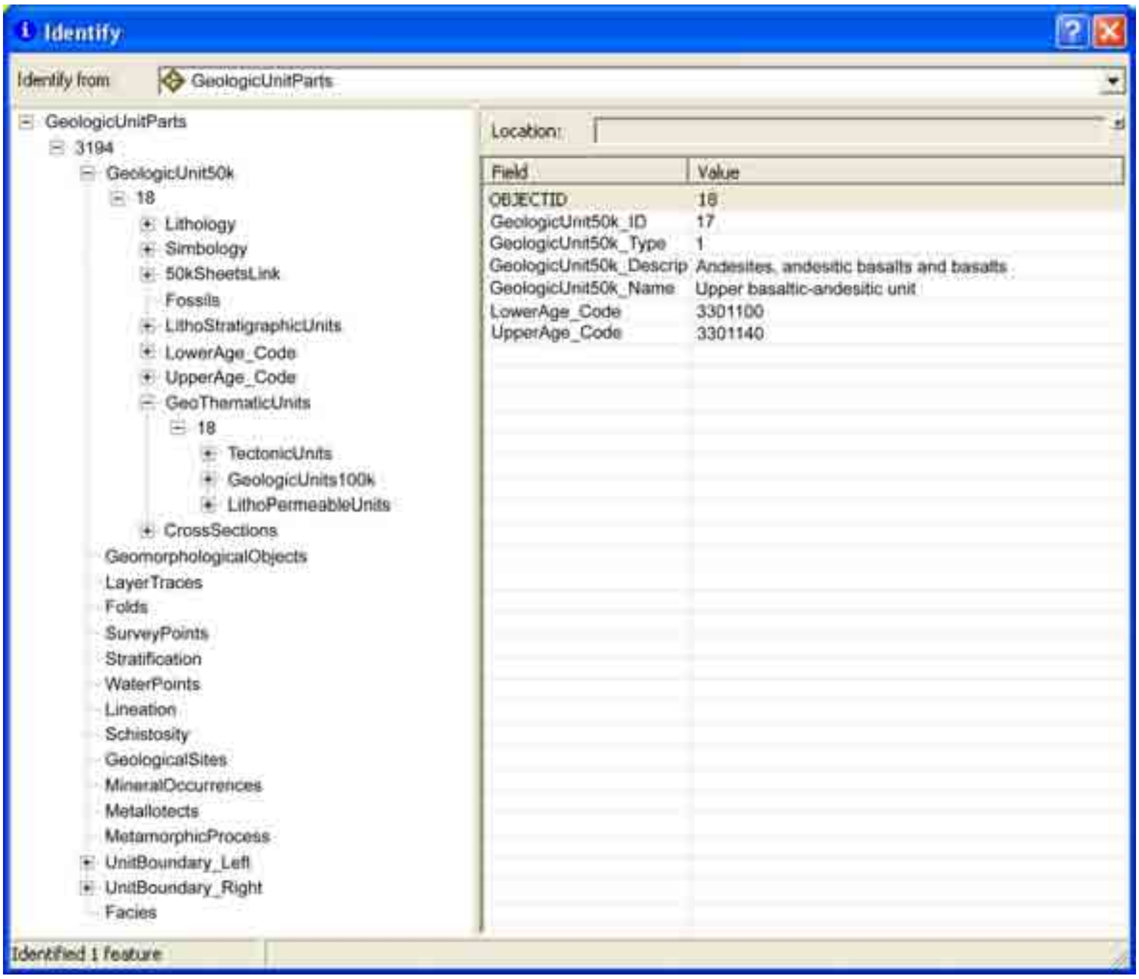


Figure 16. Window showing the results of the query about a Geologic Unit Part. The sign + to the left of each class indicates the presence of related occurrences of other classes.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

