European conference of the Czech Presidency of the Council of the EU
TOWARDS eENVIRONMENT
Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe
http:/www.e-envi2009.org/proceedings/
J. Hřebíček, J. Hradec, E. Pelikán, O. Mírovský, W. Pillmann, I.Holoubek, T. Bandholtz (eds.)
Masaryk University, 2009

An Holistic View of Coverage Model and Services for SISE-SEIS

S. Nativi[a] **and A. Woolf**[b]

[a] *Italian National Research Council - IMAA, and Univ. Florence at Prato, Prato IT 59100, Italy (stefano.nativi@cnr.it)*

[b] *STFC Rutherford Appleton Laboratory, Oxon OX11 0QX, UK (andrew.woolf@stfc.ac.uk)*

Abstract: There has recently been a growing recognition of the need for interoperability – not only in traditionally 'geographic' disciplines (e.g. mapping, physical geography) but also in more science-focussed disciplines like Earth systems science, geology, meteorology, oceanography. These latter application domains bring a fresh viewpoint to the nature of spatial information. While the conventional geographic disciplines have a primary focus on a 'feature' view (points/lines/polygons with thematic attributes), the scientific domains utilise a conventional 'coverage' view (thematic properties distributed over some region in space and/or time). For integrated information systems like SEIS and SISE, however, a more harmonised model is required. Fundamentally, the feature- and coverage-views are complementary, and any general framework must support both viewpoints. A framework for harmonisation is offered through the observation process that lies at the heart of much Earth science data. By recognising that an observation process samples the natural environment (which may be represented through a feature view) and generates a coverage result, we are able to reconcile these hitherto different approaches. This harmonized model supports the different types of discovery and access services useful to serve the heterogeneous SEIS use cases.

Keywords: SDI; Coverage model; Observation & Measurement model, Feature model.

1. INTRODUCTION

The growing area of environmental informatics is concerned with providing integrated access to a range of advanced information and processing resources for the environment. Both the US and European premier scientific unions are recognising this – the American Geophysical Union (AGU)[1] has established an Earth and Space Sciences Informatics (ESSI) Focus group, and the European Geosciences Union (EGU)[2] has recently created a new scientific Division for ESSI. Primarily, interoperability and metadata are identified as two key technologies in discovering and enabling access to usable information and processing resources for the environment. Interoperability is achieved by adopting a Service-Oriented Architecture (SOA) approach and applying international standards, as far as service interfaces and data models are concerned. International standard organizations,

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

[2] http://www.egu.eu/

such as ISO TC211[3] and OGC[4], have been working to develop such a standards baseline for the geospatial information domain. This approach has been supported and adopted by important Spatial Data Infrastructure (SDI) initiatives such as the European INSPIRE[5].

Earth science is at the forefront of applying advanced computing technology to the solution of pressing environmental problems. GMES and GEOSS initiatives represent a couple of significant cases in point. In fact, important environmental problems demand an integrated modelling of coupled physical processes, global datasets, and a multi-disciplinary coordinated approach (e.g. biologists working together with climatologists to determine the impact of warming on species distributions).

However, the information modelling approach used by the SDI initiatives and GMES and GEOSS is novel – for the most part – to the environmental and earth science communities, and there are significant challenges in bridging the conceptual gap. In earth science there have been some applications of this modelling approach – notably in the geosciences and atmosphere/ocean. A valuable example is represented by the OGC GALEON IE (Geo-interface for Air, Land, Earth, Oceans NetCDF Interoperability Experiment)[6] which deals with specifying and using standard interfaces to foster interoperability between data systems used by the traditional GIS community and those in the community referred to as the Fluid Earth Sciences (FES, mainly oceanography and atmospheric science). These attempts to bridge legacy information environments in earth science with SDI have both proven the general feasibility and identified weaknesses and challenges.

Holistic interdisciplinary approaches and lack of common data models and semantics are important research challenges to be addressed in order to achieve a Single Information Space for the Environment (SISE) in Europe [Juceviciene, 2008].

