UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA

An Ontology of the Physical Geography of Portugal

Catarina da Conceição Gonçalves Rodrigues

MESTRADO EM ENGENHARIA GEOGRÁFICA E GEOINFORMÁTICA
Especialização em Sistemas de Informação Geográfica

2009

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA

An Ontology of the Physical Geography of Portugal

Catarina da Conceição Gonçalves Rodrigues

Dissertação orientada pelo Prof. Doutor Mário J. Gaspar da Silva

MESTRADO EM ENGENHARIA GEOGRÁFICA E GEOINFORMÁTICA
Especialização em Sistemas de Informação Geográfica

2009

Agradecimentos

Esta tese de mestrado foi um longo processo ao longo do qual foram várias as pessoas que participaram. Quero agradecer aos meus professores e colegas de pós-graduação em especial à Raquel, sem a qual esta dissertação não teria sido possível, ao Marcírio por toda a atenção e ajuda, ao Nuno por todas as discussões que possibilitaram um melhor trabalho,à Patricia e à Paula por toda a ajuda dispensada. Como é apanágio, quero agradecer à familia pelo incentivo dado. Finalmente quero agradecer ao meu orientados sem o qual, certamente, esta dissertação não teria sido realizada.

Resumo

Com o advento da Web Semântica é cada vez mais premente desenvolver novas formas de partilha de conhecimento que melhorem a interoperabilidade entre sistemas de informação geográfica (SIG). A modelaąão e representação do conhecimento geográfico sob a forma ontológica é uma das novas possibilidades. Esta dissertação estende uma representação ontológica, de acesso livre da geografia de Portugal, a Geo-Net-PT01, acrescentando-lhe, aos já caracterizados domínios da geografia humana e da web portuguesa, o domínio da geografia física. Para tal, foi incorporado no meta-modelo de informação pré'-existente suporte para exprimir informação geográfica geo-referenciada numericamente.

Desenvolveu-se uma metodologia para produção de modelos ontológicos incorporando o conhecimento do domínio da geografia física recorrendo ás fontes de produção de informação geográfica existentes. Esta metodologia foi utilizada na produção de uma nova versão da ontologia geográfica de Portugal, a Geo-Net-PT02, que agora incorpora dados sobre mais cerca de 24.000 entidades geo-referenciadas do território português.

Palavras chave: Ontologia, Informação geográfica, Domínio físico

Abstract

The advent of the Semantic Web raised the need for development of new methodologies for information sharing that improve interoperability among geographic information systems (GIS). The modeling and representation of geographic knowledge in the ontologic form is one of the new possibilities. This dissertation extends an open source representation of the geography of Portugal, Geo-Net-PT01, adding to the previously characterised domains of the human geography and the Portuguese Web, the domain of physical geography. To that purpose, the existing information meta-model was extended with support for expressing geo-referenced information in numeric form. A method for production of ontologic models incorporating the knowledge from physical geography from existing geographic information producers. This methodology was used in the production of a new version of the geographic ontology of Portugal, Geo-Net-PT02, which now incorporates data on over 24,000 geo-referenced entities in the Portuguese territory.

Keywords: Ontology, Geographical information, Physical domain

Contents

List of Figures

List of Tables

Chapter 1

Introduction

With the raise of globalization and awareness of spatial location, geographic information (GI) is assuming a larger relevance on our daily life. GI has a growing role on representations of the world (gathering space and time dimensions).

In parallel, we are witnessing the diversification of Internet access forms and a growing use of the web. Nowadays, having both a computer and a phone line is no longer a requirement to access the Internet. Access can be made through several kinds of portable devices.

The combination of these factors is the driver for a growing demand for GI services and motivates the development of new applications, requiring new access and representation methods; so, we are in a period of strong innovation on methodologies to handle this information.

Considering this, we foresee GI acquiring new uses and becoming a core element of Semantic Web technologies, therefore integrating two different knowledge domains:

1. One that existed for thousands of years (Geography/Cartography),
2. One that existed for the last five decades (Computing), but is already part of our daily lives.

To Antoniou & Harmelen (2004), with the evolution of computer science and its algorithms, a correctly modeled ontology contributes to the replacement of men by machines in routine tasks. Machines can be trained to be *smart*, learning to recognize and associate concepts. This need of ontological representations of geographic information is turning analog maps increasingly obsolete.

According to the W3C[1], the Semantic Web

> is a vision for the future of the web in which information is given explicit meaning, making it easier for machines to automatically process and integrate information available on the web.

So, the Semantic Web is a web of data. With the development of the Semantic Web the ontological representation of information becomes an important method for organizing shared knowledge, and consequently the representation of spatial information. According to Hendler (2001), the Semantic Web can be considered as

> the aggregation of a large number of small components of ontologies that appear interconnected amongst themselves.

It is necessary to emphasize that the semantics were not only related to the content of a domain, but also to how it links with other domains in the web. Therefore, it is essential that the domains made available are expressive so that the machines or agents are capable of processing and understanding the real meaning of the data.

A definition of ontology, which is gaining prominence, from the W3C[2] is

> ontology, defines the terms used to describe and represent an area of knowledge. Ontologies are used by people, databases, and applications that need to share domain information (a domain is just a specific subject area or area of knowledge, like medicine, tool manufacturing, real estate, automobile repair, financial management, etc.). Ontologies include computer-usable definitions of basic concepts in the domain and the relationships among them [...]. They encode knowledge in a domain and also knowledge that spans domains. In this way, they make that knowledge reusable.

[1] http://www.w3.org/RDF/FAQ
[2] http://www.w3.org/2003/08/owlfaq.html

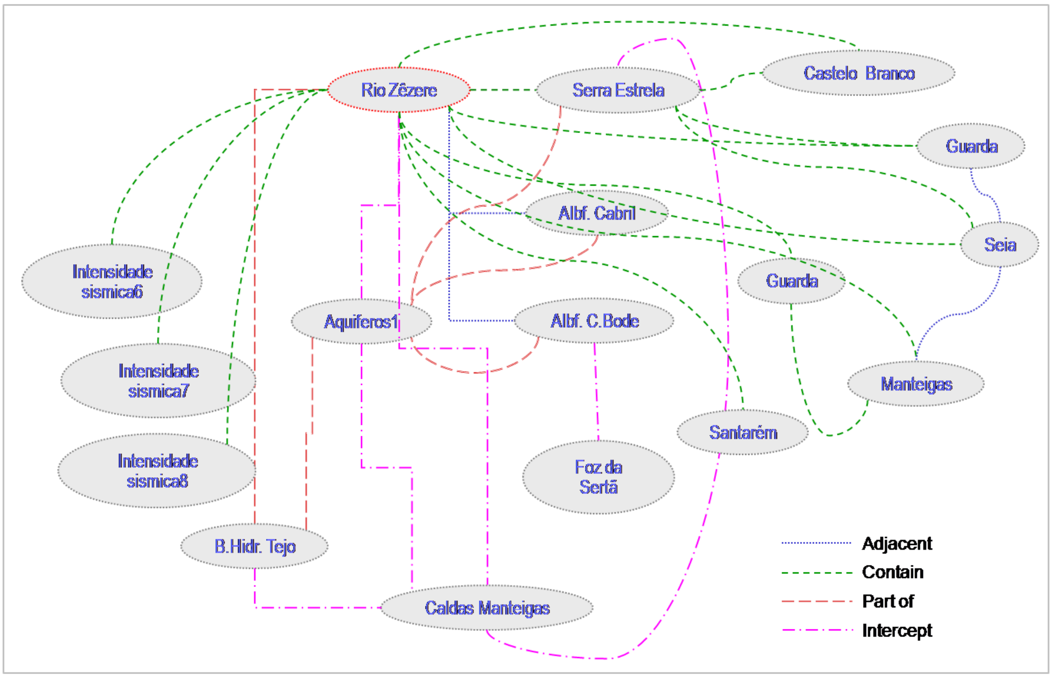

Figure 1.1: Excerpt from an object model for the physical geography of Portugal

Ontological representations of GI may be described in OWL format (XML-based), which is the Semantic Web standard for this purpose, therefore fulfilling the information interoperability requirement.

Some of the GI about Portugal was already integrated in a first public ontology of Portugal. The focus of the initial version of this ontology was the administrative domain of geographic information (the human geography). Administrative information is well structured and has strict relationships (e.g. *a parish is part of a municipality*). However, the GI of the physical geo-domain has a more complex structure (e.g.*a river could touch a mountain and intercept another mountain*), Figure 1.1 shows that. This information has a richer and more complex set of relationships (e.g. *crosses*, *contains*, *intersects* and *is-adjacent-to*): one river could *contain* an island, *cross* a mountain and *intercept* or be *adjacent* to a municipality. The modeling and integration of the physical domain into a new version of the Geo-Net-PT01, to be called Geo-Net-PT02, corresponds to a need for making this ontology more comprehensive.

To support the creation of geographic ontologies, the XLDB Group at LASIGE, University of Lisbon[1] developed the Geographic Knowledge Base (GKB). GKB was created under the project GREASE (Geographic Reasoning for Search Engines)[2]to support the conversion of geographical information into semantic knowledge. GREASE researches methods, algorithms and software architectures for assigning geographic scopes to Web resources and for retrieving documents using geographical features.

GKB integrates geographic data and knowledge from multiple sources under a common meta-model. This environment includes tools for deriving knowledge and to generating ontologies. GKB contains geographical information together with geographical attributes of network resources, such as Web sites and Internet domains.

1.1 Objectives and contributions

The scope of this dissertation is the construction of an ontological model for representation of the physical geography, and use it for the creation of a new Geo-Ontology of Portugal. To achieve this purpose, the following objectives have been defined:

1. Characterization of the main physical geography information requirements in web applications;
2. Systematic identification of authorities providing data in the geo-domain for Portugal; and
3. Collection and harmonization of that geographic information as one ontological representation.

The work in this dissertation resulted in several contributions:

1. Identification of relevant sources of geographic information for Portugal;
2. Development of a new information model for physical geography;

[1]http://xldb.fc.ul.pt/
[2]http://xldb.fc.ul.pt

3. Development of an ontological representation of GI, with a focus on the domain of physical geography, following recommended guidelines and formats for representing information on the Semantic Web;

4. Integration of this representation in the previous version of the public ontology of Portugal.

Geo-Net-PT02 is published, like Geo-Net-PT01, by Linguateca [1], a distributed language resource center for Portuguese.

1.2 Methodology

The creation of an ontology is a process with several phases. As in any other project of development of a large and complex body of information this is not a linear process. The construction of a Geo-Net-PT02 was segmented in six phases. The phases have been iterated, and it has been necessary at several points to backtrack to earlier steps.

Figure 1.2 illustrates the phases that constitute the process of creating a Geo-Ontology, as well as their interdependencies. Some of the phases have multiple sub-tasks. Some of the proposed work phases have been executed in parallel, since some depend on the progress of others. The main stages were the following:

1. **Systematic identification of concepts and authorities with data in the domain**. In this initial stage, a survey of all elements and concepts required to the development of the work has been performed.

2. **Definition of the conceptual model of the domain of physical geography and relationships with the administrative domain**. This phase involved the revision of the meta-model of GKB (i.e. the concepts used to describe the geographic entities in the knowledge base), and the design of the model representing concepts and relationships of the physical geography domain. Afterwards, I defined new relationships inferred from

[1] http://www.linguateca.pt/

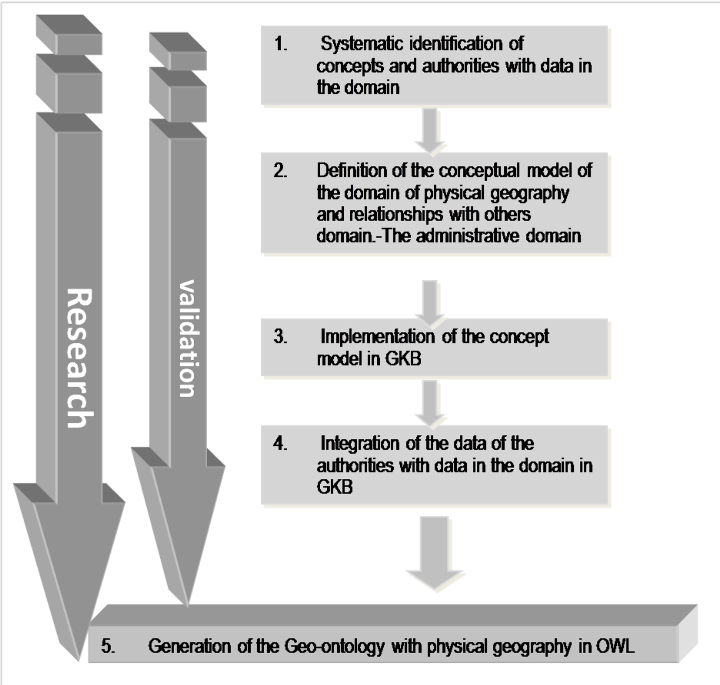

Figure 1.2: Ontology construction phases

the analysis of the collected data. In the next step of this phase, I defined relationships between concepts in the physical geography domain and existing concepts of the administrative domain.

3. **Implementation of the conceptual model for representation of physical geography in GKB**. In this phase, I performed the implementation of the conceptual model in GKB.

4. **Integration of the data in GKB**. In this phase all the physical domain data that was being identified and collected was cleaned and transformed to incorporation in GKB.

5. **Generation of the Geo-Ontology with physical geography in OWL**. After the integration of the data, the output files can be generated. In this phase, I tested the ontology to identify possible inconsistencies. After the resolution of these problems I generated the final ontology in the OWL format as it is now distributed.

The bibliographical research phase took place along all of the ontology creation and validation process. The validation of the collected data was also a process that was present along all the other phases of the construction of the ontology. After the generation of the ontology, it was again thoroughly reviewed.

1.3 Layout of the thesis

Chapter 2 presents the background information for the construction of a geographic ontology. The chapter introduces several standards related to geographic ontology construction as well as tools and methodologies needed for this purpose. It also provides a review of the state of the art on processes and methodologies for information representation. Geographic Knowledge Base (GKB) is also descriebed in this chapter.
The third chapter describes the extension to GKB to accommodate the new geo-physical domain model. The new model was used on the production of Geo-Net-PT02.
The fourth chapter presents the created physical ontology of Portugal, describes its sources of information,and the transformations of the collected data for insertion into GKB and the validation process.
Finally, the fifth chapter presents the main results, draws the conclusions and provides some directions for future work.

Chapter 2

Background and Related work

There are several definitions of Geography; According to Ptolemy (150 BC)

> the purpose of geography is to provide 'a view of the whole' Earth by mapping the location of places.

Another possible definition is

> ...concerned with the location or spatial variation in both physical and human phenomena at the Earth's surface (Kenzer (1989)).

Both definitions are based in the georeferencing and or geocoding of the information from different places of the Earth. Geography could be described by the scientific study of the location and spatial and time variation in both physical and human phenomena on Earth. Hence, geography could be separated in two domains:

1. Physical Geography; its main objective is the study of natural features on the surface land, or the study of the conditions of nature or natural landscape. It involves aspects from other sciences like Hydrography, Paleogeography, Geomorphology, Oceanography, Biogeography, Climatology, and Relief,

2. Human Geography; has its focus on the human study and description of the interaction between society and space. It helps to understand the man who lives in the area. One can understand the object of human geography

> as a critical reading of human perceptions and change on the area that comprises, in pass of time, as well as the impact of space on society and the study of man in which we live in sense of the relationship of man with the space. The limits of the municipality, parishes, and others limits with no physical frontiers as being assumed part of the human geography.

This chapter presents the two types of geographic references and its components and methodologies of representation, as so the main standards and to the understanding of the geographic information represented and modeled in the Geo-Net-PT02. In this chapter is presented some of the current work related, as so the Geographic Knowledge Bases (GKB).

2.1 Geographic references

To Hill (2006) georeferencing is

> relating information to geographic location.

