

A conceptual holding model for veterinary applications

Nicola Ferrè¹, Werner Kuhn², Massimo Rumor³, Stefano Marangon⁴

¹*Istituto Zooprofilattico Sperimentale delle Venezie - GIS Unit, Legnaro (PD), Italy;* ²*Institute for Geoinformatics, University of Muenster, Muenster, Germany;* ³*3DGIS, Rovigo, Italy;* ⁴*Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro (PD), Italy*

Abstract. Spatial references are required when geographical information systems (GIS) are used for the collection, storage and management of data. In the veterinary domain, the spatial component of a holding (of animals) is usually defined by coordinates, and no other relevant information needs to be interpreted or used for manipulation of the data in the GIS environment provided. Users trying to integrate or reuse spatial data organised in such a way, frequently face the problem of data incompatibility and inconsistency. The root of the problem lies in differences with respect to syntax as well as variations in the semantic, spatial and temporal representations of the geographic features. To overcome these problems and to facilitate the inter-operability of different GIS, spatial data must be defined according to a “schema” that includes the definition, acquisition, analysis, access, presentation and transfer of such data between different users and systems. We propose an application “schema” of holdings for GIS applications in the veterinary domain according to the European directive framework (directive 2007/2/EC - INSPIRE). The conceptual model put forward has been developed at two specific levels to produce the essential and the abstract model, respectively. The former establishes the conceptual linkage of the system design to the real world, while the latter describes how the system or software works. The result is an application “schema” that formalises and unifies the information-theoretic foundations of how to spatially represent a holding in order to ensure straightforward information-sharing within the veterinary community.

Keywords: geographical information system, holding, conceptual model, data harmonisation.

Introduction

Geographical information systems (GIS) have become widely used in the veterinary domain for epidemiological mapping and/or spatial data analysis (Sanson et al., 1991; Norström, 2001; Pfeiffer et al., 2008; Ferrè et al., 2011). Usually, these GIS use a geometry-centric meta-model, in which the items of interest are primarily geometric (points, lines, polygons, etc.), with additional attributes, often stored in database tables without any other essential spatial properties than those essential to formalise the characteristics of geographical phenomena in order to ensure that the information can be interpreted without ambiguity and also effectively communicated (Burrough and McDonell, 1998). According to the Open Geospatial Consortium (OGC), spatial information semantic may include, but is not limited to, metadata (describing the contents, valid features and attribute values, including guidelines for the capture of these data), symbol sets for feature representation, rule sets

for feature portrayal, relationships and dependencies between features, and association with other spatial information (Buehler and McKee, 1998). Users trying to integrate spatial data stored without these proprieties, frequently face the problem of data incompatibility and inconsistency expressed as differences in syntax or semantic, spatial and temporal representation (Tóth et al., 2012). In particular, match-making spatial information not supported by adequate spatial properties for their re-use, presents problems of (i) different file storage formats and grammar rules; (ii) name heterogeneities (different expressions for the same concept); (iii) conceptual heterogeneities (different concepts expressed by the same symbol); (iv) level of granularity; and (v) different spatial representations (e.g. different geometries like three-dimensional (3-D) *versus* 2-D ones, linear *versus* polygon representation, geometrical shift, etc.) (Cliff and Carl, 2009). For example, the “holding” concept is recognised and described in terms of well-established ideas within the veterinary community thanks to common cultural background. However, there are many ways to spatially describe a holding, and different terms can be used for different levels of resolution. Many perceived geographical phenomena used by veterinarians as explicit entities (such as “holding”, “classified bivalve mollusc area”, etc.) do not have an exact form, their extent may change with time, and the type of spatial

Corresponding author:

Nicola Ferrè

Istituto Zooprofilattico Sperimentale delle Venezie - GIS Unit
Viale dell'Università 10, 35020 Legnaro (PD), Italy
Tel. +39 049 808 4253; Fax +39 049 883 0268
E-mail: nferre@izsvenezie.it; crev.nferre@gmail.com

property describing them at one scale is likely to be different from that at another.

Conceptual modelling has been suggested as a way to create a more “intelligent” approach to spatial information. Conceptual modelling is the abstraction process describing all relevant static and dynamic aspects (i.e. all rules and laws, etc.) of the universe of discourse (for a specific phenomenon). Abstraction implies the need for simplification of the analysed system and for assumptions about what is not known about the system. In other words, a conceptual model is a type of abstraction that uses logical concepts and hides the details of implementation and data storage (ISO/IEC, 1998). Conceptual models have been used as tools (i) to describe and structure geographical phenomena; (ii) to assess the relevance of existing rules for description, formats, data models, etc.; and (iii) to improve them so they convey a common conceptualisation through which it is possible to develop tools capable of bringing heterogeneous datasets together (Frank and Goodchild, 1990; Groot and McLaughlin, 2000).

Conceptual models may be implicit or explicit. The former is the cognitive knowledge that refers to something understood within an information community though not clearly or directly stated, i.e. acquired more or less directly from the interaction with other members of the information community (Kuhn, 2005). Because of this, implicit models are particularly good at providing quick answers that have worked well in the past, e.g. a user representing a holding as a point just because the location of the farmhouse is represented by a single point in most animal health information systems (Durr and Froggatt, 2002). The explicit model, on the other hand, lets the user know what can be expected from the data (Kuhn, 2005). This approach has many advantages as it not only improves the understanding of a problem by providing a formal model that facilitates communication, but also allows easy upgrading and management of the system so it can be used for other systems (Nativi and Federici, 1994). In addition, making a conceptual model explicit facilitates database maintenance and offers a level of description for scientific or economic analyses that is useful beyond the implemented system itself (Worboys and Duckham, 2004). For instance, if we take the specific example of an information system for animal holdings, veterinarians need to know what kinds of holding are included in a database, how precise their geographical positions are and what additional information is supplied. Furthermore, when the data from such an informa-

tion system are being combined with other records, e.g. data collected at the European level, the conceptual model allows assessment of the possibilities for inter-operability.

The standard mechanism used to make conceptual models explicit is based on information system ontologies, i.e. specifications of system concepts (Kuhn, 2001). With regard to an animal holding, for example, the concept would primarily involve the area used for this purpose, but it does also comprise its attributes. The main elements of ontologies are objects and processes including their specific properties (Frank, 2013). For instance, by classifying a holding as an object (with properties such as position, animal capacity, etc.) and the animal conditions as processes (with properties like health status, disease presence, etc.), the conceptual model can be related to other models. The detailed decisions involved in this type of classification are often intricate, as they must capture what every term in the database used stands for. They may also go beyond physical objects and processes to involve the cognitive and social reality, e.g. by including regulations concerning restriction areas.