Geospatial information models for interoperability recognize three important concepts: feature, coverage and the more general map. ISO TC211 introduced two fundamental concepts to map both discrete and continuous real world phenomena: features and coverages. For Earth Sciences, a 'coverage' or field view of information is very predominant – much earth science data is regarded as a field over some region of space and/or time, rather than a discrete spatial object with attributes. Multiple coverage types exist; mainly, they are disciplinary related. Virtually, any geospatial data may be viewed as an instance of a coverage type. Different coverage types are characterized by different coverage domains and coverage functions.

There exists a clear need to develop an holistic approach to model, discover, access and use environmental coverage data types. For SEIS-SISE, it is particularly important to rely on effective and flexible models and service interfaces for coverage datasets. This holistic approach must be clearly harmonized with the General Feature Model (GFM) adopted by the international standardization frameworks for geo-information –e.g. ISO and OGC.

The manuscript is structured as follows. The first section briefly discusses the coverage concept as introduced and used by standard data models for geo-information, outlining the different coverage types that characterize important geospatial communities. Having in mind these different views, the second section discusses an holistic coverage and feature model for SEIS-SISE and the need for one or more types of coverage access service. The final section summarizes the manuscript conclusions.

2. THE COVERAGE CONCEPT

The coverage concept was defined to summarise the different conceptual and physical representations of a traditional image, going further by enlarging the spectrum of geospatial information that can be represented this way. Hence, the "coverage" term refers to any data

[3] http://www.isotc211.org/

[4] http://www.opengeospatial.org/

[5] http://inspire.jrc.ec.europa.eu/

[6] https://sites.google.com/site/galeonteam/Home

representation that assigns values directly to spatial position. Thus, a coverage may be seen as a function from domain –commonly a spatial, temporal or spatiotemporal domain– to an attribute range. A coverage associates a position within its domain to a record of values of defined data types. According to ISO TC211, coverage is a subtype of feature that has multiple values for each attribute type, where each direct position within the geometric representation of the feature has a single value for each attribute type [ISO 19123].

Therefore, a coverage may be generally modeled through: a domain (the well-defined set of direct position the coverage deals with); a co-domain or attributes range; a coverage function (the rule that associates each element from the domain to a unique element in the co-domain).

2.1 Coverage Types

By way of example, we consider a number of coverage types that are of widespread interest within the oceanographic and meteorological communities.

As with many earth science disciplines, these communities collect data through observational campaigns or deployment of monitoring instruments. Both the measured values and the location/time at which they are measured are important – fundamentally this represents a coverage view. Moreover, it is usual for practitioners to classify coverage types into classes based on the geometry and topology of the discrete coverage domain.

There are very good scientific reasons why this should be so. Physical processes occur in the natural world across a wide spectrum of spatial and temporal scales, and considerable science informs the design of experimental sampling strategies. Conversely, the geometry and topology of observation sets are a fundamental determinant of the scientific uses to which they may be put. Moreover, the properties of the instruments used to generate data themselves place constraints on their interpretation (e.g. as regards accuracy, precision, calibration, required post-processing, etc.). These two factors – the scientific utility of a sampling regime, and the limitations of an observing process – lead to a natural, scientifically important, classification of data types along these axes. Quite often the two are highly correlated (certain instruments generate certain samplings), and so scientific communities of practice adopt more abstract conceptual information classes that nevertheless reflect artifacts of sampling or instrument-type. Within the meteorological and oceanographic communities, broad information classes based on measurement-set geometry and topology have almost universal acceptance. We next describe some examples:

For example, the US National Oceanographic and Atmospheric Administration (NOAA) is developing a plan for a Global Earth Observing Integrated Data Environment (GEO-IDE) to integrate measurements, data and products and create interoperability across data management systems. In the GEO-IDE Concept of Operations[7], the following ‘structural data types’ are defined: *Grids, Moving-sensor multidimensional fields, Time series, Profiles, Trajectories, Geospatial Framework Data, Point Data, Metadata*.