Assuming that the georeferencing is the spatialization of the elements on Earth, today are there two forms to georeferencing Geographic Information (GI):

1. The numeric way of georeferencing information as coordinates, when information is given as coordinates or ontology terms/identifiers it is said to be georeferenced. Georeferenced model geospatial footprints which show on a map a particular spot of the Earth where some features are located,

2. The codified way of georeferencing GI is referred through locations names or codes (without coordinates): when a class has geographic information as part of its attributes is said to be geocoded. Geocoded entities are usually specified relatively to other georeferenced entities (e.g. *The Faculty of Sciences is in Campo Grande, near the Campo Grande metro station).*

The georeferenced information could be geocoded; if we define the relative terms; e.g. *what are the distance to a near feature.* One of the most important standards for georeferencing GI is the spatial reference, which consists of several components described below.

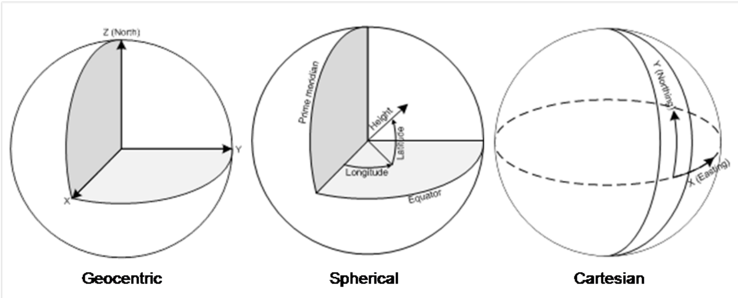

Figure 2.1: Coordinate systems

2.1.1 Numeric coordinates

To representing the numeric coordinates some rules and standards have been developed.

Coordinate system

Coordinate system is a referential system for describing a spatial property relatively to a system of axes. Figure 2.1 shows tree ways of doing this:

1. Geocentric coordinate system, it is based on a normal (X, Y, Z) coordinate system with the origin at the center of the Earth. This is the system that Global Position System (GPS) use internally for its calculations. However, it is very unpractical for humans to work with it, due to the lack of common concepts of east, north, up, down. It seldom directly displayed to the user before converting to another coordinate system.

2. Spherical Coordinate System. It specifies the position of a point P by the combination of its distance to the origin (radius r), the angle j between the x-axis and the line OQ, and the angle q between the radius and the z-axis.

3. Cartesian coordinate system or Spherical is also the Geographic coordinate system; this is probably the best well-known. It is based on angles relative to a prime meridian and Equator usually as Longitude and Latitude, as so the heights are usually given relative to either the mean sea level or a well-know georeferenced point called the datum). Cartesian coordinate system it defined as a *flat* coordinate system placed on the surface of Earth.

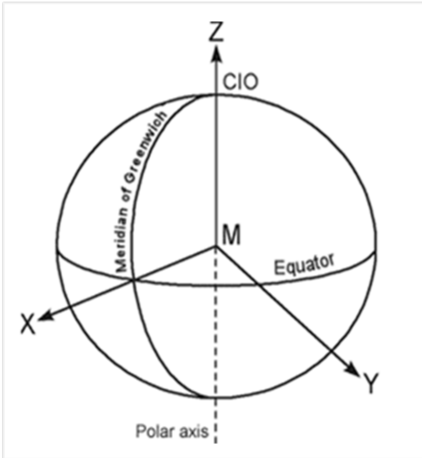

Figure 2.2: Conventional terrestrial system (CTS)

Global Terrestrial Reference System is a geocentric Cartesian coordinate systems that have their origin in the Earth's center of gravity, (see Figure 2.2. CTS has its point of origin at M, its Z axis passing through the North Pole (designed as the Conventional International Origin CIO), its XZ plane passing through the mean zero point in Greenwich, and its Y axis completing the definition of a right-handed coordinate system. Figure 2.2 shows the point of origin M, and illustrates the axis of the CTS.

Ellipsoids

With the exception of the Geocentric Coordinate System (defined relatively to the center of the Earth) all systems use a height system relative to the surface of the Earth. Earth surface of as being relatively at the mean sea level, beside the form of the Earth is often called by an Ellipsoid but we don't get a spheroid or even an ellipsoid; the really shape of the Earth planet are the geoids shape (Figure 2.3).

Because of gravitational changes often caused by large masses e.g. mountain ranges, Earth is irregular with variations around 100 meters, Figure 2.4 shows that variations in Continental Portugal.

The center and orientation of the ellipsoid is what we call the datum. So the datum is a set of basic parameters (numerical quantities or geometric entities)

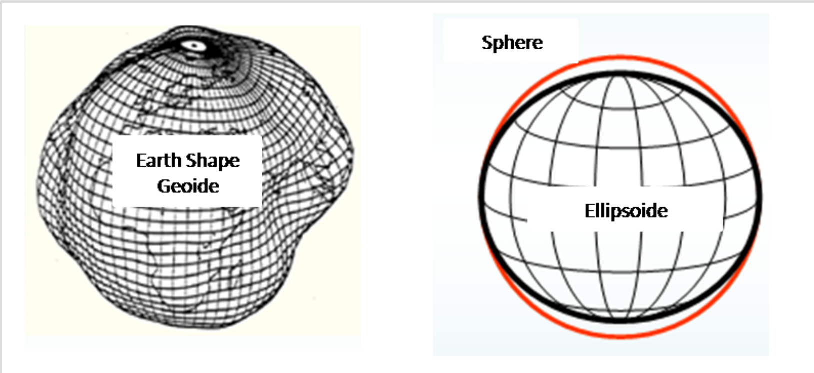

Figure 2.3: Shape of the Earth

who defines an ellipsoid and through the use of a set of points on the ground that we relate to points on the ellipsoid, we define the center of the Earth.

A geodetic datum is a set of parameters (including axis lengths and offset from the true center of the Earth). For each region, a different datum can be chosen so that it best matches average sea level, therefore terrain features. The data acquisition for a map involves surveying, or measuring heights and distances of reference points as deviations from a specific datum. The shape, size and location of the center of the ellipsoid for the center of mass of the Earth are determined to ensure that the surface of the ellipsoid to adapt our best to land surface in a given country or region.

The most common is the Datum of the World Geodetic System 1984 (WGS84), a global datum, which uses an ellipsoid of revolution equipotential and geocentric to express the geographical, coordinates (latitude, longitude and ellipsoidal altitude). For its definition a set of reference stations has been used, distributed throughout the entire globe. The origin of the WGS84 coordinate system is located in the center of mass of Earth. Is the system used by the GPS systems. It is a good approximation of the entire world and with fix-points defined almost all over the world. However, these do not include points in Europe. To Europe the most well-new is the European Datum 1950 (ED50) and, after 1989 the standard and official datum is defined by European Reference Frame [1], sub

[1] http://www.euref-iag.net/

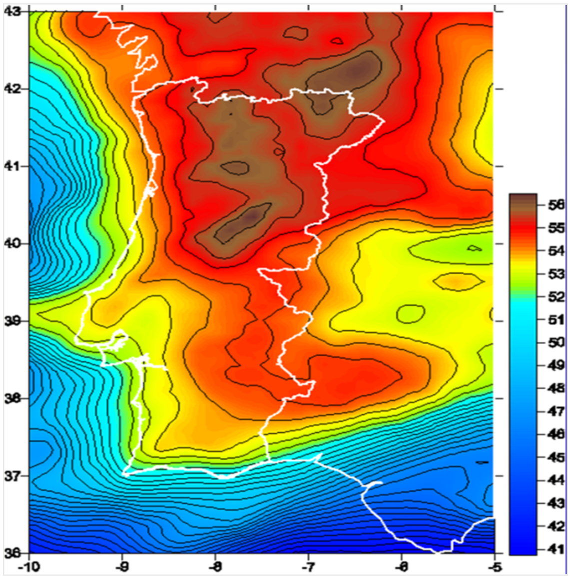

Figure 2.4: Continental Portugal geoide

commission of the International Association of Geodesy of the International Association for Geodesy (IAG)[1]. They have defined a global system to be used for all the members. ETRS89 is a global system of reference recommended by the EUREF established through space techniques. The establishment of the ETRS89 in Continental Portugal was a result of international campaigns, with the objective of binding the Portuguese network conveniently to the Euronet. That work has carried by Instituto Geográfico Portugês (IGP) [2].

Besides the European reference system ETRS89, older national systems are used in European countries. In Continental Portugal the most used data are the datum Lisbon and datum 73. The datum Lisbon, a geodetic datum, is the oldest and has its origin in the Castle of São Jorge, Lisbon. The datum73 use an origin point more central to their Geodesic, the Apex Melriça, both date using the ellipsoid of Hayford-Gauss.

Spatial reference

[1]http://www.iag-aig.org/
[2]http://www.igeo.pt/

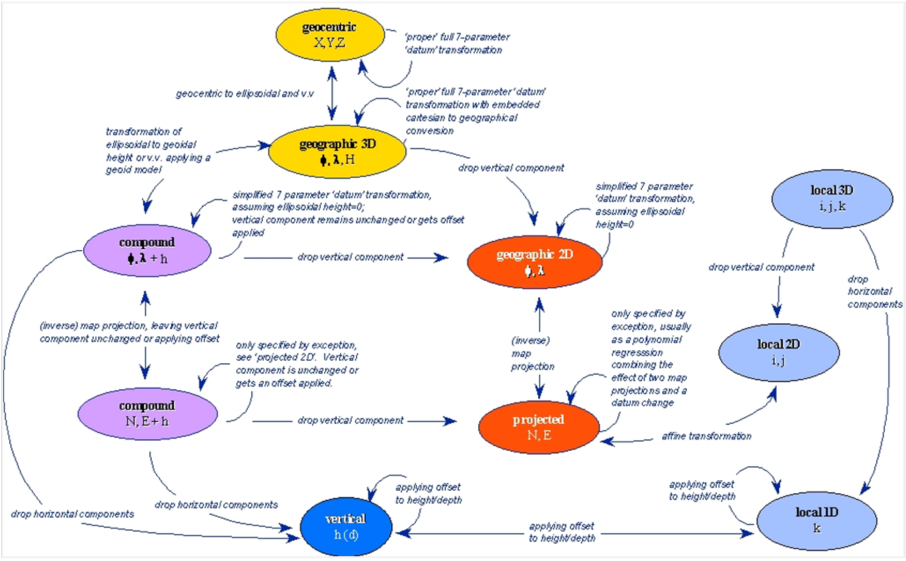

Figure 2.5: Coordinate system and transformations

The spatial reference is a combination of all the above. It defines an ellipsoid, a datum using that ellipsoid, and either a geocentric, geographic or coordinate system. The projection always has also a geographic coordinate system associated with it. The European Petroleum Survey Group (EPSG) has a huge set of predefined spatial references, each given with a unique ID. These ID are used throughout the industry. Spatial References are often defined in a Well-known format defining all these parameters.

There are forms to make possible the transformation of the coordinate systems, to enable their interoperability. Figure 2.5 shows examples of the coordinates systems and their transformations.

All of these concepts are more developed and explained by Casaca (2005) and Matos (2001).

2.1.2 Representing place names

Terms are the basic semantic units for conveying concepts. They are usually single-word nouns, since nouns are the most concrete part of speech. Verbs can be converted to nouns -cleans to cleaning, reads to reading, and so on. Adjectives and adverbs, however, seldom convey any meaning useful for indexing.

One of the most important types of GI is place names. Almost all the GI could be georeferenced through references to place names on Earth, e.g. *The Tagus River flows in Portugal and Spain, in the European Continent.* There are several ways of representing places names; one of these are based on a *controlled vocabulary*, which consists on a list of explicitly enumerated terms, published by a registration authority. All terms in a controlled vocabulary should have an unambiguous, non-redundant definition (Schmidt, 2006).

This form must follow at minimum two rules to respect the definition:

1. If the same term is commonly used to mean different concepts in different contexts, then its name is explicitly qualified to resolve this ambiguity,

2. If multiple terms are used to mean the same thing, one of the terms is identified as the preferred term in the controlled vocabulary and the other terms are listed as synonyms or aliases.

Another form of data representation is a *taxonomy*, which is a collection of controlled vocabulary terms organized into a hierarchical structure (Garshol, 2004). Each term in a *taxonomy* is in one or more parent-child relationships to other terms in the taxonomy. There may be different types of parent-child relationships in a taxonomy (e.g., whole-part, type-instance), but good practice limits all parent-child relationships to a single parent and of the same type. Some taxonomies allow poly-hierarchy, which means that a term can have multiple parents. This means that if a term appears in multiple places in a *taxonomy*, then it is the same term. Specifically, if a term has children in one place in taxonomy, then it has the same children in every other place where it appears. One example of a *taxonomy* is the Yahoo website.

A *Thesaurus* is another form of data representation. A formal definition of a thesaurus (Garshol, 2004) designed for indexing is:

- A list of every important term in an domain of knowledge,

- A set of related terms for each term in the list.

Term relationships are links between terms that often describe synonyms, near-synonyms, or hierarchical relations. Synonyms and near-synonyms are indicated by a Related Term. The way the term ´´coordinate" is related to the term ´´geography" is an example of such a relationship. Hierarchical relationships are used to indicate terms which are narrower and broader in scope.

Another definition for a Thesaurus is given by Gonzalez & Lima (2001), where a thesaurus is a lexical database that implements an ontology, where the described terms should be interpreted as concepts and its relationships constitute the essence of the description of these concepts. A thesaurus is a closed language restricted normally by three relationships (ISO 19136:2007)[1]:

1. Equivalence (Used For and USE),
2. Hierarchy (Broader Term/Narrower Term),
3. Associativity (Related Term).

One example of a thesaurus is the UNESCO Thesaurus[2], which was developed by UNESCO for use in the indexing and retrieval of information in the UNESCO Integrated Documentation Network. It was first published in 1977. A second edition was issued in 1995. Booth use a controlled vocabulary developed by the United Nations Educational, Scientific and Cultural Organization, which includes subject terms for the following areas of knowledge:

- Education,
- Science,
- Culture,
- Social and human sciences,
- Information and communication,
- Politics, law and economics.

[1] http://www.iso.org/
[2] http://www2.ulcc.ac.uk/unesco/

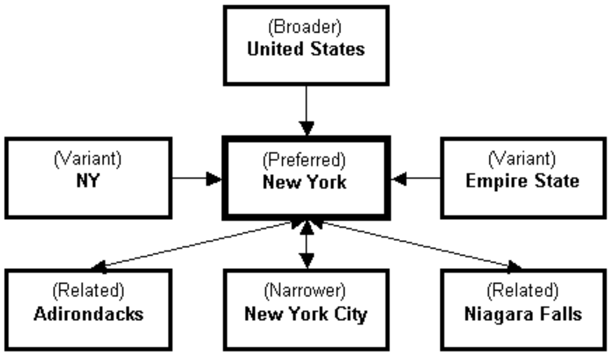

Figure 2.6: A UNESCO Thesaurus example

It also includes the names of countries and groupings of countries:

- Political,
- Economic,
- Geographic,
- Ethnic and religious,
- Linguistic groupings.

The UNESCO Thesaurus allows subject terms to be expressed consistently, with increasing specificity, and in relation to other subjects. It can be used to facilitate subject indexing in libraries, archives and similar institutions. As in other subject thesauri, the terms in the UNESCO Thesaurus are linked together by the three types of relationships referred in ISO 8879.

- Hierarchical relationships, which link terms to other terms expressing more general and more specific concepts, e.g. broader terms and narrower terms,
- Hierarchically related terms are grouped under general subdivisions (known as "micro-thesaur´´), which in turn are grouped into the areas of knowledge covered by the Thesaurus,
- Associative relationships, which link terms to similar terms (related terms) where the relationship between the terms is non-hierarchical,

- Equivalence relationships, which link *non-preferred* terms to synonyms or quasi-synonyms which act as *preferred* terms(see Figure 2.6).

They also include scope notes which explain the meaning and application of terms, and French and Spanish equivalents of English preferred terms. The UNESCO Thesaurus has not a Portuguese version, so its information could not be used by applications in Portuguese.

Another example of a thesaurus is the Getty Thesaurus of Geographic Names (TGN, a structured vocabulary that includes names and associated information about both current and historical places around the globe[1]. The focus of TGN records are only places and their various forms of presentation (historical, common alternative and in a multilingual form), the place's parent or position in the hierarchy, other relationships, geographic coordinates, notes, sources for the data, and place types, which are terms describing the role of the place (e.g., inhabited place and state capital). There may be multiple broader contexts, making the TGN polyhierarchical. In addition to the hierarchical relationships, the TGN has equivalent and associative relationships.