Ontologies are engineered with specific design languages and implemented through standard, semantic Web languages such as the resource description framework (RDF; <http://www.w3.org/RDF/>) or the ontology Web language (OWL; <http://www.w3.org/TR/owl-features/>). Implementation permits automated reasoning and assessment of the compatibility of a particular ontology with other ontologies (and the conceptual models representing them). For example, it should be possible to compare an ontology of animal holdings in one region with others within a country and, indeed, with holdings all over the continent.

Background information

The idea of developing an explicit GIS-based model of the location of holdings has not, as far as we know, been presented in the veterinary domain. Our model has been built in accordance with directive 2007/2/EC of INSPIRE, i.e. the European infrastructure for spatial information in the European Union (EU), which has now reached its implementation phase. This directive (and additional regulations) has resulted in an infrastructure for spatial informatics in Europe that supports policies and regulates activities related to the environment. It is underpinned by a framework of standards and rules for implementation with the aim of:

- (i) signposting data provided by public and private organisations to given metadata standards;

- (ii) making data available with given specifications;
- (iii) publication of information through specific Web services making it possible to view and download these data; and
- (iv) facilitating inter-operative data-sharing among EU bodies.

The data specification component specifies data models and a number of other requirements intended to support sharing of spatial information among the EU bodies. In particular, INSPIRE data specifications provide the essential, technical documentations that prescribe detailed and precise definition of data content, data quality, data consistency, reference systems and metadata for 34 selected themes. Directive 2007/2/EC identifies some themes that fall under the veterinary sector namely: agricultural and aquaculture facilities (AF), species distribution, area restriction and reporting units, human health and safety, production and industrial facilities, and sea regions (a complete description of the themes is provided in the document “Definition of annex themes and scope” available at the INSPIRE Web portal: <http://inspire.ec.europa.eu/index.cfm/pageid/2>).

The INSPIRE thematic working group (TWG) for AF regulates and supports requirements, design and characteristics for all agricultural activities, such as cultivating soils, producing crops as well as holding and rearing animals. The AF theme has developed specifications to facilitate inter-operability of all geographical information related to AF topics among the EU member states. It covers areas that might be lacking in other themes that also address agricultural and aquaculture concepts, but at different levels of detail (INSPIRE, 2012b). INSPIRE defines two types of model: the core and the extended. The former contains specifications for legally binding data and covers the most essential geophysical feature types that are relevant for the community domain (essentially to minimise the work load of obligatory data transformations). The extended model was developed to fully support the various cases identified by the information community. We find these two applications overly limited with respect to description of veterinary issues as they are mainly centred on environmental aspects and integration issues with other INSPIRE themes. In particular, the AF theme identifies a “holding” as an example of a holding attribute derived from the list of statistical classifications of economical activities in Europe (NACE). We have also found that the proposed schema shows limitations with respect to breeding methodology. For example, it discusses only techniques where animals are constrained inside an infra-

structure (“AgriBuilding” or any other “Structure”) or a delimited portions of land or water (the “Plot” feature) and do not consider particular types of breeding techniques, e.g. those practiced by nomads or by bee hive keepers, common fell grazing, etc.

This paper proposes a conceptual model of animal holdings for GIS applications in the veterinary domain. This is done as it would be useful to extend the generalist model proposed by INSPIRE towards the specificity of the veterinary domain. The proposed conceptual model of holdings for veterinary applications is aimed to achieve a generic, extendable and inter-operable, conceptual model that relies on geographical and veterinary concepts. The model discussed here is designed to serve as a basis for any type of animal holding data and veterinary analyses with known location.

Model creation

The INSPIRE generic, conceptual modelling for developing data specifications is based on the general feature model (GFM) defined by the International Organization for Standardization (ISO) under 19109:2005. Information GFM models emphasise the conceptual significance of the data, not just the structure. The aim for the digital representation is to reflect the language used by practitioners in the application domain. The advantage of this approach is that (i) semantics, rather than the geometry, are primary; (ii) the feature case carries properties (such as attribute, association/relationship, operation and behaviour); and (iii) a feature may have more than one geometry associated with it, each representing different spatial properties and implementing different representations of a property (e.g. at different scales, or using different geometry models) (ISO, 2005). We produced two models based on the conceptual modelling process. The purpose of the first, the “essential model”, was to provide a description of the thinking behind the system, its design and theoretical basis with a further aim to establish conceptual linkages to the real world. The essential model is a meta-model that contains the defining schema, which specifies the concepts, terminology, operations and assumptions needed to specify the basic constructs of the other model, i.e. the “abstract model”. This latter model describes the spatial data according to veterinary informatics (the domain) specifying objectives, inputs, outputs, content, assumptions and simplifications of the modelled information. It is an application-model that defines the types of features and processes that produce datasets

of geographical information (Cliff and Carl, 2009; Kotiadis and Robinson, 2008). The result of the abstract model is an application schema that is used to capture data about features. For the analysis of the “essential model” we took into account both the AF documentation and the core European veterinary legislations on animal health, animal disease and animal movement as expressed by the following directives: 64/432/EEC, 82/894/EEC, 88/571/EEC, 90/425/EEC, 90/426/EEC, 90/539/EEC, 91/068/EEC, 91/496/EEC, 92/102/EEC, 2005/94/EC, 2006/88/CE, 2008/1242/EC and 2008/1251/EC.

The AF specification defines a holding as “the whole area and all infrastructures on it under the control of an operator to perform agricultural or aquaculture activities”. Three different parts can be identified in this definition: (i) “the whole area and all infrastructures on it”; (ii) “under the control of an operator”; and (iii) “to perform agricultural or aquaculture activities”. The first part is composed of two nouns, area and infrastructure, and a verb binding them together. From this sentence, we infer that there could also be some physical structures within this “area”. The second part introduces the role of the “operator” identified by the specification as a person with overall responsibility, not only for management (i.e. for looking after and raising animals, planning strategies for maximum yield, etc.), but also for legal aspects, such as permissions, identification of responsible parties, etc., connected with the “area”. These activities are important for INSPIRE because they differentiate the “holding area” from everything else, including land use and land cover, where there is no need for an “operator”. The last part of the definition deals with the type of activities carried out on the “area” and INSPIRE classifies the sentence further by presenting it together with a list of exclusions differentiating the AF activities from those of other TWGs that also cover activities with agricultural or aquaculture aspects, but with different foci.