The ESRI ‘ArcMarine’ Data Model for marine data includes classes like *Instant, Location Series, Time Series, Profile Line, Track, Sounding, Survey, {Ir}Regularly Interpolated Surfaces, Mesh Volume, etc*. File formats such as netCDF and NASA Ames utilize data models that reflect these structures (e.g. netCDF four-dimensional gridded lat-lon-height-time variables, or NASA Ames time-series at a point). The netCDF Common Data Model (CDM) and the Climate Science Modelling Language (CSML) adopt very similar classifications, Table 1.

3. AN HOLISTIC APPROACH

In an object-oriented approach every business entity is an object type. Analogously, in the ISO feature-based approach every business entity is a feature type. Actually, in this case

[7] https://www.nosc.noaa.gov/dmc/docs/NOAA_GEO-IDE_CONOPS-v3-3.pdf

they are geographical features: an abstraction of a real-world phenomenon that is associated with a location relative to the Earth [ISO 19107].

Therefore, in the ISO GFM framework a coverage is simply harmonized treating it as a feature sub-type; this is often implemented by defining “logic” features rather than physical ones. The sub-typing connection is quite general. Hence, there is a clear need to investigate such association in order to outline an holistic approach for implementing a more effective and flexible harmonization model. The next chapters will investigate this issue proposing a model for an harmonization solution.

Table 1: Comparison of CDM and CSML v2 feature types [Caron, 2008]

CSML Feature Type	CDM Feature Type
PointFeature	PointFeature
PointSeriesFeature	StationFeature
TrajectoryFeature	TrajectoryFeature
PointCollectionFeature	StationFeature at fixed time
ProfileFeature	ProfileFeature
ProfileSeriesFeature	StationProfileFeature at one location and fixed vertical levels
RaggedProfileSeriesFeature	StationProfileFeature at one location
SectionFeature	SectionFeature with fixed number of vertical levels
RaggedSectionFeature	SectionFeature
ScanningRadarFeature	RadialFeature
GridFeature	GridFeature at a single time
GridSeriesFeature	GridFeature
SwathFeature	SwathFeature

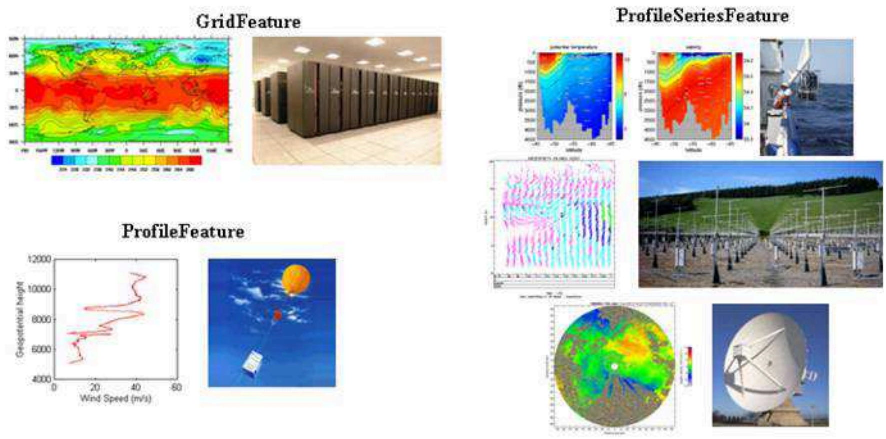

Figure 1: Illustration of some coverage types

3.1 The General Conceptual View for Geo-information

Geoinformation is used to describe objects, phenomena or processes related to the Earth [Molenaar, 1984]. This is commonly done in the form of state description at a certain moment or, for processes, as a series of state descriptions. These may refer to one or more aspects characterizing terrain objects (geographical features). These aspects are given as thematic attributes in relation to object shapes (geometric data). In the most straightforward form the thematic attributes are directly linked with positional data –in a given N-dimensional domain. Hence, the positional data serve as a vehicle to link different types of thematic data or to link data obtained at different moments. A higher information level can be obtained by the introduction of terrain objects (i.e. features class instances) [Molenaar, 1984]. In fact, thematic attributes are not linked directly to the positional data, but to the terrain objects (feature instances).