2.1.3 Computer machine representation of geographic information

Geographical information is representing in many types of applications, and available to large amount of users. To make possible the representation of geographic information through the computer machine some standards has been created.

Standard for Unicode Transformation Format-8(UTF-8)

To make them usable by all a common encoding for geographic references must be adopted. In early systems, there were hundreds of different encoding systems for representing here numbers in coordinates systems or the location names. No

[1]http://www.getty.edu

Figure 2.7: Encoding examples

single encoding could contain enough characters (e.g. only the Europe requires several different encodings to cover all its languages). Even for a single language like English no single encoding was adequate for all the letters, punctuation, and technical symbols in common use. These encoding systems also conflicted with one another. That is, two encodings can use the same number for two different characters, or use different numbers for the same character. Whenever data is passed between different encodings or platforms, that data always runs the risk of corruption. There are variously types of encodings (M.Davis, 2008) (see Figure 2.7):

1. ASCII,

2. ISO8859,

3. WINDOWS 1552,

4. UTF-16,

5. Big 5.

The Unicode standard[1] is a character coding system designed to support the worldwide interchange, processing, and display of the written texts of the diverse languages and technical disciplines of the modern world. In addition, it supports classical and historical texts of many written languages. To Unicode.org *Unicode*

[1] http://www.unicode.org/standard/standard.html

provides a unique number for every character, no matter what the platform, no matter what the program, no matter what the language. In Web standards data is mostly represented in the XML format (Harold, 2005) and UTF8 is the standard for encoding Unicode chars. It is an octet (8-bit) lossless encoding of Unicode characters. UTF-8 encodes each Unicode character as a variable number of 1 to 4 octets. UTF-8 is the default encoding for XML.

The advantages are:

- UTF-8 can encode any Unicode character. Usually most encodings can be converted to Unicode and roll back with no loss of information. Character boundaries are easily found from anywhere in an octet stream,

- A byte sequence for one character never occurs as part of a longer sequence for another character. For instance, US-ASCII octet values do not appear otherwise in a UTF8 encoded character stream. This provides compatibility with file systems or other software.

Relational Databases

GI information can be represented in the form of a database; they can be seen as a database possible to geocoding or georeferencing. This Geo-Ontology is based in a relational database. That is a database that conforms to the relational model, and could also be defined as a set of relations or a database built in a relational database management system (RDBMS)) (Schmidt, 2006).

More strictly, a relational database is a collection of relations or tables. A relation is defined as a set of tuples that all have the same attributes. In a relational database, all of the data stored in a column should be in the same domain (i.e. data type). In the relational model, the tuples should not have any ordering. In this type of database neither the rows nor the columns should have an order to them.

2.2 Geographic data types

Geo-Net-PT02 integrated information from the named physical geography as so the human geography, trough the first version of Geo-Net-PT01. So, all the alphanumeric data represented in this Geo-Ontology is possible to represent in a vector map. Vector data is what most people think of when they represent GI a map. The lines on a road map, the points that indicate cities, and the polygons that enclose a state are good examples of a vector map and its components. This type of data looks the most realistic to people, and when used in GIS applications can generate the most accurate data. Much of the data that comes from IGP (topographic data) and IA (Environment Institute) is in some type of vector format.

Vector is a generic term; there are many vector formats such as DLG, SDTS, DXF, SHP, etc. Some of this is specific to only one software. Vector format is recommended when accuracy is needed (e.g. boundary information), or very realistic looking maps are desired. To be integrated in the Geo-Ontology is necessarily the vector data have alphanumeric data, is the coordinates of the vector and the alphanumeric data, and the metadata who integrate the Geo-Ontology. Not all of the formats have that, besides this all of the data used in GIS applications have this duality of information. This GI has a vector linked to a database; the named alphanumeric information, the coordinates could be alphanumeric information or not. The GIS format has the coordinates of the features implicit, as so the scale. They are different vectors types. To Hill (2006) The Geographic Information can be represented by:

- Lines,
- LineString,
- LinearRing,
- Polygon,
- Point,
- MultiPoint,
- MultiLineString,
- MultiPolygon,

- MultyGeometry.

In this master thesis I group all of this vector types in three different types:

- Lines; consist in Lines, LineString, LinearRing, LineString and MultiLineString,
- Points; consist in Points and Multipoint,
- Polygons; consist in Polygon and MultiPolygon.

I do not represent the MultyGeometry because in my opinion this is not a need type of vector.

The Figure 2.8 show examples of the types of data used, all of this data are characterized to be a vector data. The lines in this figure represents the Rivers, this data is characterized for a set of ordered coordinates that represent the shape of geographic features.

The GI, reservoir represented in a Point. To Hill (2006)

> the shape point of the GI is "A zero-dimensional abstraction of an object represented by a single X, Y coordinate."

Figure 2.8 shows the localizations of some features in a point shape, e.g. the localizations of a water point or the location of a service cover on a medium scale map or a city localization. Same picture shows to the shape polygon (area) e.g. the administrative limit, or the limit of some features reservoir.

2.3 Metadata for geographic information

Metadata is critical for sharing GI, they turn possible searching to see if the resources we need already exist. Metadata describes GI resources in the same way a card in a library card catalog describes a book. When we found a resource with a search, its metadata will help you decide whether it is suitable for your purposes. To make this decision, you may need to know how accurate or current the resource is and if there are any restrictions on how it can be used, what is the year or the font of data, with this information we could make the evolution of the GI as well know some particularities of the data.

According to Federal Geographic Data Committee (FGDC) [1]

[1] http://www.fgdc.gov/

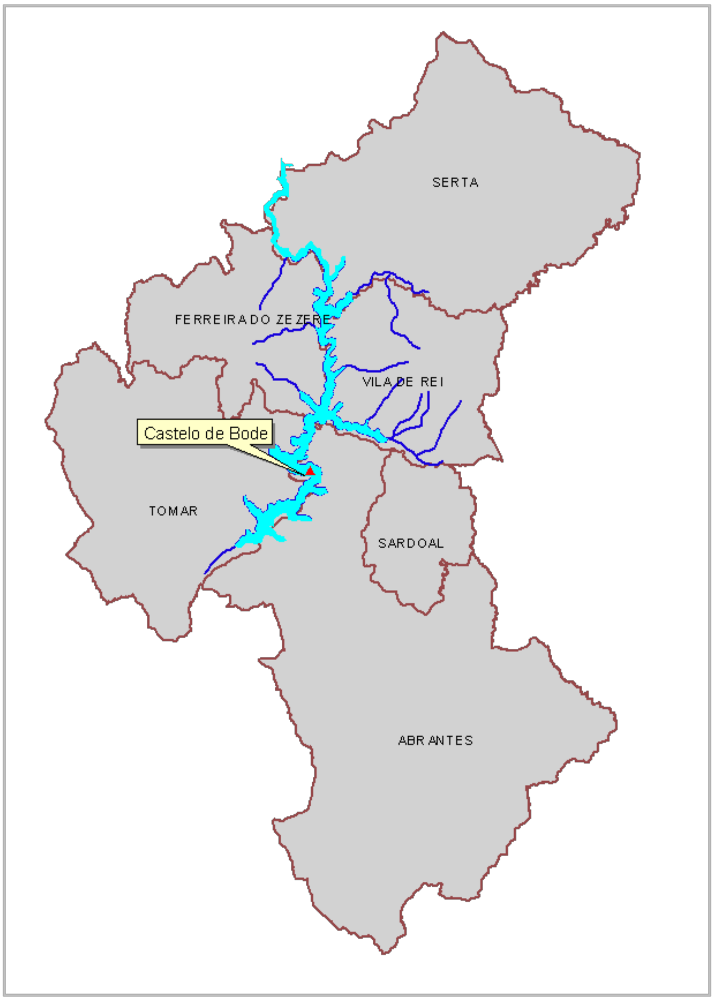

Figure 2.8: Example of polygon and point shape

a metadata record is a file of information, usually presented as an XML document, which captures the basic characteristics of a data or information resource. It represents who, what, when, where, why and how of the resource. Geospatial metadata are used to document geographic digital resources such as Geographic Information System (GIS) files, geospatial databases, and earth imagery. A geospatial metadata record includes core library catalog elements such as Title, Abstract, and Publication Data; geographic elements such as Geographic Extent and Projection Information; and database elements

> such as Attribute Label Definitions and Attribute Domain Values. The FGDC is tasked by Executive Order 12906 to develop procedures and assist in the implementation of a distributed discovery mechanism for national digital geospatial data. Geospatial metadata are critical to data discovery.

To make the metadata understanding for all has been created standards, have a well-known metadata standard make possible people from different communities, industries, and countries to understand the documentation because the standard acts as a dictionary, defining terminology and the expected values.

There are a number of metadata content standards to choose from. The FGDC Content Standard for Digital Geospatial Metadata (CSDGM) aims to provide a complete description of a data source. It is quite detailed, other states and regions have created their own metadata standards to try to simplify the information that should be recorded.

For example, the European Committee for Standardization (CEN)[1] and the Australia New Zealand Land Information Council (ANZLIC)[2] have created their own metadata standards and guidelines. The ISO standard 19115, Geographic Information Metadata, was designed for international use. It attempts to satisfy the requirements of all existing metadata standards. It allows for either general or detailed descriptions of data sources, makes some allowances for describing resources other than data, and has a small number of mandatory elements. While the 19115 standard has been finalized, work on the new ISO standard 19139, Geographic Information Metadata Implementation specification, is initiated; it will integrate content from several different ISO standards and produce an XML schema that specifies how to store ISO metadata in XML format. When ISO 19139 has been finalized, the FGDC will adopt a profile of that standard to replace the existing CSDGM; other regions of the world will also adopt profiles of ISO 19139 to replace their existing metadata standards.

[1] http://www.cen.eu/cenorm/homepage.htm
[2] http://www.anzlic.org.au/

2.4 Web data representation

With the raising of the Web with the Web 2.0 and the Semantic Web more standard to the representation of the data in the Web is developed. Several tools and languages have been proposed in the past for Ontologies and Semantic Web implementation such as RDF. This chapter presents some of these standards more pertinent to this master thesis.

Uniform Resource Identifier (URI)). To Berners-Lee (2005) URI provides a simple and extensible means for identifying a resource. The specification of URI syntax and semantics is derived from concepts introduced by the World Wide Web global information initiative, which's use of these identifiers, dates from 1990 and is described in "Universal Resource Identifiers in WWW"[1]. The syntax is designed to meet the recommendations laid out in "Functional Recommendations for Internet Resource Locators" and "Functional Requirements for Uniform Resource Names".

Resource Description Framework (RDF). RDF second Brickley & Guha (2004) is often called a language but according to Antoniou & Harmelen (2004) it is essentially a data-model to represent information on the Internet. RDF descriptions are models of archives; they are models or data sources, also known as metadata. This is a technology and recommended endorsed by W3C. Its main objectives are:

- Create a simple model of data, with formal semantics,
- To use the URI-based vocabulary and an XML-based syntax called RDF XML,
- To support the usage of XML.

RDF descriptions have three basic components (Lassila & Swick, 1998):

1. Resource or object,
2. Property or a relation between resources,

[1]http://www.ifla.org/

3. Statement or the properties of resources.

eXtensible Markup Language (XML). To W3C (2007) XML is a simple, very flexible text format derived from the Standard Generalized Markup Language (SGML, ISO 8879:1986)[1]. Both SGML and XML are *meta* languages, they are used for defining markup languages. A markup language defined using SGML or XML has a specific vocabulary (labels for elements and attributes) and a declared syntax (grammar defining the hierarchy and other features). Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere.

Geography Markup Language (GML) (ISO/TC211)[2] is a Language Encoding Standard to handle ontologies of geographical information, of the responsibility of OpenGis Consortium (OGC)[3]. Some of the implementation specification for GML that OGC developed although being based in XML does not follow RDF as OWL. Is an XML grammar for expressing geographic features.

GML serves as a modeling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. Like most XML based grammars, there are two parts to the grammar:

1. The schema, describes the document,

2. The instance document contains the actual data.

The objective of GML is to encode geographical information as well as their spatial and non spatial properties. This encoding is to be done in an open way, independently of the platform used as well as sharing the schema of organization of GI.

This language is one of the bases for the communication of GI. GML allows easy transmission of documents on the WWW, because there is a large pool of process in XML documents. IGML serves as a modeling language for geographic systems as well as an open interchange format for geographic transactions on

[1] http://xml.coverpages.org/
[2] http://www.iso.org
[3] http://www.opengeospatial.org/

the Internet. As with most XML based grammars, there are two parts to the grammar - the schema that describes the document and the instance document that contains the actual data. A GML document is described using a GML Schema. This allows users and developers to describe generic geographic data sets that contain points, lines and polygons. However, the developers of GML envision communities working to define community-specific application schemas are specialized extensions of GML. Using application schemas, users can refer to cities, rivers, and continents instead of points, lines and polygons. If everyone agrees to use the same schemas they can exchange data easily and be sure that a road is still a road when they view it, so the interoperability between systems turns effective.

2.5 Ontologies

Gruber (1993) defines an ontology *an explicit specification of a conceptualization.* Fensel (2000) details this definition, asserting that a conceptualization refers to an abstract model of some phenomenon in the world, which identifies relevant concepts from that phenomenon. A conceptualization explains the intended meaning of the terms used to indicate relevant relations (Guarino, 1992). The restriction to be explicit means that the concepts and relations between them are explicitly defined, that is, there is no ambiguity in the ontology. A geographic ontology is an ontology where the domain is the geographic domain.

Their growing importance arises from their role in:

- Formal modeling of specific information of domains,
- Understanding the content of databases,
- Information sharing, data interoperability.

They are essential for supporting the communication between systems of several types, such as cadastral engines or spatial search engines. A representation of the geographical information as an ontological representation must obey some rules and include some information. An ontology has multiple components (see Figure 2.9:

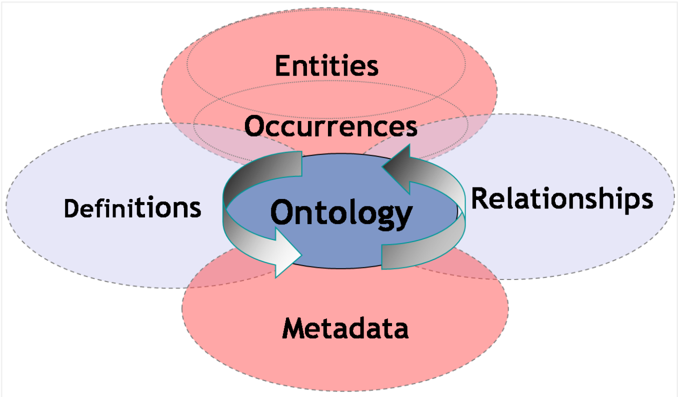

Figure 2.9: Components of an ontology

1. Entities or Occurrences,
2. Relationships between Entities and Occurrences,
3. Definitions of Entities,
4. Metadata of Entities.

Only with all of these components one representation could be ontological.

The similarities among the controlled vocabularies, thesauri, taxonomies, meta-models and ontologies are (Garshol, 2004):

1. All approaches help structure, classify, model, and or represent the concepts and relationships pertaining to some subject matter of interest to some community,
2. They are intended to enable a community to come to agreement and to commit to use the same terms in the same way,
3. There is a set of terms that some community agrees to use to refer to these concepts and relationships,
4. The meaning of the terms is specified in some way and to some degree,
5. They are fuzzy, ill-defined notions used in many different ways by different individuals and communities.

On the other hand according the same author, the major differences that distinguish these approaches are:

1. How much meaning is specified for each term,
2. On the notation or language used to specify the meaning, and
3. Taxonomies, thesauri, ontologies have different but overlapping uses.

2.5.1 Ontlogies representation

Several tools and languages have been proposed in the past for Ontologies such OWL.