Our “essential model” extends the area concept developed by INSPIRE into an idea that refers to a surface specialised for animal keeping. Specifically, we intend to exclusively associate the area concept with land parcels organised for animal confinement. Accordingly, we assume that a “holding area” exists if we observe, at a given location in space and time, properties that can serve as animal confinements. We thus need to recognise two observable properties of reality to allow the identification of an area as such: the presence of (i) animals; and (ii) permanent or semi-permanent structures (e.g. buildings, shelters)

dedicated to animal confinement (the simple delimitation of a parcel of land with fence, walls or mile stones is not sufficient and does not meet this condition). These two properties are usually seen simultaneously, but one can also think of dedicated confinement areas with no animals inside or, alternatively, areas with animals but without structures dedicated to confinement. Examples of the latter include nomadic sheep keeping, where flocks move along given paths whose boundaries are spatially undeterminable, the common fell grazing where sheep flocks that belong to several different farms graze freely without fences or walls, or the area surrounding a bee farm. According to this analysis, we assert that an “area” can be seen as an abstract element that can be described by two mutually exclusive observable elements: (i) “area with structure” and (ii) “area without structure”. For the latter, the presence of animals is a necessary condition.

Like INSPIRE, veterinary informatics considers the operator’s role a constitutional element of the holding concept. For veterinary legislation, the operator is a natural person (with legal attributes) who is responsible for:

- (i) adherence to the veterinarian regulation requirements, e.g. the operator shall keep the competent authority continually informed about the status of the establishment and activities under his/her control;
- (ii) supervision of the health status of the animals under his/her control; and
- (iii) exercise of due care to minimise the risk of introduction, development and spread of disease so as to protect human and animal health as well as the environment.

The holding concept incorporates the understanding that the operator establishes the size of the farm from the legal point of view, decides the type of structures needed, the level of sanitary measures that should be adopted and other elements that shape the “area” in question. To complete our analysis, we must also define the type of activities performed in an animal holding and do so with consideration to those indicated by INSPIRE, which considers all agricultural activities as part of the conceptual model of a holding. We propose narrowing down the type of activities taking place in a “holding” to exclusively mean “the keeping of animals”. The verb “keep” refers here to the list of activities mentioned by the veterinary legislation, which is used as reference (i.e. to hold, to keep, to breed, to handle and to rear animals). In accordance with the assumptions supporting our analysis of the

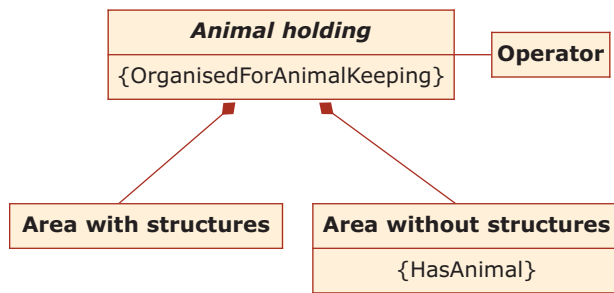


Fig. 1. Holding essential model schema.

“essential model”, we propose the following definition “holding is an area organised by an operator to perform activities solely related to animal keeping”.

To give to the “essential model” sufficient rigor for unambiguous representation and to formalise and unify its informatic-theoretic foundations within the veterinary community, a process is necessary to also develop the “abstract model”. To that end, we begin by presenting the OGC that was developed to link real-world facts with their representation as a collection of geometrical features (Cliff and Carl, 2009).

The OGC abstraction process considers nine levels of abstraction. The first five layers (moving from the “real world” to the “project world”) deal with the abstraction of real-world facts, while the final four layers (going from points to feature collection) deal with mathematical and symbolic models of the world and are meant to be modelled using GIS software. We

deal here with the first five layers of abstraction and aim to continue the presentation in a follow-up paper. We describe the main results that can be obtained with each layer ending with the conceptual model of a holding, defined according to veterinary informatics, which comprises the final layer.

In the “real world layer”, we describe what for us determines a “holding”. The items that characterise this term from the veterinary informatics point of view have in part been extracted from the “essential model” defined above and in part from the veterinary European legislation. Abstraction of the term “holding” results in the following items:

- (i) agricultural establishment;
- (ii) premises;
- (iii) stable(s);
- (iv) open-air farm; and
- (v) any installation where animals are held, kept or handled.

We see that these items (except for “open-air farm”) are characterised by the presence of a structure dedicated to confining animal(s). We associate these items with the condition defined as “area with structure”, while the “open-air farm” goes with the condition described as “area without structure”. From the characterisation of our universe of discourse in the “real world layer”, we abstract the terms that a veterinarian easily identifies as the essence of the idiom “animal holding area”. The “conceptual world layer”, the first