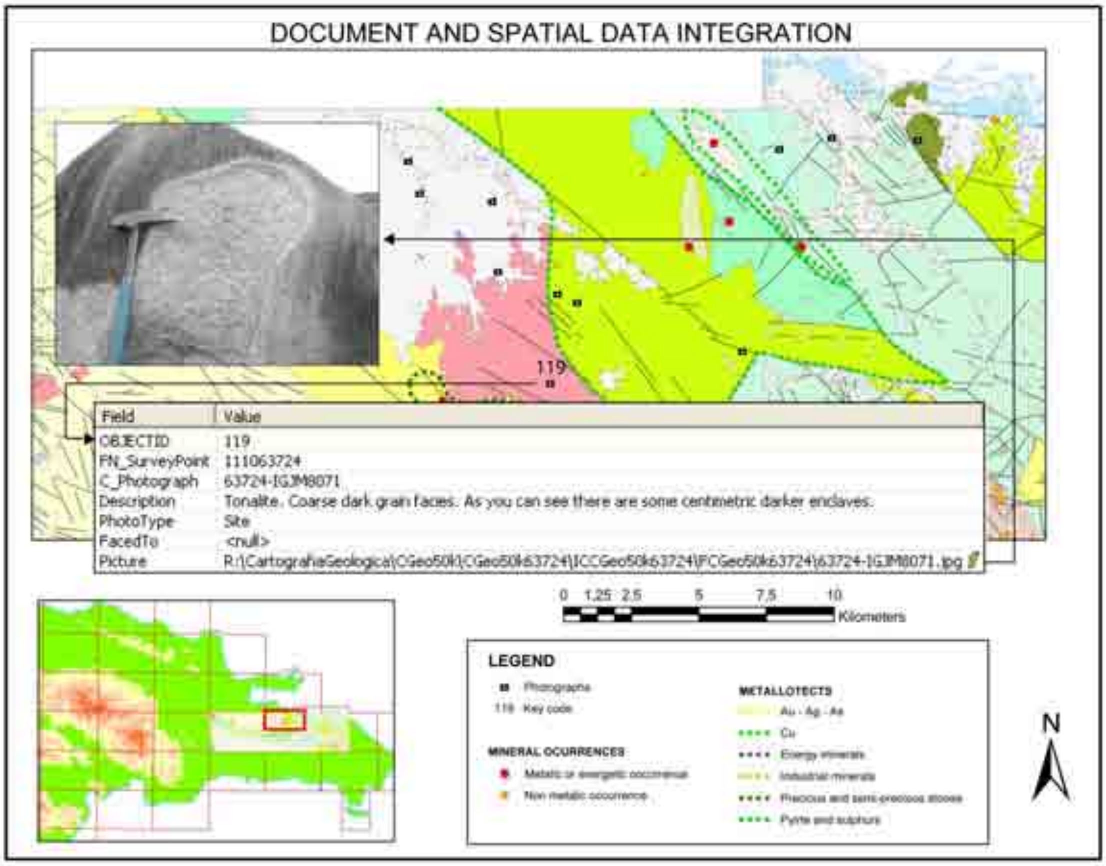


Figure 17. Composition with Singular Geologic Units, Metallotects, Mineral Occurrences and Photographs; window with the description of the photography and its location, and the image of the Photograph of a survey point numbered as 119.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