Ontology Web Language (OWL) (W3C, 2001) is a language for ontologies standardized for the Semantic Web and with a well defined structure. OWL adds more vocabulary to RDF to describe properties and classes, such as relationships among classes. It allows a richer description of the properties, characteristic of the properties, and enumeration of classes. OWL is a structured language in a way to be compatible with the design of World Wide Web. OWL uses URIs; and RDF, adding the following potentialities to the ontologies:

- Capacity of being distributed through many systems,
- Scalability to the needs of the Internet,
- Compatibility with Internet standards for accessibility and internationalization,
- Openness and extensibility.

According Antoniou & Harmelen (2004) a language based ontology needs to have some requirements:

- A well-defined syntax,
- A formal semantics,
- A convenience of expression,

- Efficient reasoning support,
- Sufficient expressive power.

2.5.2 Ontlogies construction

The construction of an ontology, whether geographic or not, proceeds in several phases. To Antoniou & Harmelen (2004) it is possible to distinguish eight main stages in the development process:

1. Determine the scope or domain,
2. Consider reusing,
3. Enumerate important terms of the ontology,
4. Define taxonomy or classes and class hierarchy,
5. Define properties of the classes,
6. Define facets or the internal structure of concepts,
7. Define instances,
8. Check for anomalies.

The first tree steps are defined for the purpose of the producer of the ontology. For developing class hierarchies there are several approaches. Independently of the used approach, we usually start by defining classes. From the list of the more important terms for the ontology we select the terms these describing the objects having independent existence rather than terms that describes these objects. These terms will be classes in the ontology. For each property in the list we must determine which classes it describes (Uschold & Gruninger, 1996).

After the first five stages, the ontology will only require the concepts provided by RDF Schema and do not use any of the additional primitive in OWL (Antoniou & Harmelen, 2004). This will change in step six, which includes enriching the previously defined properties with facets:

- Cardinality,

- Required values,
- Relational characteristics.

Phase seven is the population of the ontology.

2.6 Projects

There are several applications of geographic ontologies, such as object catalogues for cartography. The catalogue of objects from the Digital Geographic Information Exchange Standard (DIGE)[1] was prepared by and issued under the authority of Digital the Geographic Information Working Group (DGIWG), as well as the proposed South African spatial feature standard catalogue (Maphumulo, 2003).

Semantic Web search engines are another application of the ontological representation of physical geography. The SPIRIT project was a research project with the objective of creating a search engine which can obtain geographical information included in a "query". The main goal was the development of a geographic information retrieval system, including the creation of an interface for mobile devices. The project developed tools and software that can be used to create search engines and web sites that recognize geographic terminology. In order to demonstrate and evaluate the project results, a prototype of a spatial search engine was built and used as a platform to test and evaluate new techniques in geographic data retrieval.

A similar project to SPIRIT is the project Geographic Reasoning for Search Engines (GREASE) from Faculty of Sciences, University of Lisbon. GREASE [2] researches methods; algorithms and software architecture for a system for helping users find Web pages with a scope matching their current location from hand-held computers. GREASE designates as "location" the meta-data or knowledge about the geographic position and as "scope" the manually or automatically generated meta-data describing the area of interest of web pages. The project develops a location-aware web search engine that matches against users' queries by joining"

[1]https://www.dgiwg.org/digest/
[2]http://xldb.fc.ul.pt/

the pages' information scopes and location data obtained from positioning devices connected to their hand-held computers.

Another project related with ontologies and geographic information is the 'Ontology' Geonames, although the Geonames be regarded by many as a geographic ontology, it is not. As already pointed out in the previous background chapter, an ontology integrates in its definition the relationships between entities and/or occurrences, Figure 2.9. In Geonames this relationships are implicit and not explicit. We could considerer the Geonames as a good search tool for spatial data, but not as a ontological representation of the geographic information or a Geo-Ontology.

2.7 Geographic Knowledge Base (GKB)

One example of an environment for developing Geo-Ontologies is the Geographic Knowledge Base (GKB). GKB was created by the project GREASE.

The main purpose of GKB is to provide a common place for integrating data from multiple sources under a common schema, supporting mechanisms for storing, maintaining, and exporting the assembled knowledge about geographic entities and Web resources.

GKB is an environment for integrating geographic data and generating Geo-Ontologies. It is implemented based on a model able to support the storage of the information from administrative and physical domains (Chaves, 2005).

The first version of GKB (GKB 1.0, henceforth) contains two instances: one stores data about the geo-administrative and network (web) domains of Portugal and the other contains data from all over the world. The information stored in GKB 1.0 is exported to ontologies in the OWL format.

One of these ontologies (Geo-Net-PT01) contains the integrated geographic administrative information of Portugal and is freely available from XLDB page[1]. Applications using semantic web technologies can easily work with the ontologies exported from GKB.

[1]http://xldb.fc.ul.pt/wiki/Geo-Net-PT02

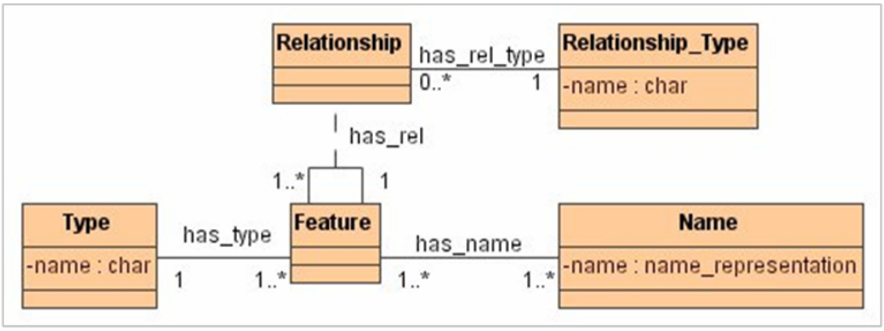

Figure 2.10: GKB Meta-model

GKB 1.0 is implemented on top of a MySQL relational database. The information is from various sources of information. In addition to the database, GKB has two sets of tools:

- Converters, which load data from the various source formats into GKB and perform data normalization in order to maintain a single unified view of all the information,
- Generators, which output the GKB contents as ontologies.

GKB Metamodel

A model is an abstraction of phenomena in the real world, and a metamodel is another level of abstraction, highlighting properties of the model itself. In GKB the metamodel describes the classes and relationships that are used to modeling geographic information. The GKB metamodel supports representations of multiple information domains related to geography, such as the administrative and physical domains. In each of these domains, the information is organized using concepts from the metamodel (e.g. features, feature types and their relationships).

So, metamodel is defined as the representation of the model, which is the conceptualization of entities and relationships between them.

This metamodel allows the representation of spatial and non spatial data. The class model of the metamodel has three main classes (see Figure 2.10).

1. Type or class,

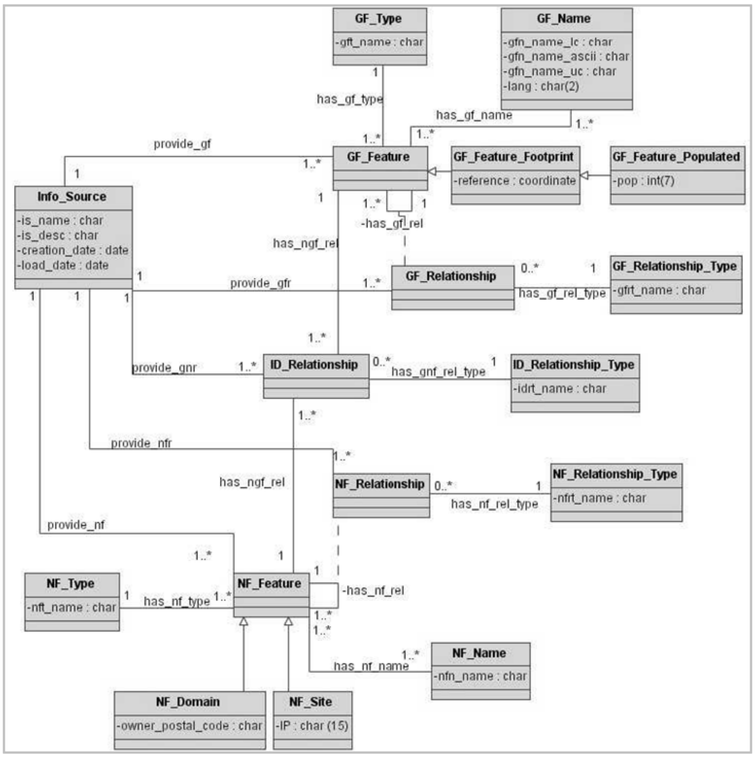

Figure 2.11: Complet Metamodel for the first version of Geo-Net

2. Feature or instance,

3. Name, which stores the name of the instances.

Additional meta-classes support Relationships between Types, Attributes of Types, Coordinates of Features, Info-Sources and Relationship-Types.

Network domain

To make possible the representation of entities from the network domain, the metamodel of GKB included two more meta-classes:

1. Site,

2. Domain.

Figure 2.11 shows these two meta-classes and their relationships with the other Domains of Geo-Ontologies.

MetaModel for all Domains

Metamodel construction implies the analysis, construction and development of rules, constraints, models and theories applicable and useful for modeling. This metamodel is the design and formalization of the database to integrate the conceptual model. Due to the nature of the data to represent in the Geo-Ontology, this metamodel must have specific characteristics to allow the representation of spatial and non spatial data.

To create this Geo-Ontology spanning multiple domains also requires the representation of the relationships among domains; Figure 2.11 shows these domains and relationships. Only one component meta-class is shared; Info-Source (both physical and administrative domains) which is used to track the provenance of each stored feature. The domain of each meta-class is encoded in the first two characters of meta-class name. To recognize the Domain the two first letters show the Domain:

- GF; for Geographic Features,
- NF; for Network Features.

2.8 Summary

Today the market has a large number of tolls to represent Geographic Information. The GI could be stored under a simple Database or under a *Thesaurus*. One of the more well-know toll was de GIS (geographic information system). However the most number of these do not has interoperability has a priority, so GI turns static information. GIS is used by a great number of areas of knowledge like geography or telecommunications,but the interoperability of information is not guaranteed, as so the relationships between types and or instances is not mandatory or explicit. We can consider an ontological representation like a development of these forms of representation.

Chapter 3

Modeling the Geo-Physical Domain

This chapter details the proposed model for geographic knowledge and the creation of the new version of Geo-Net-PT, ontology which extends it with geo-physical domain. The chapter presents the extensions to the meta-model developed for GKB 2.0 and the model of the physical domain including newly defined relationships and relationship-types.

3.1 GKB 2.0 Meta-model

Figure 3.1 shows the domains represented in Geo-Net-PT02 as implemented in terms of a common meta-model that was designed to enable the integration of data from the three domains of knowledge:

1. Geo-administrative domain (geographic),
2. Geo-physical domain (geographic),
3. Network domain (www).

Figure 3.2 represents a class diagram common to these domains of knowledge, including the geo-physical domain presented in this thesis. The tree main classes in the GKB meta-model are:

1. *Type* or *Class*, which defines geographic concepts of features-types
2. *Feature* or *Occurrency*, which defines instances of concepts

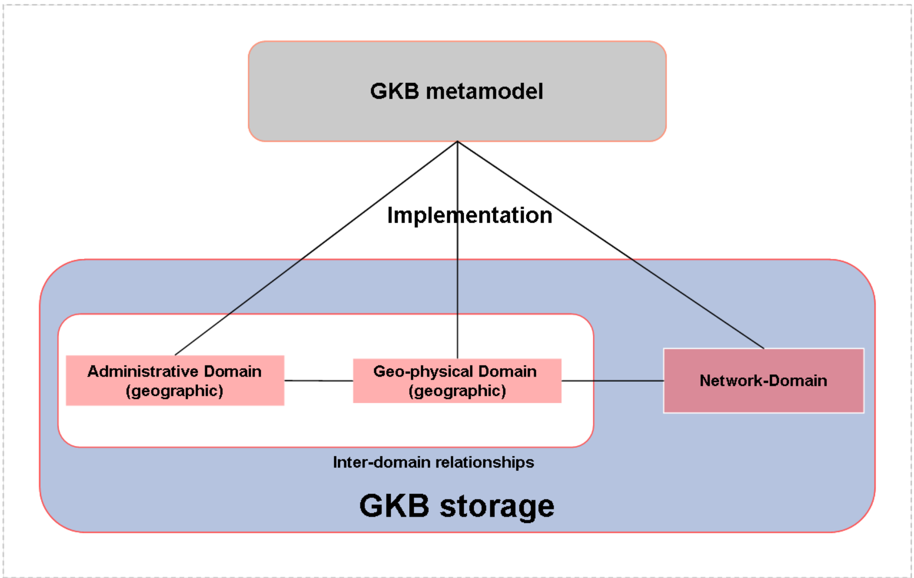

Figure 3.1: Geo-Net-PT02 Domains

3. *Name*, which defines the names of the occurrencies.

Table 3.1 shows some examples of types and the 3 letter identifiers that have been assigned to them in GKB (see Appendix B for the complete list). The class *Features*is the main class which represents the real objects. Another main class is *Name*, which holds the names of *Features*. Instances of identical names are shared by *Features*. For instance the name "Caramulo" is a mountain range type and water source type.

The class *Name* has additional attributes besides the name representation to support multilingual designations:

- Language code, and
- Country code.

"Tagus River" could be represented in an Iberian Geo-Ontology as a single feature. With these two characteristics we can represent all the languages. The code for Portuguese is POR (ISO 639-2) or PT (639-1); and the Tagus River is represented in Portuguese as *Tejo*, while Spanish codes are SPA (ISO 639-2) or ES (639-1) and the same river is represent as Río *Tajo*.

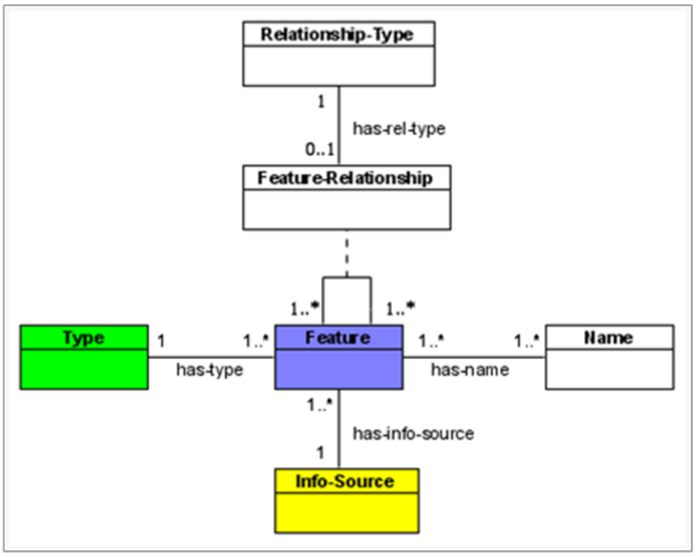

Figure 3.2: meta-model of the GF defined to GKB1.0

The meta-model of GKB 1.0 presented in (Chaves, 2008) is based mainly in the relationships between geographic features (GF) and did not include types relationships; in the new meta-model relationships between the Types are also represented. For instance in GKB 2.0 it is possible to store that rivers flow to seas or other river. Figure 3.3 represents how the relationships among types and features are modeled.

Relationship-type is a class that defines the supported relationships, such as "‘part-of"’. Relationship-types have attributes that enable agents to make inferences and be a knowledge repository. There are four Boolean attributes:

1. Transitive; if "A" is related to "B", and "B" is related to "C", then if the relationship is transitive, "A" is related to "C",

2. Commutative; if "A" is related to "B" and "‘B"’ is related to "‘A"’, then if the relationship is commutative is given the relationship for all "‘A"’ and "‘B"’,

3. Symmetric; if "A" is related to "B", if the relationship is symmetric, then "B" is related to "A",

PHY-ID	Type
ADG	Adega reconhecida
ADS	Açude
APR	Área protegida
AQU	Aquífero
BHD	Bacia hidrográfica
BIO	Biótopo
DUN	Dunas
FER	Ferrovia
GRT	Grutas
LAG	Lagoa
NAS	Nascente
PAU	Paúl
PRC	Classes de precipitação
RIO	Rio
TEP	Temperatura
RCT	Recurso turistico

Table 3.1: Types examples

4. Reflexive; if "‘A"’ is related with "‘A"’ holds.

For example if defined that the *Tagus* intercepts the *Natural Reserve of the Tagus estuary* GKB clients can infer that the *Natural Reserve of the Tagus estuary* is intercepted by the *River Tagus* and that if a River is contained in a *Continent*, the *Tagus River* is also in a *Continent*.