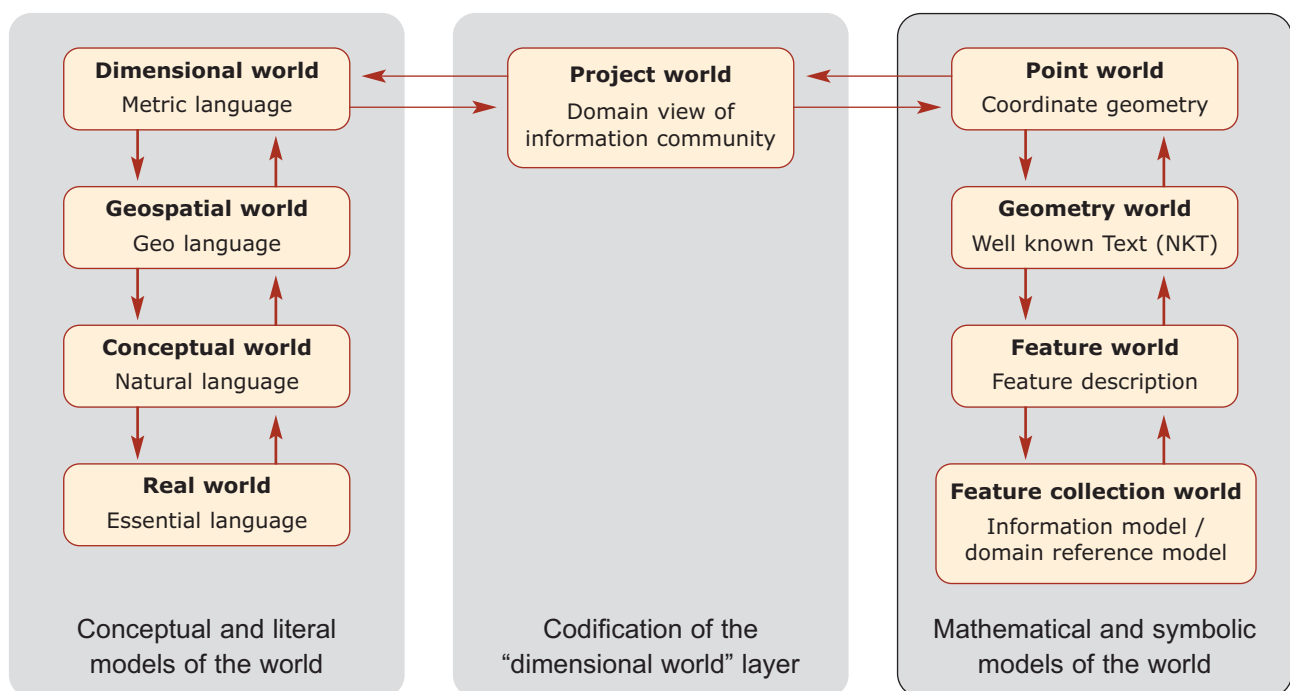


Fig. 2. OGC layers of abstraction.

level of abstraction of the real world, identifies the following three classes:

- (i) “Structures” including ground-level construction (e.g. buildings) or lower than surface construction (e.g. fish pond), intended for animal containment and used for that, as well as any other permanent construction such as ancillary buildings or semi-permanent structures, e.g. huts, mobile boxes, cages, etc.;
- (ii) “Plots” comprising all outdoor areas dedicated to animal movement that are characterised by the presence of physical boundaries for animal containment. This physical delimitation can be composed of man-made manufactures such as fences, walls, electrified ropes and any other physical barriers or be natural barriers, e.g. a river, copses of bushes, etc.; and
- (iii) “Open areas” dedicated for the movement of animals without specific, physical boundaries. Generally, in this type of area animals feed themselves with the products they find. Examples include parks, gardens, natural and semi-natural, urban, green spaces (including woodlands, urban forestry, scrubs, grasslands, etc.) wetlands, rock areas (e.g. cliffs, quarries and pits), green corridors (including rivers and canal banks), roads and cycle paths.

All possible types of “animal holding area” can be described by these classes, alone or in various combinations, such for example “only structure(s)”, “structure(s) and plot(s)”, “structure(s) and open area(s)”, only open area(s) (Fig. 3).

With the next level of abstraction, the “geospatial world layer”, the description of the type of animal holding areas defined above becomes greatly simplified with only a selection of essential characteristics visible. For example, from complex objects, e.g. a layer hen tunnel, which comprises different identifiable properties such as walls, doors, roof, etc., the abstraction selects only those parts that are of interest for veterinary informatics. In this case, the tunnel wall is the main characteristic and therefore becomes the “foot-print” representing the tunnel. In addition, the dimensionality of the features recognised in the “geospatial world” varies from 0 to 3, depending on whether they are seen as points, lines, surfaces or solids.

The feature recognised within the “geospatial world layer” incorporates a given degree of detail, or scale, into the process of identifying and representing them. The very act of identifying a particular pattern means that scale, extent, and resolution are involved (Goodchild, 2011). The choice of scale affects the type of pattern observed critically since patterns that appear at one level of resolution may be lost at lower



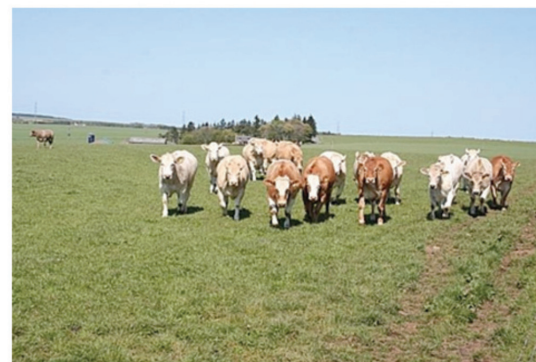
Only ‘Structure’



‘Structure’ and ‘Plot’



‘Structure and ‘Open area’



‘Open area’

Fig. 3. Possible types of animal holding area.

or higher levels (Gibson et al., 2000). In order to support both conventional, multi-scale mapping and spatial investigation in the exploratory and the quantitative analysis steps, we introduced different levels of detail in this layer so that an animal holding area can be represented. This means that different features, each one with its own geospatial rendering and with specific binary topological relationships (such as contained or part of or other spatial relationships), can be represented. We adopted multiple representations of a holding area based on the notion that (i) spatial abstraction is achieved by reducing the number of objects by means of their aggregation or elimination (Mustiere and Smaalen, 2007); and (ii) spatial analysis can be used to support the downscaling¹ of spatial data (Goodchild, 2011).

Before describing the features recognised within the “geospatial world layer”, we should introduce the idea of fiat boundaries, i.e. human demarcations that do not physically exist in nature (Worboys and Duckham, 2004). Examples include the national borders or the delimitation of postal and electoral districts. For fiat boundaries, a registry provides the exact location and extension, and surveyors are required to indicate its extension in the real world (Smith and Varzi, 2000). There are different types of fiat boundaries, for instance those drawn by epidemiologists to define restriction areas in case of an outbreak and those based on mathematical definitions, e.g. the “convex hull” or “convex envelope”, i.e. the minimal convex polygon that includes a preselected set of features. In our context, the “convex hull” polygon, calculated from a set of cases, can be considered a rough representation of the location of an object and as a very simple descriptor of its shape.