These two approaches reproduce two traditional and different space conceptualizations: the object and the field views (Couclelis, 1992; Goodchild, 1992; Peuquet, 1984). The object approach –sometimes called as geo-relational, boundary or "bird-eye view"– considers space as being 'empty' and populated with discrete entities embedded in space, while the field approach –sometimes called as tessellation, direct-position or "worm-eye view" – considers the space as being continuous, and every location in space has a certain property (Ledoux, 2008).

The ISO GFM approach adopts the object/geo-relation approach, focusing on feature classes and instances definition and description; geometric data are structured using feature types (i.e. geometric objects). While, the ISO coverage realm refers to the field/direct-position approach, focusing on the definition and description of coverage domains and mapping functions (i.e. coverage functions). Tessellation or polygonal mesh types are introduced to manage data structures –realizing commonly used implicit geometries, such as regular grids. Figure 2 depicts a general harmonization model for the two approaches.

Geographical features are characterized by a geometry (i.e. shape) and a thematic description. The feature sets defined through the geometric characteristics are called "feature types"; while, the sets defined through the thematic characteristics are called "feature classes".

Locations of a field domain are mapped to thematic attributes applying a coverage-function. A domain is comprised of a set of direct positions. This set may be infinite – realizing continuous coverage; commonly, the set consists of a finite collection of points or geometric objects (e.g. tiles); they locate the samples or "ground truth" specimens of a field.

It is possible to find the spatial structure of a coverage thorough its positions topology. The topology is built up through the connectivity of neighboring domain elements (e.g. tiles). Neighboring coverage domain elements are connected within a feature instance if they have the same attribute values (Molenaar, 1991). Feature classes geometries are represented by connected domain elements –see Figure 2. This leads to "recognize" and extract feature class instances from coverages. The depicted schema supports also the other way around –i.e. to get coverage from feature class instances.

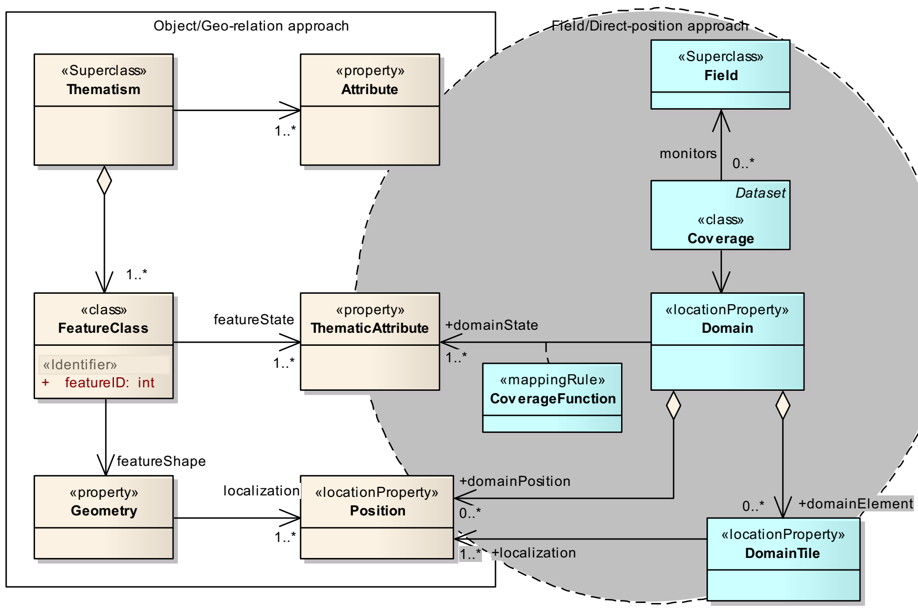

Figure 2. General approaches for Geo-information modelling

3.2 The harmonization context: a new view

In the SISE-SEIS context, an effective harmonization model can be motivated by considering the process of observing or measuring the environment. This is captured by the OGC/ISO Observations and Measurements (O&M) conceptual model [ISO 19156, 2008]. In natural language, the model states (see the dotted box in Figure 3) that an *Observation* is an action whose result is an estimate of the value of some *Property* of a *FeatureOfInterest* (FOI) obtained using a specified *Procedure.* Considering the general geo-information concepts previously discussed, the O&M model can be easily used to develop an harmonization framework for Environmental information systems. Figure 3 shows the context view of this harmonization model.