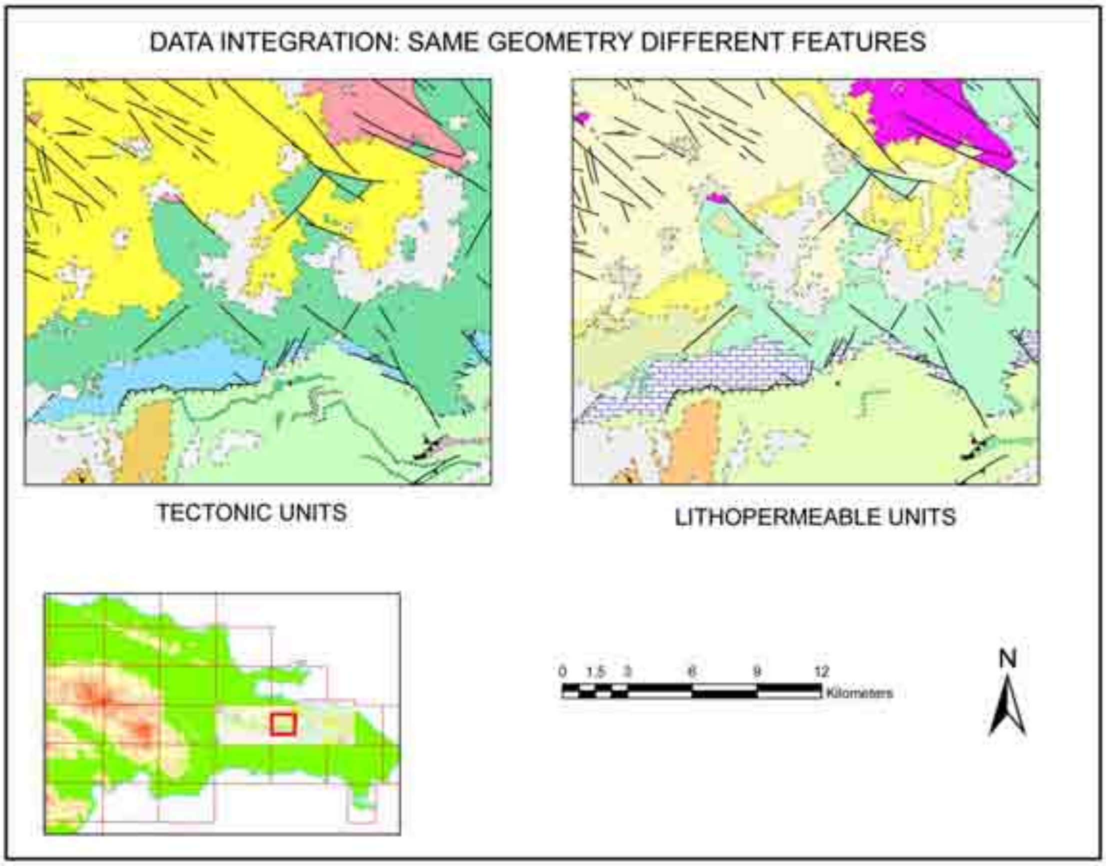


Figure 18. View of two different classes of Geologic Units generated from one single feature class.

Pérez Cerdán, F., Mancebo Mancebo, M.J., Lopera Caballero, E., Díaz de Neira, A., Monthel, A., Hernáiz Huerta, P.P. and García Senz, J. (2012): "Data Model of the ongoing Geoscientific Mapping of the Eastern Cordillera of the Dominican Republic", GeoFocus (Artículos), nº 12, p. 200-242. ISSN: 1578-5157

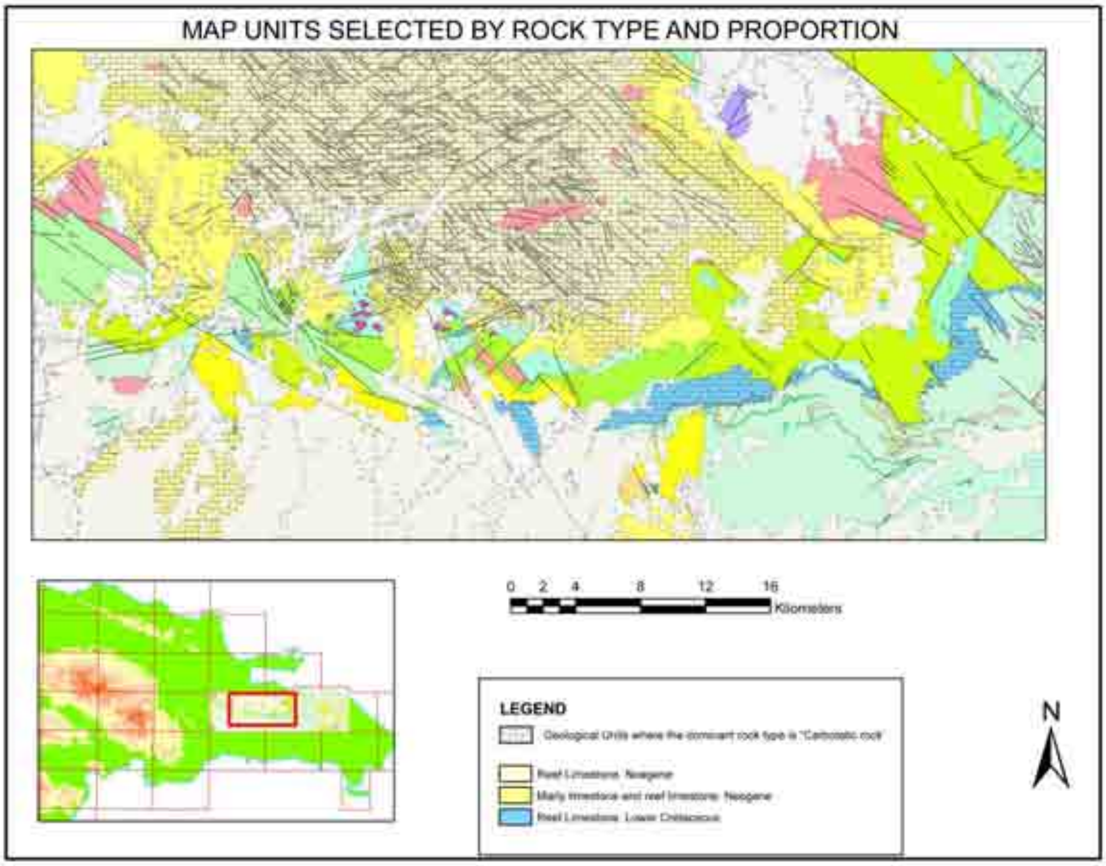


Figure 19. Map showing the selected Geologic Units by means of a query that combines the rock type (CARBONATIC ROCK) and the proportion (Main) within the Unit.