The types of feature identified within the “geospatial world layer” are the following:

- (i) “Structure”, which is the 2-D, graphical representation of objects represented by polygons describing their “footprints”;
- (ii) “Plot”, which is the 2-D, graphical representation of objects represented by polygons that correspond to the boundaries of installations used for animal confinement;
- (iii) “Open area”, which is the 2-D, graphical representation of objects represented by polygons, whose shape can be given by a mathematical function or manual editing;
- (iv) “Site”, which is the 2-D, graphical representation that describes the complete extension of land dedicated to animal confinement and thus represents the generalisation of the extent of the holding

area obtained from a set of polygons that symbolise the facilities². (In this case, one of the “sites” will be appointed the principal site with the others functioning as subordinate sites); and

- (v) “Holding”, which is a 0-D feature (point) that provides a synthetic, geographical representation of the area in question. (It was introduced in the feature list to accomplish the requirement that it be represented according to the data model adopted by the majority of veterinary organisations (Durr and Froggatt, 2002; INSPIRE, 2012b). This synthetic representation can be a point placed inside a “site polygon” (i.e. the “principal site”) or outside any polygon that composes the holding area. There can be only one “holding point”. In case the “holding area” comprises more than one “site”, it would be valid to relate the point to the “principal site”, while it would not be valid to define the point as a “centroid” of all “site polygons”).

To complete the “geospatial world layer” analysis we add the topological relationships among the recognised features. The “structure”, “plot” and “open area” features have the part of spatial relationship with the “site feature”. Note that the “holding feature” must be considered as independent with no topological relationships established with the other features.

The “dimensional world layer” represents the next level of abstraction. This layer is important for our analysis since it equips the features with metrics such as length (of a border) or size (of an area). While there are no constraints for unary relations (e.g. length of a building), some consideration is necessary for binary relations (e.g. distance between two buildings). The feature, recognised in the “geospatial world layer”, can be located in one of four different geographical entities: (i) land; (ii) fresh water (rivers and lakes); (iii) brackish

¹Inserting missing details using properties that can be identified from the coarse spatial data. In a geostatistical framework one, the most common tools for downscaling is kriging or Gaussian process regression.

²The operator can “draw” the shape of a “site” by means of the convex hull function for all objects organised for animal confinement (“structure”, “plot”, “open area”). In case empty space between the different objects that comprise the animal holding area is representative (greater than 500 m (Atkins and TripleDev, 2012), the “site” feature must be represented with different polygons, one for each group of features (the distance between objects that belong to a “site” polygon must be less than 500 m, while the distance between the “site” polygons must be greater than 500 m).

water; and (iv) sea water (lagoon, sea and ocean). Each of these entities has different physical properties that characterises the geographical space and the related “dimensional world”. The literature distinguishes between two different types of geographical space properties: isotropic and anisotropic. Even if these types of space properties are always defined in relation to a certain level of resolution or generalisation of the geographical units involved (isotropy can be observed on a certain scale and for a defined number of criteria), we can generalise the concepts and use these two types of geographical spaces to associate the necessary metrics for our analysis. In particular, we associate:

- (i) isotropy with land and brackish or salt water (sea water); and
- (ii) anisotropy with fresh water.

The hypothesis of isotropic space for land and brackish water or the sea derives from the assumption there are no directional dependencies among its parts, elements and places, while such dependencies do exist for freshwater (hierarchy nodes and axes that organise the circulation flows of the medium in question are typical signs of anisotropy). For this reason, in case the holding area is a freshwater fish farm, a specialised association of the spatial object with a hydrographic system must be implemented.

With the “project world” layer, we have reached the final level and model now the spatial extent of the recognised features with a series of structures (specification) that provide sufficient rigor for their unambiguous representation. There are two popular approaches for modelling geospatial features: vector and coverage. The first of these can model the spatial extent of a feature with point, lines, polygons and other geometric primitives. The second approach refers to a function from a spatial, temporal or spatiotemporal domain to an attribute range associated with one of the elements in a regular array of points or cells. These concepts are not mutually exclusive, e.g. many landscape components may be viewed alternatively in terms of vector or coverage according to what best suits their specific discipline or sub-discipline. In this work we have followed the INSPIRE data specifications for modelling the agricultural and aquaculture geospatial features. The model (spatial object type) adopted by INSPIRE is the feature type. The term “feature type”³ is the formal term defined by ISO (ISO, 2005) for what is generally called the “vector data model”. INSPIRE defines the “feature type” with a stereotype⁴ (INSPIRE, 2012a). This specification standardises the spatial object type, but it does not assume that two engineers will apply them

identically to model the same information. Therefore, some substructures of the feature type structure must be added to the model. We decided to conceptualise the recognised features using simple, primitive geometric shapes defined by the “simple feature model” proposed by OGC (Herring, 2011), i.e. for the features “structure”, “plot”, “open area” and “site” the geometry adopted is a polygon and for the holding feature the geometry is a point. With the OGC “simple feature model”, the specification of the envelope (the geometric primitive that defines the spatial extent of the phenomena recognised within the “project world” by means of a bounding box), the type of geometry (the OGC “simple feature model” uses the “well-known types” (WKTs) of geometry for representing each feature type, WKT is a text mark-up language for representing vector geometry objects) and the spatial reference systems (the detailed model for spatial referencing by coordinates can be found in the OGC Abstract Specification reference (Cooper, 2010)) are explicit. Another required specification is the meta-information that allows the feature information to be understood and exploited. As meta-information schema we adopted the INSPIRE metadata schema (EC no. 1205/2008 of 3 December 2008). The last required specification is a list of attributes that are associated with each feature type. Tables 1-5 present the grammar for each feature recognised in our model (as far as possible we adopted the same attribute-value pairs grammar proposed by the INSPIRE AF specification).

To complete the “project world” schema, it is necessary to describe the data capture criterion for the recognised features. With the caveats that the information imported must be managed carefully to avoid overlaps and inconsistencies with information stored in existing datasets, the following list presents the types of capture that can be adopted:

Cadastral data or land registry (the registry that

³A specific feature type representing a class of real world phenomena (e.g. road) expressed in an application schema. The feature type, together with the attributes, the relationships and the behaviour of the feature are the four aspects of the general model for features defined by ISO 19109:2005.

⁴A stereotype can be thought of as a “template” to be populated at each point in time of a feature instance corresponding to that type. A complete description of the INSPIRE stereotype is available in the INSPIRE Web portal (<http://inspire.ec.europa.eu/schemas/>).

Table 1. Site attributes.

Feature type: site		
Property name	Value type	Description
Geometry	Polygon	A polygon is a planar surface defined by one exterior boundary with none or more interior boundaries.
Activity	Data type	Classification of the economic activity of the holding, according to the NACE rev. 2.0 coding (e.g. livestock production or aquaculture).
Data type	Data type	Identifies the group of animals of the same species kept on the specific holding. This attribute is related to the “Farm animal species” Data type* that contains the code lists** of animal names.
Type of capture	Data type	Identifies the type of capture adopted. This attribute is related to the “Type of capture value”*** code list that contains list of the types of capture.
Valid from	Date	Time when the site opened.
Valid to	Date	Time when the site closed.