According to the O&M specification, *the FOI of an observation may be any feature having properties whose values are discovered by observation. In general, this will be of a type from catalogue representing the application domain for an investigation.* Normally, the FOI will be a so-called 'domain feature', representing an identifiable real-world spatial object. However, important for harmonizing feature and coverage views is to consider the case where the FOI exists only for the purpose of 'sampling' the physical environment (the *SamplingFeature* class of Figure 3). *SamplingFeature* is a FOI which may realize several observations concerning any identifiable feature. Examples include a weather balloon measuring temperature as it ascends through the atmosphere, or a moored tide-gauge measuring sea-level time-series at a location within a harbor. In those cases the observations sample a field (see Figure 1); thus, they are modeled through a *FieldSampling* class in Figure 3. In many of these cases, the result of the observation is a discrete Coverage (see Figure 1 and Figure 3) – a set of attribute values (e.g. temperature, sea-level) over a spatial, temporal, or spatiotemporal domain (e.g. the trajectory of a weather balloon, or the time instants of recorded sea-level). This realization provides the key step towards a model for harmonizing feature and coverage views.

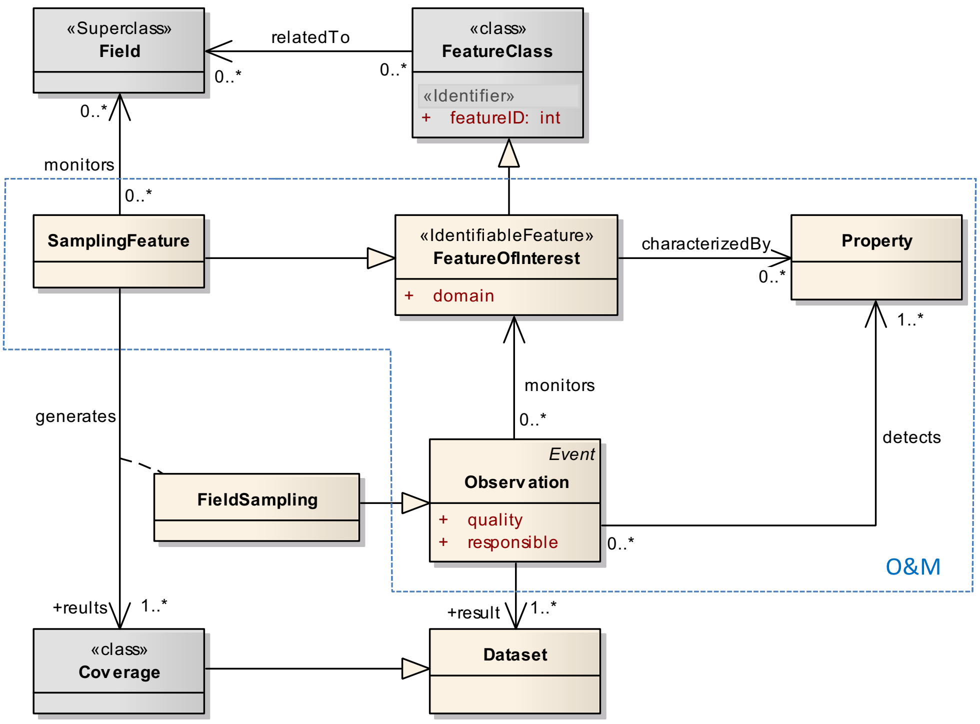

Figure 3 Harmonization model: context view

3.3 The Harmonization Model

We outlined above the key general pattern that recurs in many applications of O&M in environmental applications: the *FeatureOfInterest* is a *SamplingFeature*, and the result of the Observation is a Coverage. This realization provides a mechanism to integrate feature and coverage views through a harmonized model utilizing the O&M framework. Recalling the O&M model, an essential constraint is that the Observation result must be consistent with the observed property. For the case of a *SamplingFeature*, therefore, we propose that the observed property (*Phenomenon* class, Figure 4) is equivalent to the range of the Coverage result (or more strictly the semantic rangeType of the ISO 19123 CV_Coverage).In addition, we recognize that the *SamplingFeature* is intended to incorporate geometric aspects of the sampling regime, and propose that it should be equivalent to the domain geometry of the Coverage result (see again Figure 4). In the specific case of a coverage domain sampled through a set of tile elements, the survey procedure includes tessellation implementation information (*TesselationProcedure* call, Figure 4), such as: tile shape, geometry type, distribution, etc.