Through footprints association information is possible find which features are within 10m, or North, or the closest feature. Figure 3.4 represent the complete meta-model, this figure shows the association of footprints function performed by the footprint and its metadata that turns possible interoperability of the data represented.

GKB 2.0 meta-model also allows the specification of preferred names, alternative names and historic ones, among others. Due to the nature of the hydrologic data, it is also necessary to support evolutions. Evolution is often done based on historic texts in which the local name of a flood has changed. For instance, instead of Lisbon, Olissypo may be found in the text. The meta-model allows Olissypo to be represented as an historic he name of Lisbon.

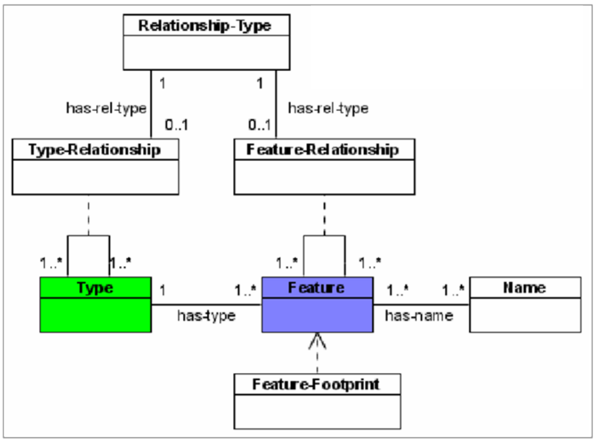

Figure 3.3: GKB 2.0 meta-model

3.2 Relationships

Spatial relations may include aspects purely mathematical, cognitive, linguistic and psychological considerations. According Egenhofer (1989) the exploration of spatial relationships is a multi-disciplinary effort involving engineering, cartography, computer science and other fields of researchers.

To define the relationships necessarily to formalize a Geo-Ontology it is necessary use a theory about spatial relations applicable to real-world problems. To Abler (1987) refers that the one of the goals to the National Center of for Geographic Information and Analysis is the development of a coherent, mathematical theory of spatial relationships.

Spatial relations could be grouped into three categories:

1. Topological relations, which are under topological transformations of the reference objects (Egenhofer, 1989)

2. Metric relations in terms of distance and directions (Peuquet & Zhang, 1987)

3. Relations concerning the partial and total order of spatial objects (Kainz, 1990).

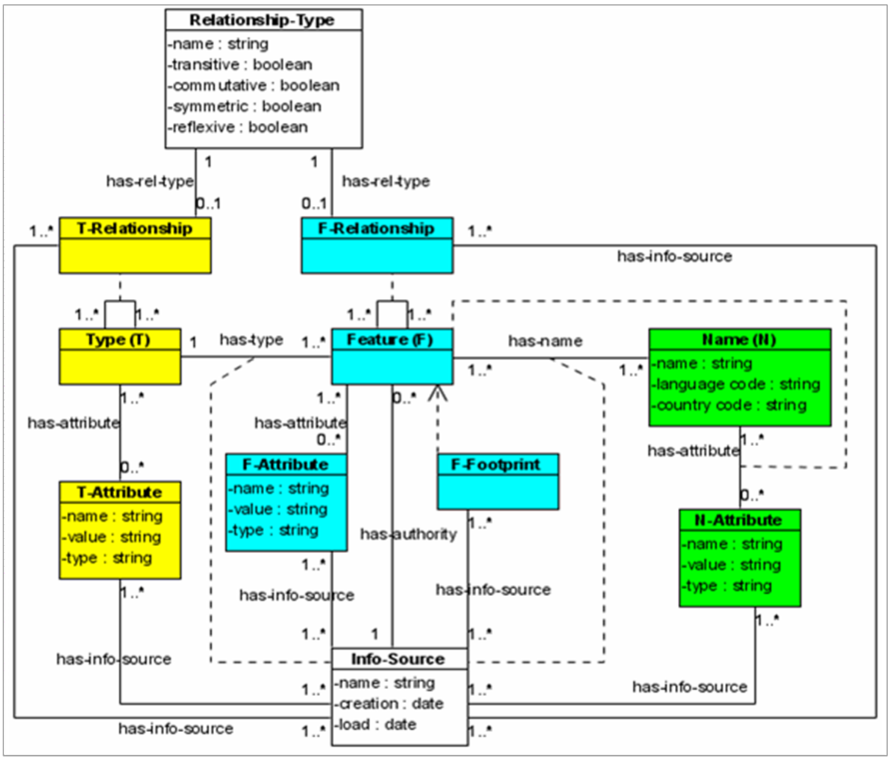

Figure 3.4: Complete GKB 2.0 Meta-model

The definitions of these relations could be in terms of set operations to describe the topological relations, Appendix D - Topology rules shows topologycal rules difined by ESRI.

Figure 4.4 shows eight defined examples of topological relationships; same of these could be considered a variant e.g. contain and contained.

The relationships represented in this Geo-Ontology are of two types:

- Explicit relationships,
- Implicit relationships.

Explicit relationships are formalized in the model; implicit relationships could only be inferred with spatial queries to the GKB. The explicit relationships (formalized) defined in GKB 2.0 are:

Contain/contained:The relationships contain/contained is a symmetric relationship, so it could be of two types:

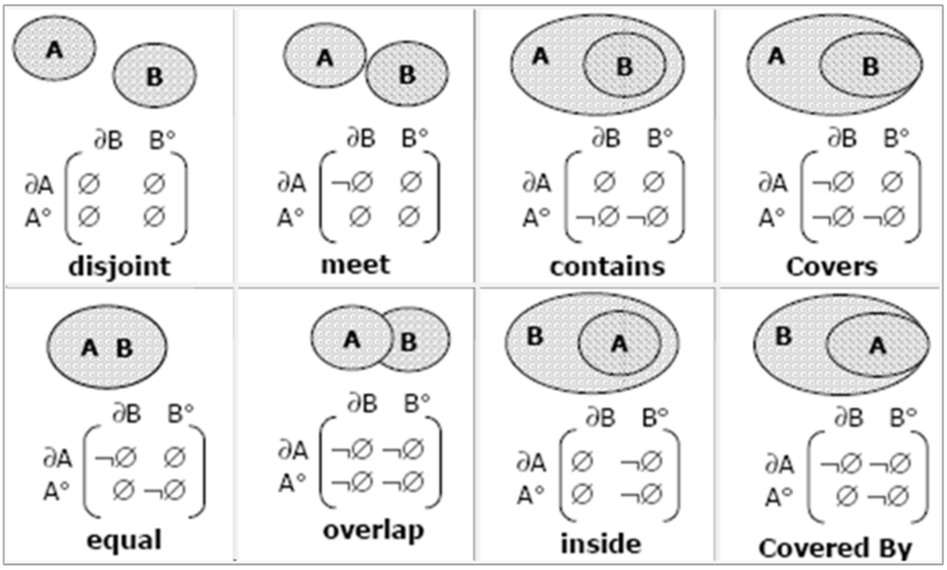

Figure 3.5: Topological relationships

- Completely contained (within); selects features or types if they fall completely within an other feature or type,

- Completely contains; selects features or types that completely contain one or more feature or type.

The **Part-of relationship**: could be inferred thought the relationships contain/contained, as shown in Figure 3.6, if *Rio Tejo* is completely contained by type Bacia hidrográfica do Tejo, so the feature *Bacia hidrográfica do Tejo* completely contain feature *Rio Tejo*.

Intersects: Selects features and types that intercept other features and or types. Intersection implies that at least one point is common to both the selector and the target or one of them is completely within the other, buy not completely contained.

Figure 3.7 shows an example; Type *Rede ferroviária* intercept Type *Bacia hidrográfica*.

Adjacent: This relationship describes features contiguous, adjoining or abutting to other features. In geometry, adjacent sides of a polygon join at their vertex. The administrative types are always adjacent between the same types of

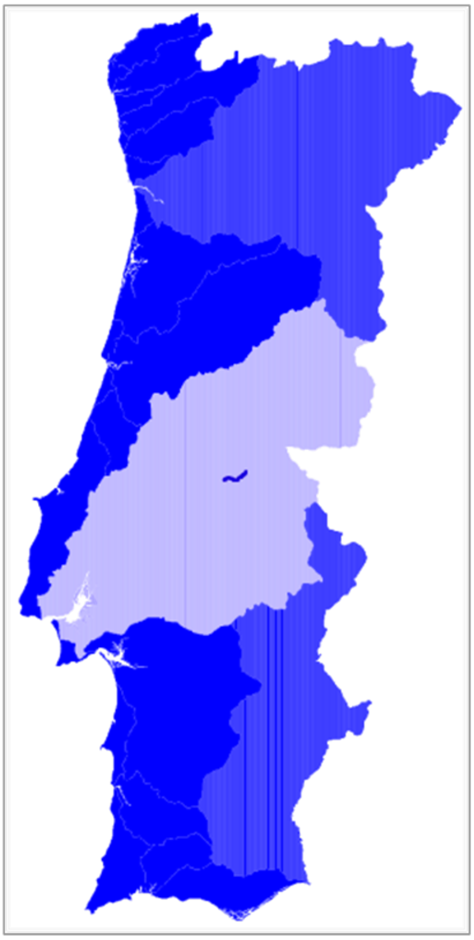

Figure 3.6: Feature *Bacia hidrográfica do Tejo* completely contains Feature *Albufeira de Belver*

features. In physical domain this type of relationships is uncommon. Features of this domain do not have precise borders, e.g. the area of a lake changes with the seasons.

That relationships is illustrate in Figure 3.7, representing hydrographic basins of Portugal, all of which are adjacent to other basins.

With explicit relationships is possible to represent and formalize all of the relationships between the types represent in the Geo-Net-PT02; some other relationships like part-of could be inference with these types of relationships.

Implicit relationships include metric relationships (Peuquet & Zhang, 1987), such as:

- Are Within Distance Of; a selected feature are within a specified distance of others features,
- Proximity; a larger and shorter distance
- Calculus of routes (the path closest, cheaper, or faster)
- N/S/E or W, the geographical directions
- Local Center

3.3 Summary

With the relationships presented and modeled in this version of GKB is possible the representation of almost the relationships between geographic shapes. With the proposed model is not necessarily the segmentation of the geographic information, all the information could be modeled under the same meta-model.

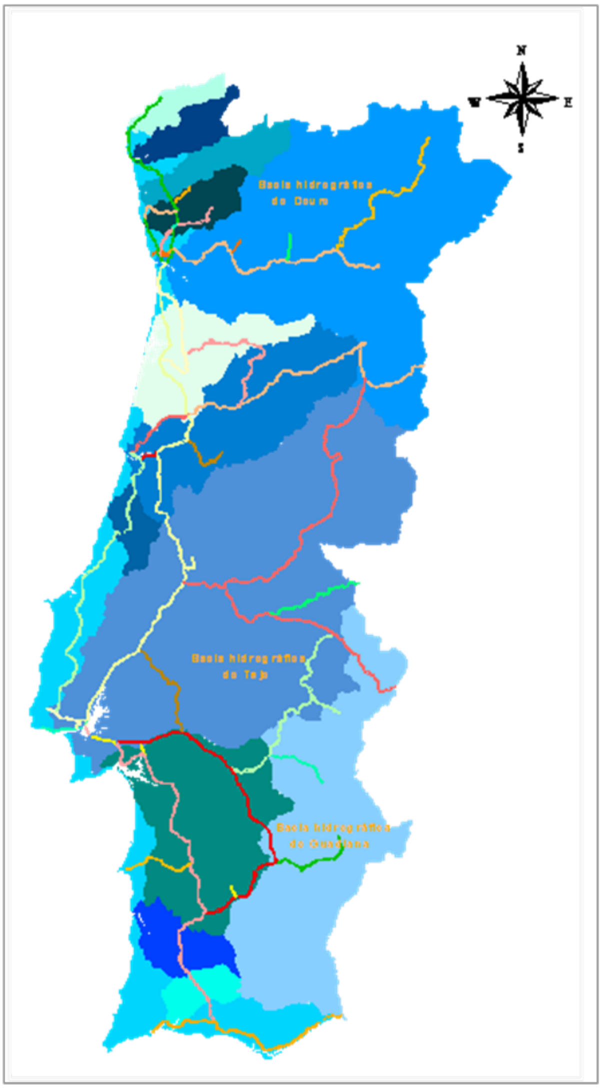

Figure 3.7: Examples of relationships between Types

Chapter 4

Physical Ontology of Portugal

This chapter characterizes the production process, the data and the information sources used for the generation of the physical domain of Geo-Net-PT02.

Geo-Net-PT02 is a free Geo-Ontology of Portugal, which adds among other features, data from the domain of physical geography to its previous version, Geo-Net-PT01. Geo-Net-PT02 adopts the formats, rules and guidelines for representing information on the Semantic Web.

This new version adds more than 24 000 features from the physical domain of Portugal (Continent and Islands).

Geo-Net-PT02 includes not only relationships between features from the geophysical domain, but also inter-domain-relationships spanning features from the administrative domain and the physical domain, e.g. *the Tagus River is adjacent to Lisbon (the city).*

This Geo-Ontology was generated following the tasks presented in the methodology proposed in this dissertation. Figure 4.1 ilustrates these processes.

This chapter is organized as follows:

1. Identification of information and their characterization,

2. Cleaning and normalization of the collected data,

3. Integration the data into the new GBK meta-model,

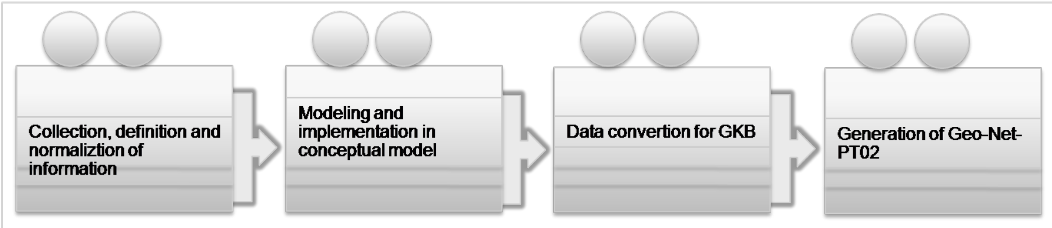

Figure 4.1: Geo-Net-PT02 generation process

4. Modelation of the collect and normalized data into ontological model,

5. Characterization of the produced Geo-Ontology.

Figure 4.1 presents Geo-Net-PT02 generation process, or the 'Making-of' process:

- Collection and treatment of information,
- Modelation of the conceptual model and integration of the collected data,
- Data conversion for GKB 2.0, and
- Generation of Geo-Net-PT02.

To perform all of the main process presented in previous figure, is necessarily perform others sub-tasks, Figure 4.2 shows these sub-tasks.

1. Research,
2. Staging area,
3. Geographic Knowledge Base (GKB).

4.1 Information sources

One of the initial objectives of this thesis was to identify producers of geographic information (GI) about Portugal of the physical domain and the collection of their information.