*A structured type of data without identity. Specification of the Data type stereotype adopted by INSPIRE is available at: <http://inspire-twg.jrc.ec.europa.eu/datamodel/approved/r937/>; **A class that provide an agreed set of values with multi-lingual names, definitions and descriptions to be re-used as values of properties. INSPIRE-governed code lists are available in the INSPIRE code list register at <http://inspire.ec.europa.eu/codeList/<CodeListName>>; ***A code list to be made available by the IZSve as soon as maintenance workflows have been formally defined according to the Technical Guidelines reference provided by INSPIRE.

Table 2. Structure attributes.

Feature type: structure		
Property name	Value type	Description
Geometry	Polygon	A polygon is a planar surface defined by one exterior boundary with none or more interior boundaries.
Type of structure	Data type	Identifies the type of construction. This attribute is related to the “Animal housing construction type” code list that contains a list of the construction types.
Type of capture	Data type	Identifies the type of capture adopted. This attribute is related to the “Type of capture value” code list* that contains list of the types of capture.
Valid from	Date	Time when the site opened.
Valid to	Date	Time when the site closed.

*To be made available by the IZSve as soon as maintenance workflows have been formally defined according to the Technical Guidelines reference provided by INSPIRE.

Table 3. Plot attributes.

Feature type: plot		
Property name	Value type	Description
Geometry	Polygon	A polygon is a planar surface defined by one exterior boundary with none or more interior boundaries.
Type of capture	Data type	Identifies the type of capture adopted. This attribute is related to the “Type of capture value” code list* that contains list of the types of capture.
Valid from	Date	Time when the site opened.
Valid to	Date	Time when the site closed.

*To be made available by the IZSve as soon as maintenance workflows have been formally defined according to the Technical Guidelines reference provided by INSPIRE.

Table 4. Open area attribute.

Feature type: open area		
Property name	Value type	Description
Geometry	Polygon	A polygon is a planar surface defined by one exterior boundary with none or more interior boundaries.
Type of capture	Data type	Identifies the type of capture adopted. This attribute is related to the “Type of capture value” code list* that contains list of the types of capture.
Valid from	Date	Time when the site opened.
Valid to	Date	Time when the site closed.

*To be made available by the IZSVE as soon as maintenance workflows have been formally defined according to the Technical Guidelines reference provided by INSPIRE.

manages cadastral parcels, i.e. single areas having homogenous real estate rights and unique ownership defined by national laws). A direct, spatial relation can be established with the polygons managed by this registry. It is possible to verify the existence of a specified, topological, spatial relation between two geometric objects (Herring, 2011), a situation that allows extraction (and eventually transformation) of the geographical information managed by the registry datasets and its integration into our “structure” and “plot” feature system. However, it is important to highlight that the “structure” and “plot” features remain independent from the registry datasets because the relation is only

established at the spatial level, and there is thus no direct association between the features of our model and the registry datasets.

Land use maps (charts depicting partitions of a territory characterised according to their functional dimensions or socio-economic purposes, e.g. residential, industrial, commercial, agricultural, forestry and recreational). A direct spatial relation might be established with objects shown on large-scale (1:1,250, 1:2,500, 1:5,000 and 1:10,000) land use maps. In case these maps are digital equipped with external object identifier, which is a unique identifier, they may be used by external applications to reference spatial

Table 5. Holding attribute.

Feature type: holding		
Property name	Value type	Description
geometry	Polygon	A polygon is a planar surface defined by one exterior boundary with none or more interior boundaries.
Activity	Data type	Classification of the economic activity of the holding, according to the NACE rev. 2.0 coding (e.g. livestock production or aquaculture).
Includes animal	Data type	Identifies the group of animals of the same species kept on the specific holding. This attribute is related to the “Farm animal species” data type that contains the code lists of animal names.
Type of capture	Data type	Identifies the type of capture adopted. This attribute is related to the “Type of capture value” code list* that contains list of the types of capture.
Veterinary code	String	Identifies the veterinary code of the holding according to a scheme provided by the “Holding registry” (92/102/CEE).
Valid from	Date	Time when the site opened.
Valid to	Date	Time when the site closed.

*To be made available by the IZSVE as soon as maintenance workflows have been formally defined according to the Technical Guidelines reference provided by INSPIRE.

objects (INSPIRE, 2012a). If so, it would also be possible to establish a spatial relation and extract (as well as transform) available geographical information from them and integrate the data into the proposed “structure”, “plot” and “open area” features. In case they are digital-raster maps, data capture can be carried out as described in the following paragraph.

Digitizing maps. The polygon that represents the boundary of “structure”, “plot”, “open area” and “site” features or the point that represents the “holding” feature can be digitised (identification and extraction of spatial data from a digital images) from raster maps, satellite images, ortophotos, etc. The process of digitising can either be manual or automatic. In the latter case, an object-oriented image detection application is required. Several different type of software for a generalist automatic detection and description of objects from RGB images (a color model in which red, green and blue are combined together to reproduce different color images) are available, e.g. ERDAS IMAGINE (<http://www.intergraph.com/>) and eCognition software (<http://www.ecognition.com/>).

Spatial function. The geometry of the “site” and “holding” features can be obtained from a spatial function as follows:

- (i) the point that represent a “holding” can be obtained from the centroid function applied to the “site” feature (in case the “holding” is composed by “sites” that are not geographically continuous, it would be valid to provide the geographical information related to the principal one);
- (ii) the boundary of “site” feature can be obtained by the “convex hull” function applied on the “structures” or “plots” or “open areas” features that comprise the “holding area”;
- (iii) the boundary of the “open area” can be expressed as a buffer around a given feature (e.g. a line that represent the path of a nomad flock or the point that represent the position of a bee-hive); and
- (iv) the “site” extent can be represented by a buffer function centre on the “holding point”, with a radius that is a function parameterised on the species and number of animals kept in the holding.

Geocoding. The “holding point” location can be indirectly obtained through the address (by geocoding provided in a database for livestock registration (Margoulies, 2001). In case of a holding composed by different, not geographically continuous structures, it would be valid to provide the geographical information related to the main one. Geocoding capture can

provide a wrong geospatial location in case the source information provided is incorrect (e.g. legal address instead of facility address).