Thus, we arrive at a harmonization model for integrating feature and coverage views: there exists a real-world ‘ultimate’ FOI having properties that may be represented as continuous coverages. An example would be ‘The Atmosphere’ having a property ‘temperature’ that is a continuous coverage over a four-dimensional (x-y-z-t) domain. However, in practice we are limited to observing this ultimate domain feature only at discrete locations – the FOI in this case is not ‘The Atmosphere’ but rather a *SamplingFeature* that exists only to provide a concrete focus for the observation. The result of such an observation is a discrete Coverage (which may be classified according to broad classes of geometry and topology, as discussed in section 2.1 earlier). The domain geometry of this coverage is reflected in the geometry of the *SamplingFeature*, and the range of the coverage is an implied thematic property of the *SamplingFeature* (providing a consistent semantic closure for the observed property).

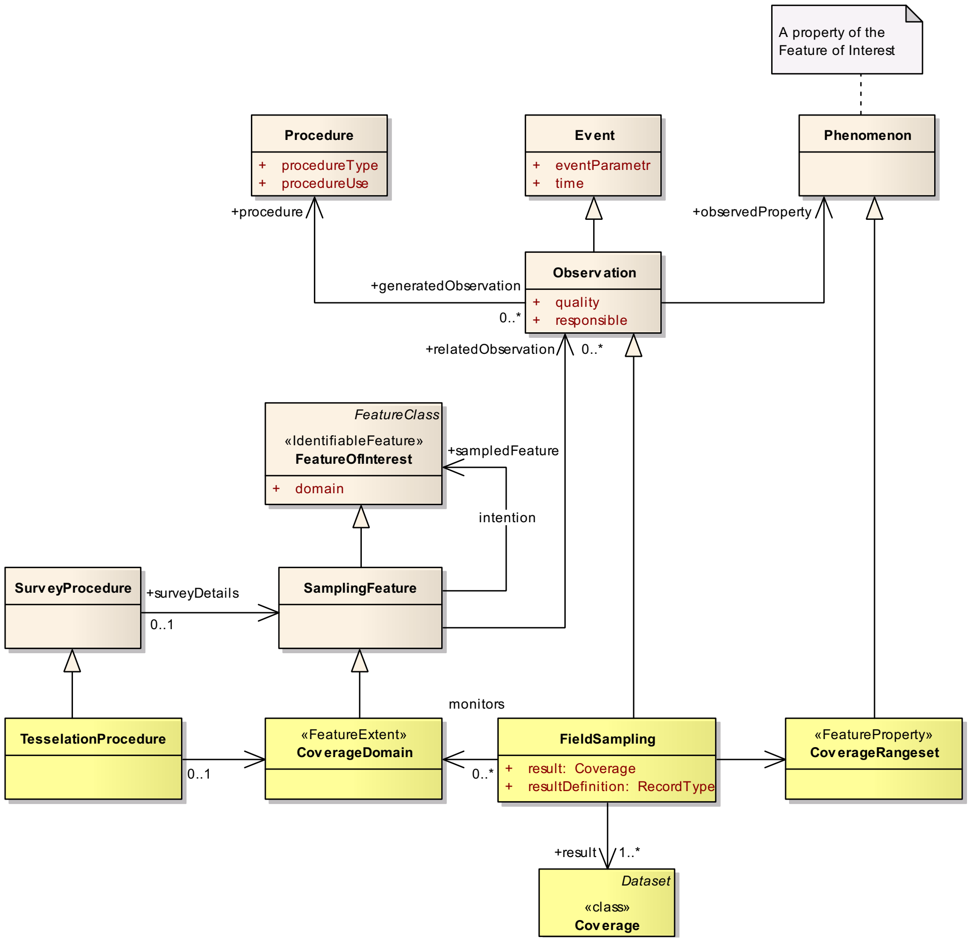

Figure 4 Harmonization model

3.2 Access services for SISE-SEIS

The harmonized model supports both the traditional feature based (e.g. FOI) and coverage based (e.g. sampled datasets) access services for an environmental information framework. For instance, standard service interfaces like the Sensor Observation Service or Web Feature Service may be more suited to the 'feature' view, while the Web Coverage Service may be more suited to the 'coverage' view. Besides, it enables the implementation of an advanced type of access services which make use of the traditional ones. These services provide the access to high level artifacts (e.g. phenomena, event, observation, procedures, etc.) avoiding to discriminate about the structure of their instances and representations. These two levels of services are extremely important to address the use cases heterogeneity characterizing complex and large "system of systems" such as SISE-SEIS.

4. CONCLUSIONS

For SISE-SEIS there exists a clear need to develop an holistic approach to model, discover, access and use environmental coverage data types. This holistic approach must be clearly harmonized with the ISO GFM. Coverage and GFM reproduce the traditional and different space conceptualizations which characterize geoinformation: the object and the field views.

We proposed a possible context to harmonize these two different views. This context is then implemented using the OGC/ISO O&M model. In fact, we recognize that in many cases with *SamplingFeatures* we get a Coverage result. Actually, the Coverage domain is the SamplingFeatures and the Coverage range is a feature property.