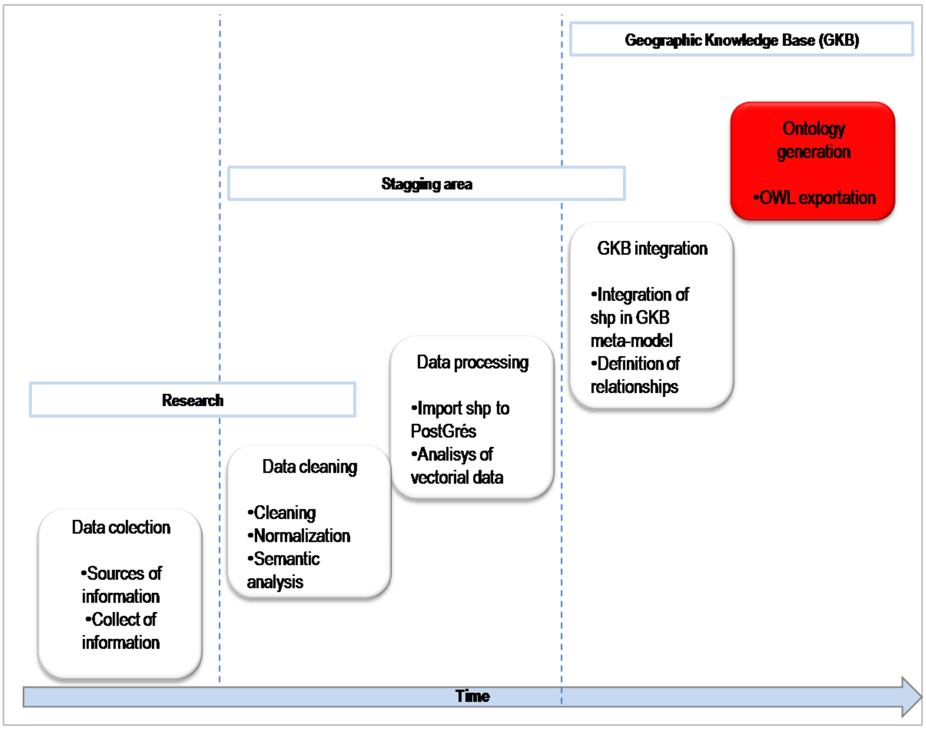

Figure 4.2: 'Making-of' sub-tasks

There are multiple sources of GI in Portugal (see Appendix D - Fonts of Information) and their scope and the quality of the information available from these providers is very heterogeneous. For this work I only selected official sources with expertise in the production of geographic information and with public data, as we intended to make Geo-Net-PT02 public. Others sources of information, and other information has been collected, but this work do not intend to integrate all information from all of sources, but only the information that have a great number of users. Other GI has not been integrated; e.g. Zonas de Litologia.

There are two types of sources of geographic information:

1. Aggregators of available data,

2. Producers.

Instituto Geográfico Português (IGP)[1] a division of the Ministério do

[1]http://www.igeo.pt/

Ambiente, do Ordenamento do Território e do Desenvolvimento Regional (MAOTDR), has the status of National Mapping Authority; is the responsible for implementing the policy of geographical information. IGP has all the rights, obligations and duties from the old Centro Nacional de Informaç ão Geográfica (CNIG) and Instituto Português de Cartografia e Cadastro.

The Sistema Nacional de Informação Geográfica (SNIG)[1], integrated in IGP, is the infrastructure of national spatial data and is intended to provide, from several points of access, the ability to search, view and explore the geographical information on the national territory. It articulates and organizes activities related GI in Portugal and in the context of the European directive INSPIRE (INfrastructure for SPatial InfoRmation in Europe)[2]. Beside the dissemination of the geo-information, SNIG has congregates all under the same infrastructure, many institutes and municipalities already collaborate with SNIG (see Appendix D - Fonts of Information).

The list of producers of information and a brief description of their information is given below.

- **Instituto Geográfico Português(IGeoE)**[3] produces all the geographic information (cartographic or not) necessary to the military forces, as so the production of information of context not only military.
- **Instituto da Água (INAG)**[4] as National Water Authority, is intended to propose, monitor and ensure implementation of national policy in the field of water resources to ensure sustainable management and the effective implementation of the water law; the Sistema Nacional de Informação de Recursos Hídricos (SNIRH)[5].

[1] http:http://snig.igeo.pt/Portal/
[2] http://inspire.jrc.ec.europa.eu/
[3] http://www.igeoe.pt/
[4] http://www.inag.pt/
[5] http://snirh.pt/

- **Instituto Hidrográfico(IH**[1] is a laboratory of the state, under the supervision of the Portuguese Navy that is dedicated to science and technology of the Sea. The task is to ensure the research and development activities related to science and techniques of the sea, taking into view the application priority in the military, and contribute, as the rule of reference laboratory at national and international level devoted to research of the oceans, for the development of the country and for the protection of the marine environment.
- **Agência do Ambiente** is an agency of the Ministério do Ambiente Ordenamento do Território e do Desenvolvimento Regional.
- The **Instituto Nacional de Engenharia,Tecnologia e Inovação (INETI)**[2] has , as a department responsible for the management and provision of geocientific information, has been to integrate and manage the large number of institutional information, from the normal activity of the institution, with a view to providing a coherent and representative. e-Geo - National System of Geocientific Information, can make available online, a number of expressive content geo-scientific, mostly derived from analog files. Conduct systematic geological mapping, inventory and enhance the national geological resources, delineate the areas of geological risk and develop information systems geological
- **Turismo de Portugal** [3] through Minisério da Ecónomia e Inova cão (MEI) has the task of design, implementation and evaluation of policies aimed at economic activities, including production of services, energy, trade, tourism, consumer protection, policies to regulate the markets and promotion of investment, as well as horizontal policies aimed at competitiveness and innovation aimed at internationalization of Portuguese companies. The information from Turismo de Portugal is about 2800 instances from 20 Types (see Table 4.1).

[1] http://www.hidrografico.pt/
[2] http://e-geo.ineti.pt/
[3] http://www.turismodeportugal.pt/

Types	**Number-of-features**
Aldeias Preservadas	217
Áreas de Paisagem Protegida	18
Casinos	8
Castelos	263
Catedrais Basílicas	23
Caves	310
Estações Arqueológicas	456
Estâncias Termais	35
Campo de Golf	65
Jardins Botânicos	7
Kartódromos	30
Marinas	26
Monumentos Naturais	5
Museus	524
Parque Nacional	1
Parques Naturais	12
Património Mundial	11
Praias	591
Reservas Naturais	9
Agro-turismo	142
Total features	2753

Table 4.1: Features Types provided by Turismo de Portugal

Types	Number-of-features
Adega Reconhecida	311
Albufeira(ponto)	1880
Aldeia Preservada	217
Aquifero	125
Área de Paisagem Protegida	18
Area Protegida	49
Bacia de Escoamento	209
Bacia Hidrografica	129
Biótopo	120
Campo de Golf	65
Casino	8
Edifício Notável	3877
Especie Notável	2432
Estabelecimento de Dormida	2847
Estação Arqueológica	456
Estância Termal	35
Ferrovias	103
Fortificação	263
Gruta	7
Jardim Botânico	7
Kartódromo	30
Lagoa	10
Local de Culto	23

Table 4.2: Extract of integrated features

4.1.1 Integrated data

This new version adds more about 24 000 features from the physical domain; Table A.2 shows some examples of the features types integrated, all of these features types are presented in Appendix A - Count of collected data.

4.1.2 Other information

Others sources of information, and other information has been collected, butt this work dont pretend integrated all of the disponible information from all of sources, but only the information that have a great number of users. Other GI that as not been integrated is for example:

- Geodesic vertices from IGP,
- ...
- ...

Other information available from this data source has been not integrated in Geo-Net-PT02, because it was not considered relevant, including tourist apartments and shopping centers.

4.2 Data cleaning

The information from the different sources is not normalized, they is not in a comprehensive format. To improve the quality of information some process has been preformed before the integration of information in GKB:

- Analysis of information,
- Cleaning of repeated information from different sources,
- Normalization the format, always with the name of Type before the Name of instance and with the words of connection,
- Collection of alternatives names.

All of the data collect to be integrated in GBK 2.0 is first cleaned and normalized in Staging area.These phases consists in several actions:

1. Cleaning of misspellings in names,
2. Normalization, all data is formatted by the same rules, all features have associated connection words defined such as e.g. *Serra **da** Estrela.*

3. Semantic analysis, what is a Type or a Name feature; e.g. to *Serra do Caramulo*, *Serra* is Type and *Caramulo* Name, so *Serra do Caramulo* is the Feature; but to *Serra da Estrela Serra* is the Type as so the Name, so in this ontology is stored as Type *Serra* and Name *Serra da Estrela*, so the Feature name is coded as *Serra Serra da Estrela*. In this phase is performed to, the search of alternative names.

The firts and second phases is manually performed in a spreadsheet editor, Table 4.3 shows the ID-Type and number of cleaned features, after the exportation of the data from the shp. After the cleaning and normalization of the alphanumeric data, all alphanumeric data(initial and alternative names and types) are integrated in the vectorial data. And now I have a shp of the vectorial information with the correct and complete alphanumeric data. For example:

> DESIGNACAO - Albufeira Idanha;ID-TYPE - ALB; NOME CORRIGIDO - Albufeira de Idanha;NOMES ALTERNATIVOS - Albufeira de Idanha-a-Nova; Barragem de Idanha; Barragem de Idanha-a-Nova.

The third phase of the process of the improvement of the data quality of Geo-Net-PT2 has been the analysis of the semantic of the name of instances, e.g. the *Type* of *Serra da Estrela* is mountain and the name of *Instance* is *Serra da Estrela* but the type of *Serra do Caramulo* is *Serra* and the name of instance *Caramulo*.

4.3 Data transformation for GKB

After the implementation of meta-model and geographic data (shp format) import, has been performed the integration of this data into GKB meta-model. This phase is performed in the Staging area (see Figure 4.2).

The conversion of the data from the fonts of information, after treatment to normalization, was made trough SQL commands.

To resume before the integration this data in GKB has been performed in some previous phases. Figure 4.3 shows the main phases performed before the integration of the data in GKB.

ID-Types	**Number-of-features**
ALB	418
ANT	2432
APF	19
APR	44
BAR	306
BHD	96
BIO	120
BRR	2
ETU	98
FER	103
FRM	105
FRT	82
GRT	7
LAG	9
NAS	220
PCP	150
PIT	2611
PRC	254
RAM	10
RCT	4889
RGN	437
RHD	129
RIB	3171
RIO	3038
RRO	122
SIS	121
TEP	177
UNP	551
VAL	117
TOTAL	19838

Table 4.3: Cleaned features

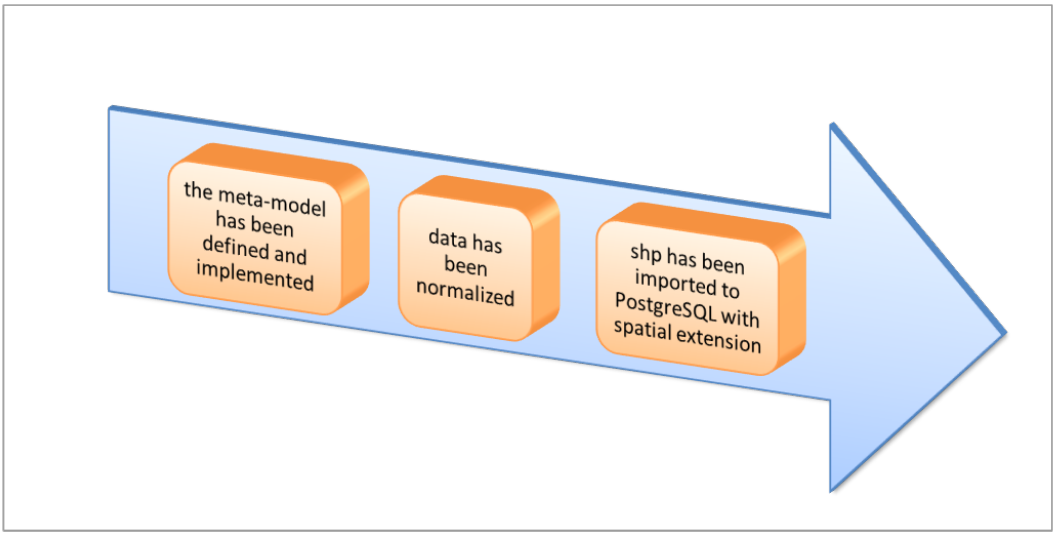

Figure 4.3: Previous phases

So, the last phase before the Geo-Net-PT02 generation has been the integration of variously components in GKB; Figure 4.2 shows the components that integrate the process to generate the Geo-Ontology.

4.4 Implementation

The logical model produces a logical database model that translates the conceptual model into a fully descriptive logical data model. As a result, whereas some join tables did not appear in the conceptual model the logical model will include all tables. The logical data model represents the realistic and normalized design of the database that has been implemented. Geo-Net-PT02 has been implemented in PostgreSQL[1], with the PostGis[2] extension, total free technology.

The physical meta-model has produce the physical data model which is the database design where is implemented. In practice this is an implemented schema which is an empty set of tables linked by relationships that are ready to receive data. The conceptual model is always prepared according to a standard formalism

[1] http://postgresql.org/
[2] http://postgis.refractions.net/

which clearly indicates how tables and their relationships have to be presented and documented. The formalization of the conceptual model has been made in SQL language.

The first step in this phases has been the importation of the Geo-Net-PT01 to this new meta-model, and to this new technology (Geo-Net-PT01 is under MySQL), to turn this possible the encoding from Geo-Net-PT01 was amended from Latin1 to UTF8.

The next step has been importation of the geo-information from shp (a vectorial format with alphanumeric data; ESRI)[1] to PostGis.

In this phase, after, the import of the shp to PostGrés as been performed a vectorial data analysis; e.g. if the same feature is in more than one source of data they are dissolved in only one. In this phase is to performed a conversion of the all the data to ETRS89.

4.5 Characterization of the produced ontology

The result of this master thesis is variously. The main result has been the generation of Geo-Net-PT02 a new Geo-Ontology, where is made the characterization of the main physical geography information requirements in web applications. This new ontology is an owl file, see Figure 4.4.

This Geo-Ontology could be centered in Types with their relationships as so in a Name centered. In a name centered the relationships formalized is between features.

Geo-Net-PT02 is a new totally free Geo-Ontology from XLDB Team of LASIGE (Large-Scale Informatics Systems Laboratory) with a focus on the domain of physical geography formatted to the Semantic Web.

Geo-Net-PT02 has integrated the previous version of the public ontology of Portugal (Geo-Net-PT01) to make a Geo-Ontology most comprehensive. In this way this new Geo-Ontology has all of the geo-domains represented and integrated. This new version has more than 24 000 features, from the geo physical domain, as so the features from the geo-administrative domain.

[1] http://www.esri.com/

```
<gn:Geo_Feature rdf:ID="GEO_PHY_145">
 <gn:names>
  <rdf:Bag>
   <rdf:li gn:name="Douro" xml:lang="PT-PT" gn:att="P" gn:is="IGeoE"/>
   <rdf:li gn:name="Duero" xml:lang="ES-ES" gn:att="A" gn:is="IGP"/>
  </rdf:Bag>
 </gn:names>

<gn:geo_type_id gn:is="IGEO" rdf:resource="#Rio"/>
<rdfs:comment>Serra de Urbião - Spain</rdfs:comment>
<gn:source_river gn:is="IGEO" rdf:resource="#GEO_PHY_120"/>
<rdfs:comment>Porto - Portugal</rdfs:comment>
<gn:outlet_river gn:is="IGEO" rdf:resource="#GEO_ADM_238"/>
<gn:tributary gn:is="IGEO">
  <rdf:Bag>
    <rdf:li rdf:resource="#GEO_PHY_400"/>
    <rdf:li rdf:resource="#GEO_PHY_401"/>
  </rdf:Bag>
</gn:tributary>
<gn:length unit="km" gn:is="IGEO">850</gn:length>
<rdf:Bag><rdf:li>
    <gn:Geo_Relationship>
      <gn:rel_type_id rdf:resource="#INTERSECTS"/>
      <gn:geo_id><rdf:Bag>
          <rdf:li rdf:resource="#GEO_ADM_238"/>
          <rdf:li rdf:resource="#GEO_PHY_300"/>
        </rdf:Bag></gn:geo_id>
    </gn:Geo_Relationship>
  </rdf:li><rdf:li>
    <gn:Geo_Relationship>
      <gn:rel_type_id rdf:resource="#ADJ"/>
      <gn:geo_id rdf:resource="#GEO_PHY_198"/>
    </gn:Geo_Relationship>
  </rdf:li><rdf:li>
    <gn:Geo_Relationship>
      <gn:rel_type_id rdf:resource="#PRT"/>
      <gn:geo_id rdf:resource="#GEO_PHY_100"/>
    </gn:Geo_Relationship>
</rdf:li></rdf:Bag> </gn:Geo_Feature>
```


Figure 4.4: Geo-Net-PT02 extract

4.6 Summary

The new Geo-Ontology congregates many sources of information and very distinct information. To generate the Ontology has been preformed some tasks for the cleaning and integration and posterior exportation of the Geo-Net-PT02. This chapter has presented all of the tasks for the construction of this Geo-Ontology.