On-site survey. Global positioning systems (GPS) technology can be used to obtain the spatial coordinates of a given location. These points can be used to represent specific elements of the holding features (FitzMaurice et al., 2007). Many veterinary field staff use GPS receivers to obtain geographical coordinates for routinely-inspected facilities. We suggest limiting the use of GPS to exclusively collect the coordinates of locations that are difficult to recognise by cartographic support, either because the scale of the map is too large or it does not represent the elements that are necessary to depict the feature in question.

According to the ISO 19109:2005 standard, the unified modelling language (UML) should be the conceptual language for application schemata. The UML diagram is widely used and it is accepted by the scientific community as the tool to design and specify spatial conceptual models (Lisboa-Filho et al., 2010).

In the proposed UML application schema there are three main interconnected conceptual blocks: the generalised part, the observable part and the external reference part. The generalised part (“holding” and “site” features) is concerned with the synthetic geographical representation of the “holding area”. It is the simplified version of the INSPIRE core model. The observable part (“structure”, “plot” and “open area”) is dedicated to animal confinement and designed to record aspect of “real world” entities that are considered of veterinary interest. Although the recognised features must be considered independent with their own geographical definition, a direct relation can be established with other geographical elements. The external reference part (cadastral parcel and existing land use object) can be exploited by using the property of spatial object that asserts the existence of topological spatial relationship between two geometric objects as they would be represented on a map (Herring, 2011). This relation could become a source of error in the accuracy of the geoposition of the elements included in the datasets and one must therefore be careful to avoid overlaps and inconsistencies with information stored in existing datasets. Included in the external reference part is also the relationship with the hydrographic system. In case the “structure” is a water installation with a water inlet or outlet device, it can be connected to a “hydro object”, identified by the INSPIRE theme hydrography (INSPIRE, 2007).

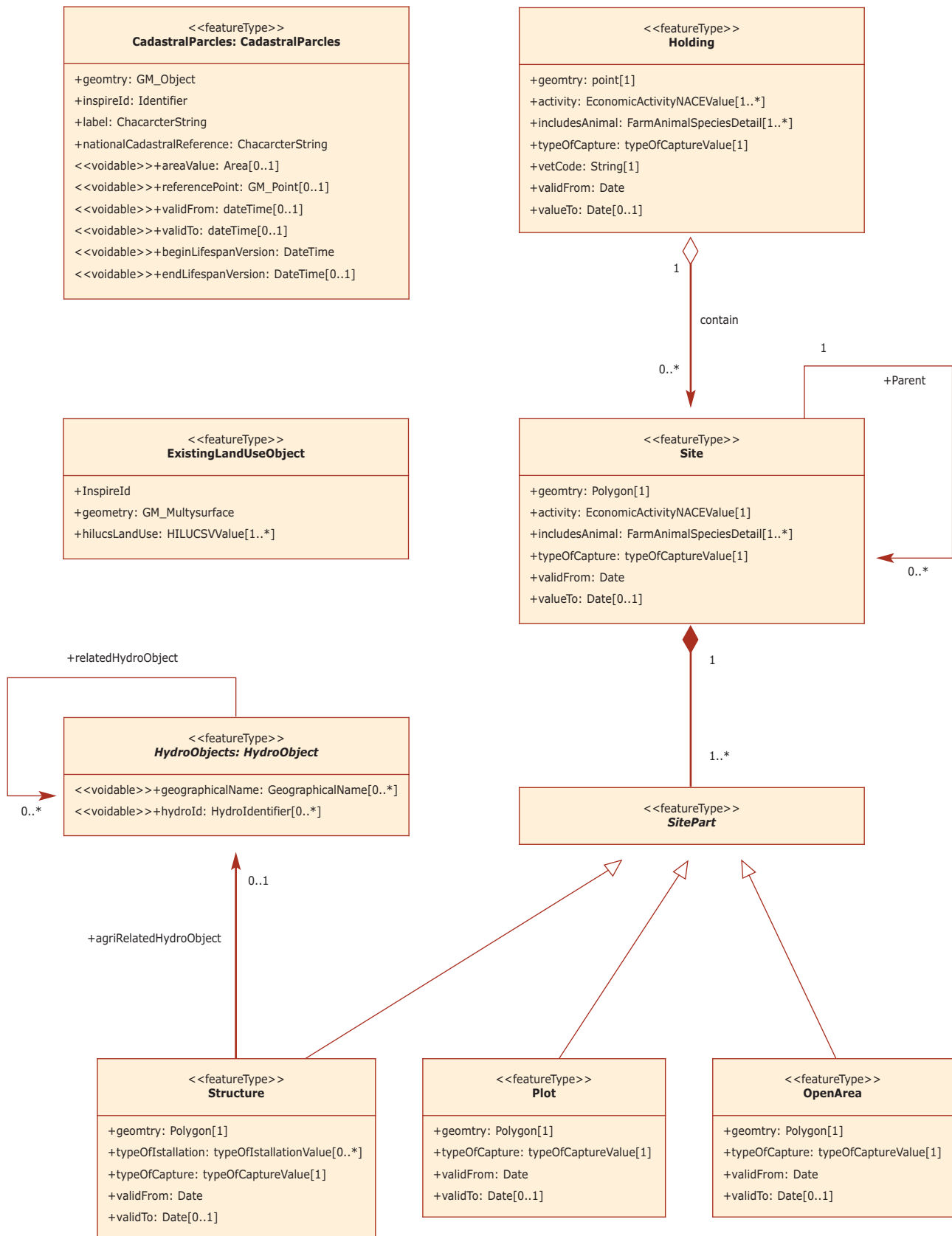


Fig. 4. UML class diagram.

Conclusion

The overriding aim of this paper was to achieve a generic and inter-operable schema that relies on geographical and veterinary concepts. It was designed to serve as the base for implementation of veterinary spatial data with known locations. We expect this application schema to be sufficiently generic to embrace all types of animal holding and fulfil the implementation of veterinary spatial data in the INSPIRE directive. We further anticipate that it will signal the development of inter-operability rules leading to a common understanding (vocabulary and semantics), so that harmonisation and sharing of veterinary spatial datasets can be achieved. We also hope that the open nature and abstraction level of the proposed model (i.e. the model provides a high-level representation and description of the structure and content of the holding geographical information without the need for mentioning GIS software or databases) will facilitate the building of bridges between datasets created and maintained during daily research activities. Using INSPIRE data specifications holds great potential for supporting policies and activities that impact both the environment and the veterinary activities.