For a Shared Environmental Information System, these complementary views are important, and it is important also to realize that they do not conflict but rather represent alternative representations of the real world. A 'feature view' regards a feature class having thematic attributes that are themselves coverages, while a 'coverage view' considers just those attributes themselves. These two views are integrated through the act of observation: extensive properties or attributes of a real-world feature are sampled (by a 'sampling feature') leading to a discrete coverage result with the domain geometry equivalent to the sampling regime. A large number of environmental thematic areas (e.g. within Annex II and III of the INSPIRE European directive) utilize this dual feature-coverage view and would benefit from application of this harmonized model within their application schemas.

ACKNOWLEDGEMENTS

The authors wish to thank Paolo Mazzetti and Lorenzo Bigagli (from CNR-IMAA, Italy) for their kind collaboration.

REFERENCES

Barton, C., Fox, P., Khalsa, S.J.S., Jackson, I., Messerotti, M., Wyborn, L., Nativi, S., (ed.s), *Report of the ESI Summit 13-14 March 2008*, Rome Italy, June 2008.

Caron, J., "Common Data Model Scientific Feature Types", *UNIDATA/UCAR presentation*, available at: http://conan.unidata.ucar.edu/Presentations/UPCsemseries/caronfeaturetalk.pdf, 2008.

Couclelis, H., People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In AU Frank, I Campari, and U Formentini, editors, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, volume 639 of Lecture Notes in Computer Science, pages 65–77. Springer-Verlag, 1992.

Goodchild, M.F., 1992, Geographical data modeling, *Computers & Geosciences*, 18(4):401–408.

ISO/CD 19156, Geographic information — Observations and measurements, *ISO/TC 211 document* N 2616, 2008.

ISO/FDIS 19123, Geographic information—Schema for coverage, geometry and functions, ISO/FDIS 19123:2005(E).

ISO/IS 19107, Geographic information—Spatial schema, ISO/IS 19107, 2003(E).

Juceviciene, V., Towards a Single Information Space for the Environment in Europe, *INSPIRE 2008*, Maribor, Slovenia, 23-25 June 2008.

Ledoux, H., FieldGML: An Alternative Representation for Fields, *annex of ISO/CD 19156, Geographic information — Observations and measurements*, ISO/TC 211 N 2616, 2008.

Peuquet, D.J. A Conceptual Framework and Comparison of Spatial Data Models. *Cartographica*, 21(4), pp. 66-113, 1984.