Chapter 5

Conclusion and Future Work

5.1 Conclusions

The main goals of this dissertation have been in large attained.

1. The identification of the main sources of information with GI from the domain of physical geography that would suit the requirement for a general ontology of the geography of Portugal,

2. The aggregation of this information and its integration under a common model with the same rules of representation,

3. Integration of GI in a unique ontological model,

4. Generation of new geographic ontology of Portugal, Geo-Net-PT02, which is freely distributed by Linguateca

There is already in Portugal an official activity, SNIG, that has the aggregation of the metadata as one of its objectives. SNIG provides the location and characterization of all the geographic information produced in the Portuguese territory concerning this territory. However, the objectives of SNIG and our objectives for Geo-Net-PT02 differ. While the SNIG is centered in the compilation of a database allowing the location and characterization of geographic information as it was produced, this and other activities concerned with the development of Geo-Net-PT02 are concerned with the creation of a common ontological model for geographic information. This information is shared and made available as an OWL document.

Geo-Net-PT02 is available from http://poloxldb.linguateca.pt. It has a total of 430 000 instances, of which circa 20 000 instances correspond to entities from the physical domain of geography of Portugal.

5.2 Future work

This work should not stop in time; Geo-Net-PT02 should be periodically undated with more Geographic Information, to better reflect current reality. In addition, further developments of the Geo-Net-PT02 should include:

- The extension of the ontological model in the domain of physica geography; more types and feature-types, more relationships, are required to make the geographic ontology more detailed,
- Automation of the processes of integration of new GI into GKB. Although the process of incorporating information from the geographic domain was in general completed, many tasks have been performed manually,
- Develop converters of Geo-Net-PT02 from the ontology format into one of the import formats accected by common GIS. This would make Geo-Net-PT02 readily useful to developers of other applications, like the watersheds plans, or cadastre in telecommunications.

Appendix A

Count of collected data

Types	**Number-of-features**
Adega Reconhecida	311
Albufeira(ponto)	1880
Aldeia Preservada	217
Aquifero	125
Área de Paisagem Protegida	18
Area Protegida	49
Bacia de Escoamento	209
Bacia Hidrografica	129
Biótopo	120
Campo de Golf	65
Casino	8
Edifício Notável	3877
Especie Notável	2432
Estabelecimento de Dormida	2847
Estação Arqueológica	456
Estância Termal	35
Ferrovias	103
Fortificação	263
Gruta	7
Jardim Botânico	7
Kartódromo	30
Lagoa	10
Local de Culto	23
Marina	26
MonumentoNatural	5
Museu	524
Nascente	220
Parque de Campismo	150
Parque Nacional	55
Parque Natural	12
Património Mundial	11

Table A.1: Count of Total features collected

Types	Number-of-features
Paul	441
Praias	591
Precipitação	254
RAMSAR	10
Rede Hidrográfica	7589
Regadio	90
Região Natural	437
Reserva Natural	9
Temperatura	177
Tipos de Paisagem	551
Tipos de Solo	588
Zonas de Geada	184
Zonas de Humidade Relativa	162
Zonas de Indice Conforto	726
Zonas Fitogreograficas	278
Zonas de Litologia	2245
Zonas de PhSolo	312
Zonas de Sismicidade	121
TOTAL	28989

Table A.2: Count of Total features collected

Appendix B

Information about the data

ADG-Adega reconhecida

PT-Casa terrea onde se guarda o vinho de interesse público (Caves vinho Porto)
EN-House of public interest where wine is kept (Caves vinho Porto)

PROJECÇÃO - Gauss ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

ADS-Açude

PT-Construção feita em rio ou levada para represar a água destinada a moinhos, regas ou ao abastecimento de populações
EN-Construction made in a river to store water for mills, agriculture or human comsumption

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

ALB-Albufeira

PT-Lagoa formada pelo mar e suas marés; Represa artificial de águas pluviais, dos rios ou do degelo
EN-Lagoon created by the sea and it's tides; Artificial dam for rain water, rivers

or thawing

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

ALP-Aldeia preservada
PT-Aldeias antigas recuperadas e preservadas para turismo
EN-Renovated and preserved villages for tourism

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

ANT-Árvore notavel
PT-Espécies arbóreasnotaveis pela antiguidade e ou raridade
EN-Species of trees notable by their age or and or rarity

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

APF-Área protegida para fauna
PT-Áreas protegidas incluem áreas terrestres, águas interiores e marinhas em que a fauna apresenta, pela sua raridade, valor ecológico ou paisagístico, importância científica, cultural e social
EN-Land and sea areas in which the fauna presents, either by their rarity, ecological value or sight, scientifical, cultural or social value

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA - Internacional
DATUM Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS militares
ORIGEM DAS COORDENADAS Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

APP-Área paisagem protegida
PT-Áreas protegidas devido ao ser valor paisagístico
EN-Protected areas due to their sight value

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA Internacional
DATUM Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS militares
ORIGEM DAS COORDENADAS Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

APR-Área protegida
PT-Áreas protegidas devido a factores de risco (importância ou vulnerabilidade)
EN-Area due to risk factors (importance or vulnerability)

ELIPSÓIDE DE REFERÊNCIA Internacional
DATUM Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS militares
ORIGEM DAS COORDENADAS Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

AQU-Aquífero
PT-Formação geológica subterrânea que armazena água e permite a sua circulação
EN-Underground geological formation that stores and allows water circulation

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

ARQ-Arquipélago
PT-Grupo de ilhas próximas umas das outras
EN-Group of nearby islands

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

BHD-Bacia hidrográfica
PT-Conjunto de terras banhadas por um rio principal e seus tributários (afluentes, subafluentes etc.)
EN-Hidrographic basin is a set of lands bathed by a main river and its tributaries (affluent, subaffluents etc.)

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

BIO-Biótopo
PT-Área povoada por um conjunto de seres vivos perfeitamente adaptados ao meio
EN-Area peopled by living beings perfectly adapted to their environment

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA - Internacional
DATUM Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

BRR-Barrinha
PT-Comunicação de uma lagoa com o mar em certas épocas do ano
EN-Conection from a lagoon to the sea during certain times of the year

PROJECÇÃO - Gauss ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

BAR-Barranco
PT-Sulco feito no solo pelas enxurradas
EN-Ridge created by rain water slides

PROJECÇÃO - Gauss ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

CAB-Cabo
PT-Saliência da costa marítima
EN-Salience of the maritime coast

PROJECÇÃO - Gauss ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

DUN-Dunas
PT-Acumulação ou monte de areia nas regiões desérticas e nas regiões litorais, sob a acção do vento de direcção constante
EN-Hill of sand in litoral and desert areas built by the action of wind

PROJECÇÃO - Gauss ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

DVT-Douro vinhateiro
PT-Região Demarcada do Douro é uma região demarcada em Portugal para a produção vinícola instituída em Setembro de 1756 por alvará de D. José I. Parte da região foi classificada como Património Mundial da UNESCO em Dezembro de 2001
EN-The Douro marked region is the region on the Douro that was marked for wine production on september of 1756 by D. José I. Part of the same region was considered world heritage by UNESCO on 2001

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)

EDN-Edificio notavel
PT-Edificios com valor, seja por antiguidade, valor arquitectónico ou utilidade
EN-Building with a high value caused by its arquitectonic interest or usefullness

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
RESP. EXEC. ORIGINAL - Atlas do Ambiente
DISPONIBILIZADOR - SNIG

ETU-Estuário
PT-Parte de um rio antes da foz, em geral com a forma de funil. É uma zona afectada pelas marés, onde se mistura água doce e água salgada
EN-An estuary is a semi-enclosed coastal body of water with one or more rivers or streams flowing into it, and with a free connection to the open sea

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

FER-Ferrovia ou caminho-de-ferro
PT-Caminho formado por dois trilhos de ferro paralelos ou carris, sobre os quais circulam os comboios
EN-Consist of two parallel rails, secured to crossbeams on wich trains travel on

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

GAN-Gândara
PT-Terreno arenoso pouco produtivo ou estéril
EN- Sandy terrain of little production value or completely sterile

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

GRT-Gruta ou caverna
PT-Cavidade natural ou artificial na rocha, de grandes dimensões
EN-A cave is a natural underground void large enough for a human to enter

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

LAG-Lagoa
PT-Extensão de água mais pequena que um lago, rodeada de terra por todos os lados e com pouca profundidade
EN-Body of comparatively shallow salt or brackish water separated from the deeper sea by a shallow or exposed sandbank

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

LEZ-Lezíria
PT-Terreno alagado pelas enchentes, nas margens de um rio
EN-Land flooded by floods, in the edges of a river

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

MAR-Marinha
PT-Represa de água do mar para a extracção do sal; salina
EN-Water dam of the sea for the extracção of salt; salt mine

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

MAT-Mata ou bosque

PT-Terreno cheio de árvores silvestres

EN-Full land of wild trees; forest

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

MNT-Monumento natural

PT-Monumento de origem natural

EN-Natural Monument

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

NAS-Nascente

PT-Local do qual se inicia um curso de água (rio, ribeirão, córrego), seja grande ou pequeno

EN-Point where groundwater flows out of the ground, Starting a river (large or small)

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

PAU-Paúl

PT-Um ecossistema lacunar

EN-Type of wetland which is subject to frequent or continuous inundation

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

PCP-Parque de campismo
PT-Local de acampamento
EN-Camping place

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

PNT-Parque natural
PT-É uma área de conservação, fora de uma área urbana, protegida por lei com o objetivo de preservar a flora e a fauna local e proporcionar um ambiente de lazer
EN-An area of land and/or sea especially dedicated to the protection and maintenance of biological diversity,used for leisure, and managed through legal or other effective means

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RIO-Rio
PT-Corrente natural de água que flui com continuidade. Possui um caudal considerável e desemboca no mar, num lago ou noutro rio
EN-Natural stream of water, usually freshwater, flowing toward an ocean, a lake, or another river

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RNT-Reserva natural
PT-Zonas de reserva natural

EN-Protected area of importance for wildlife, flora, fauna or features of geological or other special interest, which is reserved and managed for conservation and to provide special opportunities for study or research

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RSB-Reserva botânica
PT-Zonas de reserva de elementos botânicos
EN-Protected areas of botanical importance

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

SAP-Sapal
PT-Formações aluvionares periodicamente alagadas pela água salgada e ocupadas por vegetação halofítica.
EN-Type of marsh that is a transitional intertidal between land and salty or brackish water (e.g.: sloughs, bays, estuaries). It is dominated by halophytic (salt tolerant) herbaceous plants.

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

SER-Serra
PT-Terrenos acidentados com fortes desníveis, freqüentemente aplicados a escarpas assimétricas, possuindo uma vertente abrupta e outra menos inclinada
EN-Lands with strong unevennesses, frequentlywith anti-symmetrical scarps, possessing an inclined abrupt side

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

SIS-Sismicidade
PT-Zonas de grau de sismicidade
EN-Earthquake degree

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

TZE-Zona ecológica
PT-Zonas ecológicas demarcadas
EN-Marked ecological

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

UNP-Unidade de paisagem
PT-Zonas de unidades de paisagem demarcadas
EN-Marked viewing areas

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

VEG-Vegetação
PT-Plantas existentes numa certa área, que podem variar bastante, conforme o clima ou o solo
EN-Plants existing in a certain area that vary according to the soil and the station of the year

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

TEP-Temperatura

PT-Zonas de classes de temperatura

EN-Temperature class zone

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RIB-Ribeira

PT-Pequeno curso de água

EN-Small water course

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RRO-Ribeiro

PT-Curso de água perene

EN-Perrenial water course

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

VAL-Vala

PT-Vala para escorrimento de água

EN-Ditch to water rain

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RGN-Região natural

PT-Regiões de características naturais

EN-Natural characteristics area

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RCT-Recurso turistico

PT-Meios e recursos turísticos

EN-Touristic resources

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

RHD-Região hidrográfica

PT-Zonas de planos de bacia hidrográfica

EN-Hidrographic basin area

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

PIT-Pontos de interesse turistico

PT-Pontos com interesse turístico

EN-Touristic interest area

B. INFORMATION ABOUT THE DATA

PROJECÇÃO - Gauss
ELIPSÓIDE DE REFERÊNCIA -Internacional
DATUM - Lisboa (long. 09 07' 54.862" / lat. 38 42' 43,631")
COORDENADAS - militares
ORIGEM DAS COORDENADAS - Ponto Fictício (W Cabo S. Vicente)
DISPONIBILIZADOR - SNIG

Appendix C

Fonts of the collected data

Type	Source
Albufeiras	Atlas do Ambiente, Direcção-Geral do Ambiente,Preparada com elementos publicados pelo INAG, pelo IEADR ,e outras fontes
Árvores Notáveis	Estação Agronómica Nacional
Áreas Protegidas	Divisão do Atlas do Ambiente, D.G. dos Recursos Naturais, Preparada com elementos cedidos pelo Serviço Nacional de,Parques, Reservas e Conservação da Natureza
Avifauna	Divisão do Atlas do Ambiente, D.G. dos Recursos Naturais
Biótopos	Divisão do Atlas do Ambiente, D.G. dos Recursos Naturais, Preparada com elementos cedidos pelo Serviço Nacional de,Parques, Reservas e Conservação da Natureza
Classe de Sismicidade	Instituto de Meteorologia
Classe de Temperatura	Serviço Meteorológico Nacional
Lagoas	Geógrafa na Direcção-Geral do Ambiente),Preparada com elementos cartográficos do Atlas do Ambiente em colaboração com o ICN
Nascentes minerais	Atlas do Ambiente
Regiões Naturais	Estação Agronómica Nacional
Paúl	Direcção-Geral do Ambiente
Lagoas	Direcção-Geral do Ambiente
Precipitação	Serviço Meteorológico Nacional
Recursos turísticos	Direcção-Geral do Ambiente
Regadios	Direcção-Geral do Ambiente
Ferrovias	Atlas do Ambiente, Direcção-Geral do Ambiente,Preparada com elementos coligidos no IPCC, JAE, Brisa e CP

Table C.1: Sources of GI

Type	Source
Principais bacias hidrográficas	Atlas do Ambiente, D. G. dos Recursos Naturais
Recursos aquíferos subterrâneos	Direcção-Geral de Saneamento Básico
Rede hidrográfica	Atlas do Ambiente, D.G.dos Recursos Naturais
Tipos de Paisagem	Estação Agronómica Nacional
Zonas fitogeográficas predominantes	ISA
Zonas húmidas	Direcção-Geral do Ambiente,Preparada com elementos cartográficos do Atlas do Ambiente em colaboração com o ICN
Zonas de Protecção Especial	Divisão do Atlas do Ambiente, D.G. dos Recursos Naturais
Zona económica exclusiva	Atlas do Ambiente
Zonas Ecológicas	Estação Agronómica Nacional
Zonas Húmidas	Direcção-Geral do Ambiente
Zonas RAMSAR	Direcção-Geral do Ambiente
Zonas de Protecção Especial para a Avifauna	Divisão do Atlas do Ambiente, D.G. dos Recursos Naturais, Preparada com elementos cedidos pelo Serviço Nacional de,Parques, Reservas e Conservação da Natureza

Table C.2: Sources of GI

Appendix D

Fonts of Information

National Level

Estradas de Portugal
Instituto da Água, I.P.
Instituto da Vinha e do Vinho
Instituto de Conservação da Natureza e da Biodiversidade
Instituto de Gestão do Património Arquitectónico e Arqueológico
Instituto de Investigação Científica Tropical
Instituto de Meteorologia
Instituto dos Mercados de Obras Públicas e Particulares e do Imobiliário
Instituto Geográfico do Exército
Instituto Geográfico Português
Instituto Hidrográfico
Instituto Nacional de Engenharia, Tecnologia e Inovação
Instituto Nacional de Estatística
Instituto Nacional de Formação Turística
Instituto Nacional de Recursos Biológicos, I.P.
Instituto Portuário e dos Transportes Marítimos
Instituto Português de Arqueologia
Direcção Geral de Actividades Económicas
Direcção Geral das Pescas e Aquicultura
Direcção Geral de Agricultura e Desenvolvimento Rural

Direcção Geral de Edifícios e Monumentos Nacionais
Direcção Geral de Energia e Geologia
Direcção Geral de Turismo
Direcção Geral do Ordenamento do Território e Desenvolvimento Urbano
Direcção Geral dos Recursos Florestais
Gabinete de Estratégia e Planeamento
Agência Portuguesa do Ambiente
Centro Nacional de Cultural
Correios de Portugal
Estrutura de Projecto para a Reposição da Legalidade
Gabinete de Assuntos Europeus e Relações Exteriores
Gabinete de Estatística e Planeamento da Educação
Laboratório Nacional de Engenharia Civil
Secretariado Técnico dos Assuntos para o Processo Eleitoral
Serviço nacional de bombeiros e protecção civil

Regional Level

Comissão de Coordenação e Desenvolvimento Regional de Lisboa e Vale do Tejo
Comissão de Coordenação e Desenvolvimento Regional do Alentejo
Comissão de Coordenação e Desenvolvimento Regional do Algarve
Comissão de Coordenação e Desenvolvimento Regional do Norte
Comissão de Coordenação e Desenvolvimento Regional do Centro
Direcção Regional de Agricultura da Beira Interior
Direcção Regional de Agricultura da Beira Litoral
Direcção Regional de Agricultura de Entre Douro e Minho
Direcção Regional de Agricultura de Trás-os-montes
Direcção Regional de Agricultura do Algarve
Direcção Regional de Agricultura e Pescas de Lisboa e Vale do Tejo
Direcção Regional de Agricultura e Pescas do Alentejo
Direcção Regional de Economia do Algarve
Área Metropolitana de Lisboa
Associação de Desenvolvimento da Região do Alto Tâmega
Associação de Desenvolvimento Rural Integrado das Terras de Santa Maria
Empresa de Desenvolvimento e Infraestruturas do Alqueva, SA

Escola Superior Agrária de Bragança - Instituto Politécnico de Bragança
Escola Superior Agrária de Ponte de Lima
Faculdade de Ciências (U.L.)