Acknowledgements

This work has been supported by the Italian Health Ministry Ricerca finanziata dal Ministero del Lavoro, della Salute e delle Politiche Sociali Dipartimento per la Sanità Pubblica Veterinaria, la Nutrizione e la Sicurezza degli Alimenti.

References

- Atkins, TripleDev, 2012. E-PRTR Validation Tool. EOINET. Available at: <http://www.eionet.europa.eu/schemas/eptrtr/EPTRUserManual.pdf> (accessed on December 2012).
- Buehler K, McKee L, 1998. The OpenGIS Guide: introduction to interoperable geoprocessing and the OpenGIS specification. Waltham MA, OpenGIS Consortium.
- Burrough PA, McDonell RA, 1998. Principles of geographical information systems. New York: Oxford University Press.
- Cliff K, Carl R, 2009. The OpenGIS abstract specification topic 5: features. Available at: <http://www.opengeospatial.org/standards/as> (accessed on December 2012)
- Cooper P, 2010. Topic 2: spatial referencing by coordinates version 4.0 (p. 100). Available at: <http://www.opengeospatial.org/standards/as> (accessed on December 2012)
- Durr PA, Froggatt AEA, 2002. How best to geo-reference farms? A case study from Cornwall, England. *Prev Vet Med* 56, 51-62.
- Ferrè N, Mulatti P, Mazzucato M, Lorenzetto M, Trolese M, Pandolfo D, Vio P, 2011. GeoCREV: veterinary geographical information system and the development of a practical sub-national spatial data infrastructure. *Geospat Health* 5, 275-283.
- FitzMaurice PL, Freier JE, Geter KD, 2007. Geospatial field applications within United States. *Vet Ital* 43, 477-482.
- Frank AU, 2013. Ontology for spatio-temporal databases. Available at: http://www.spatial.cs.umn.edu/Courses/Fall11/8715/papers/STD12_chrono.pdf (accessed on December 2013).
- Frank AU, Goodchild ME, 1990. Two perspectives on geographical data modelling. Available at: <http://www.geog.ucsb.edu/~good/papers/130.pdf> (accessed on December 2013)
- Gibson CC, Ostrom E, Ahn TK, 2000. The concept of scale and the human dimensions of global change: a survey. *Ecol Econ* 32, 217-239.
- Goodchild ME, 2011. Scale in GIS: an overview. *Geomorphology* 130, 5-9.
- Groot R, McLaughlin J, 2000. Geospatial data infrastructure: concepts, cases, and good practice. USA: Oxford University Press, 318 pp.
- Herring J, 2011. OpenGIS implementation standard for geographic information - Simple feature access - Part 1: common architecture. OGC document. Available at: <http://www.opengeospatial.org/standards/as> (accessed on December 2013).
- INSPIRE, 2007. INSPIRE data specification on hydrography - Guidelines. Available at: <http://inspire.ec.europa.eu/index.cfm/pageid/2> (accessed on January 2013).
- INSPIRE, 2012a. Drafting team on data specification version 3.4. INSPIRE generic conceptual model. Available at: <http://inspire.ec.europa.eu/index.cfm/pageid/2> (accessed on January 2013).
- INSPIRE, 2012b. TWG agricultural and aquaculture facilities. INSPIRE - data specification on agricultural and aquaculture facilities - Draft guidelines. D2 . 8 . III . 9. Available at: <http://inspire.ec.europa.eu/index.cfm/pageid/2> (accessed on January 2013).
- ISO, 2005. 19109 Geographic information - Rules for application schema.
- ISO/IEC, 1998. Information Technology (IT) - Conceptual Schema Modelling Facilities (CSMF). Available at: http://www.aim.nl/weblog/19981200%20ISO_FCD_14481.php (accessed on December 2012)
- Kotiadis K, Robinson S, 2008. Conceptual modelling: knowledge acquisition and model abstraction. *Proceedings of the 2008 Winter Simulation Conference*, 951-958 pp.
- Kuhn W, 2001. Ontologies in support of activities in geographical space. *Int J Geogr Inf Sci* 15, 613-631.
- Kuhn W, 2005. Geospatial semantic: why, of what, and how. *Data Semantics III*, 1-24.

- Lisboa Filho J, Sampaio GB, Nalon FR, Borges K, 2010. A UML Profile For Conceptual Modeling in GIS Domain CAiSE 2010 Workshop. Hammamet, Tunisia, 18-31, DE@CAiSE'10.
- Margoulies S, 2001. Geocoder service specification. Available at: <http://www.opengeospatial.org/standards/as> (accessed on December 2012).
- Mustiere S, Smaalen J Van, 2007. Database requirements for generalisation and multiple representations. *Generalisation of Geographic Information*, 113-136 pp.
- Nativi S, Federici G, 1994. A conceptual modelling for the GIS developing. Fifth European Conference and exhibition on geographical information system – ESRI/MARI. Available at: <http://libraries.maine.edu/Spatial/gisweb/spatdb/egis/eg94102.html> (accessed on February 2013).
- Norstrøm M, 2001. Geographical information system (GIS) as a tool in surveillance and monitoring of animal diseases. *Acta Vet Scand* 42, S79.
- OGC, 1999. The OpenGIS™ abstract specification topic 10: feature collections (p. 13). Available at: <http://www.opengeospatial.org/standards/as> (accessed on December 2012).
- Percivall G, 2003. OGC reference model. Available at: <http://www.opengeospatial.org/standards/as> (accessed on December 2012).
- Pfeiffer DU, Robinson TP, Stevenson M, Stevens KB, Rogers DJ, Clements ACA, 2008. *Spatial analysis in epidemiology*. Oxford University Press, 142 pp.
- Sanson RL, Pfeiffer DU, Morris RS, 1991. Geographic information systems: their application in animal disease control. *Revue scientifique et technique (International office of epizootics)* 10, 179-195. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1760572> (accessed on February 2013).
- Smith B, Varzi AC, 2000. Fiat and Bona fide boundaries. *Philos Phenomenol Res* 60, 401-420.
- Tóth K, Portele C, Illert A, Lutz M, de Lima MN, 2012. A conceptual model for developing interoperability specifications in spatial data infrastructures. JRC, European Union, pp. 57.
- Worboys M, Duckham M, 2004. *GIS: a computing perspective*. Taylor and Francis, London, 494 pp.