Local Level

Agrupamento de Municípios de Abrantes, Constância, Gavião, Mação e Sardoal (GAT Abrantes)
Agrupamento de Municípios de Santarém e Salvaterra de Magos (GAT Santarém)
Agrupamento de Municípios de Torres Novas (GAT Torres Novas)
Associação de Municípios da Bairrada/Vouga (CM Anadia)
Associação de Municípios da Ria
Associação de Municípios da Terra Quente Transmontana (GAT Mirandela) Associação de Municípios de Castelo de Vide, Marvão e Portalegre (CM Portalegre)
Associação de Municípios do Litoral Alentejano
Associação de Municípios do Norte Alentejano
Associação de Municípios do Oeste
Associação de Municípios do Vale do Douro Norte
Associação de Municípios dos Vales do Ceira e Dueça (GAT Lousã) Grande Área Metropolitana do Porto
Câmara Municipal da Amadora
Câmara Municipal da Maia
Câmara Municipal da Marinha Grande
Câmara Municipal da Moita
Câmara Municipal da Póvoa de Varzim
Camara Municipal de Albergaria-a-Velha
Câmara Municipal de Alcanena
Câmara Municipal de Alcochete
Câmara Municipal de Aljustrel
Câmara Municipal de Arouca
Câmara Municipal de Aveiro
Câmara Municipal de Beja
Câmara Municipal de Braga
Câmara Municipal de Caminha
Câmara Municipal de Cantanhede
Câmara Municipal de Cascais Câmara Municipal de Castanheira de Pêra

Câmara Municipal de Castelo Branco
Câmara Municipal de Castro Marim
Câmara Municipal de Coimbra
Câmara Municipal de Espinho
Câmara Municipal de Esposende
Câmara Municipal de Estarreja
Câmara Municipal de Évora
Câmara Municipal de Faro
Câmara Municipal de Felgueiras
Câmara Municipal de Ferreira do Zêzere
Câmara Municipal de Gondomar
Câmara Municipal de Gouveia
Câmara Municipal de Guimarães
Câmara Municipal de Lisboa
Câmara Municipal de Loures
Câmara Municipal de Mafra
Câmara Municipal de Matosinhos
Câmara Municipal de Monchique
Câmara Municipal de Odivelas
Câmara Municipal de Oeiras
Câmara Municipal de Oliveira de Frades
Câmara Municipal de Palmela
Câmara Municipal de São João da Madeira
Câmara Municipal de São Pedro do Sul
Câmara Municipal de Serpa
Câmara Municipal de Setúbal
Câmara Municipal de Sever do Vouga
Câmara Municipal de Tomar
Câmara Municipal de Vagos
Câmara Municipal de Viana do Castelo
Câmara Municipal de Vila Franca de Xira
Câmara Municipal de Vila Nova de Cerveira
Câmara Municipal de Vila Nova de Famalicão
Câmara Municipal de Vila Real

Câmara Municipal de Vila Velha de Rodão
Câmara Municipal de Vouzela
Câmara Municipal do Barreiro
Câmara Municipal do Entroncamento
Câmara Municipal do Montijo
Câmara Municipal do Porto
Câmara Municipal do Seixal

Appendix E

Topology rules

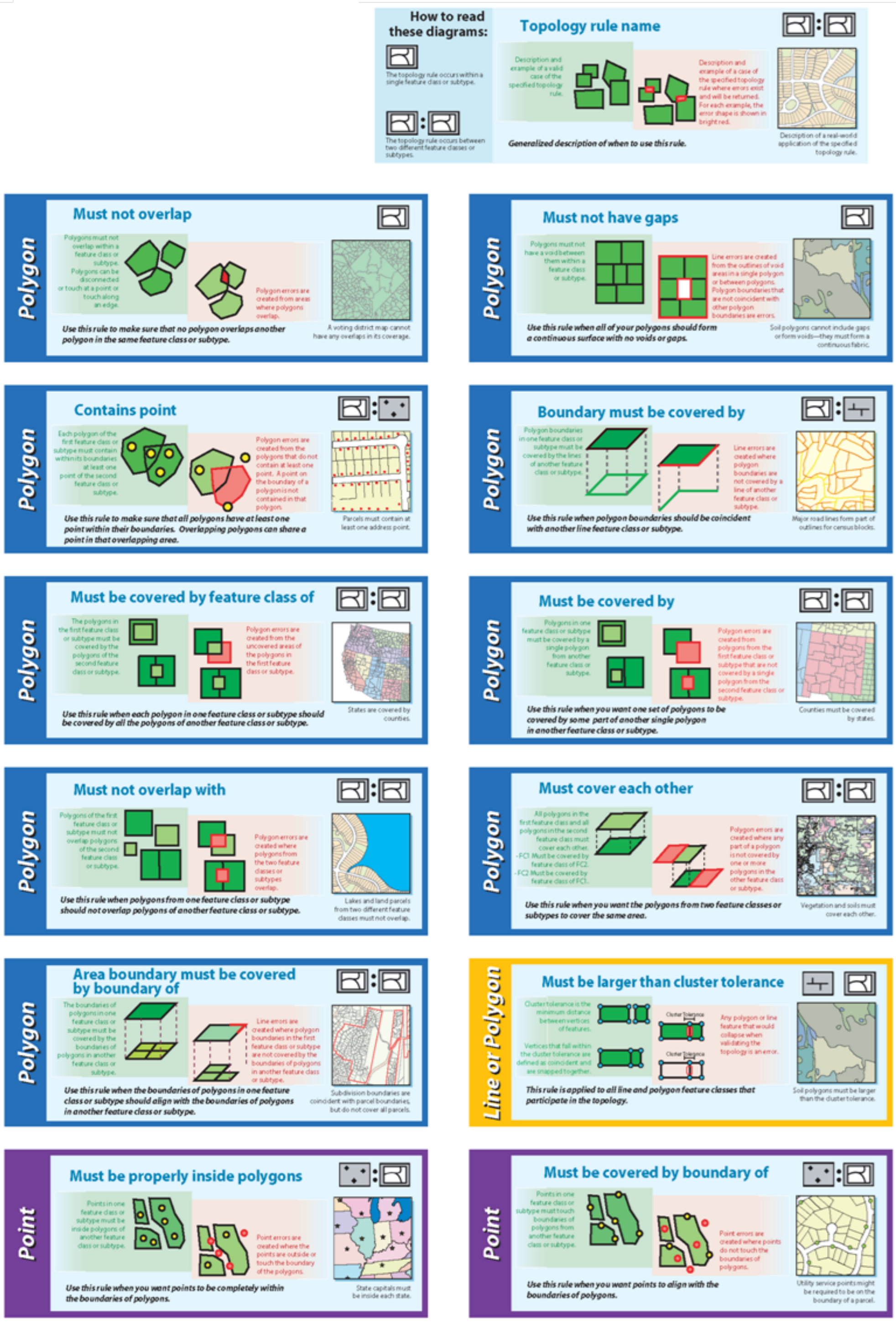

Figure E.1: Topology rules

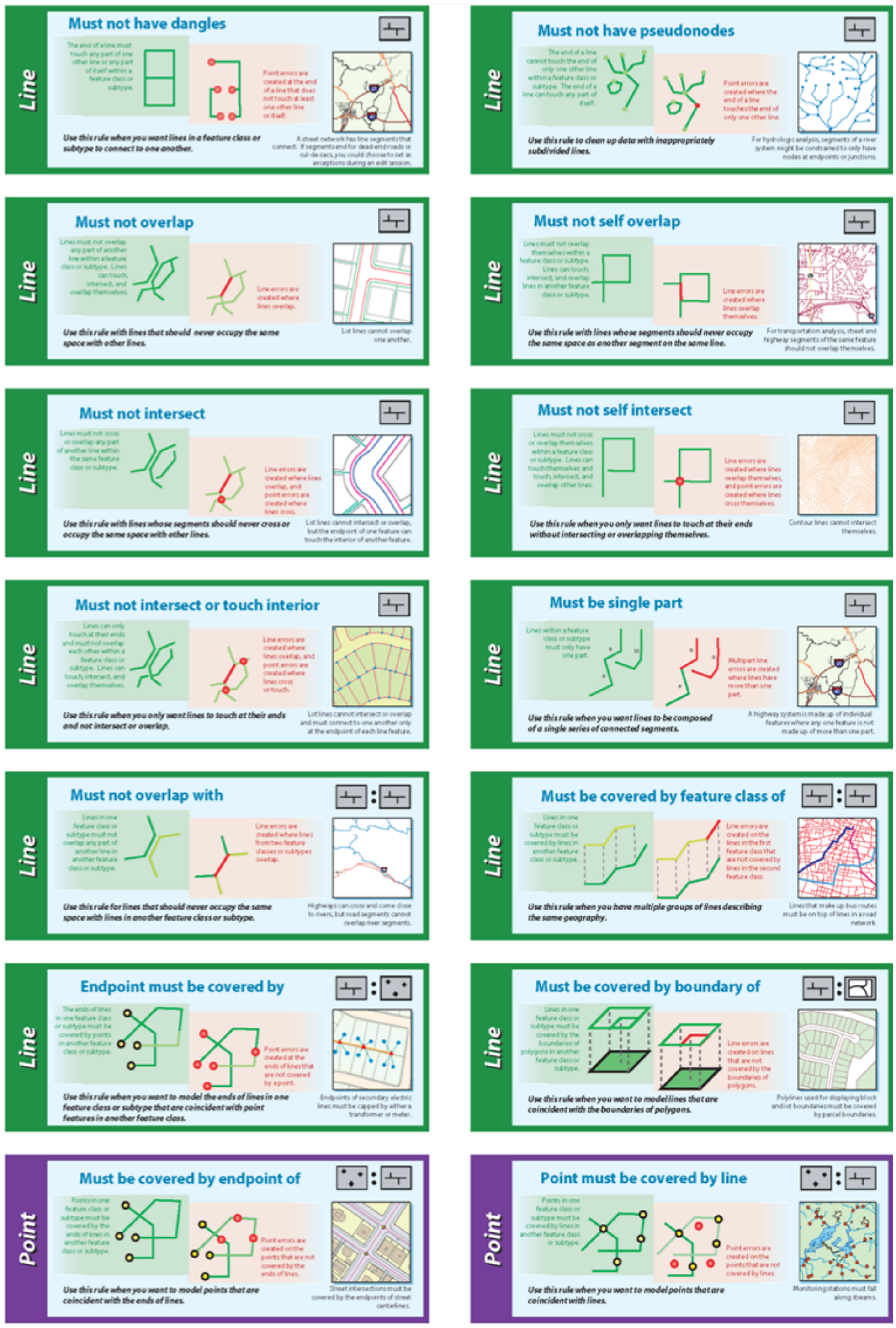

Figure E.2: Topology rules

Bibliography

ABLER, R. (1987). The national science foundation national center for geographic information and analysis. *International journal of geographical information systems*, **1**, 303 a 326. 41

ANTONIOU, G. & HARMELEN, F. (2004). *A Semantic Web Primer*. MIT Press. 1, 26, 30, 31

BERNERS-LEE, T. (2005). Uniform resource identifier (uri): Generic syntax. *http://www.rfc-editor.org/rfc/rfc3986.txt*. 26

BRICKLEY, D. & GUHA, R. (2004). Rdf vocabular description language 1.0:rdf schema. Tech. rep. 26

CASACA, J. (2005). *Topografia geral*. LIDEL - Edições técnicas,lda. 15

CHAVES, M. (2005). Gkb - geographic knowledge base. Tech. rep., Departamento de Informatica da Faculdade de Ciências da Universidade de Lisboa. 33

CHAVES, M.S. (2008). *Uma Metodologia para Construção de Geo-Ontologias*. Master's thesis, Faculdade de Ciências - Universidade de Lisboa. 39

DAVIS, M. (2008). Forms of unicode. http://www.icu-project.org/docs/papers/forms$_o f_u nicode/h8$. 20

EGENHOFER, M.J. (1989). 3rd international conference, fodo 1989 on foundations of data organization and algorithms. In *3rd International Conference, FODO 1989 on Foundations of Data Organization and Algorithms*, 457–472, Springer-Verlag New York, Inc. 41

Fensel, D. (2000). *Ontologies: Silver Bullet for Knowledge Management and Eletronic Commerce*. 28

Garshol, L.M. (2004). Metadata? thesauri? taxonomies? topic maps! making sense of it all. 16, 29

Gonzalez, M. & Lima, V. (2001). T-lex: Thesaurus com estruturação semântica e operações gerativas. *Conferencia Latinoamericana de Informática (CLEI'2001)*. 17

Gruber, T. (1993). *A Translation Approach to Portable Ontology Specifications*. 28

Guarino, N. (1992). Concepts, attributes and arbitrary relations. *Data Knowledge Engineering*, **8**, 249–261. 28

Harold, E. (2005). Encode your xml documents in utf-8. http://www.ibm.com/developerworks/xml/library/x-utf8/. 21

Hendler, J. (2001). Agents and the semantic web. *IEEE Intelligent Systems*. 2

Hill, L. (2006). *Georeferencing- The Geographic Associations of Information*. The MIT Press. 10, 22, 23

Kainz, W. (1990). Spatial relationships - topology versus order. In *Fourth Int. Syrup. on Spatial Data Handling*, vol. 2, 814–819, K. Brassel and H. Kishimoto. 41

Lassila, O. & Swick, R. (1998). Resource description framework (rdf) model and syntax specification. 26

Maphumulo, A. (2003). *Testing portability of a proposed South African spatial feature catalog standard*. Master's thesis, . International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands. 32

Matos, J.L. (2001). *Fundamentos de Informação Geográfica*. LIDEL - Edições técnicas,lda. 15

PEUQUET, D.J. & ZHANG, C.X. (1987). An algorithm to determine the directional relationship between arbitrarily-shaped polygons in the plane. *Pattern Recogn.*, **20**, 65–74. 41, 45

SCHMIDT, D. (2006). Model-driven engineering. IEEE Computer, pp 25-31. 16, 21

USCHOLD, M. & GRUNINGER, M. (1996). Ontologies: principles, methods, and applications. Knowledge Engineering Review, 11(2), (93,155). 31

W3C (2001). Web-ontology (webont) working group (closed). 30

W3C (2007). Extensible markup language (xml). 27