

DEVELOPING A MODELING FOR THE SPATIAL DATA INFRASTRUCTURE

Jan Hjelmager, Tatiana Delgado, Harold Moellering, Antony Cooper, David Danko, Michel Huet,
Henri Aalders, Alexander Martynenko

Kort & Matrikelstyrelsen, Rentemestervej 8, Copenhagen, DK-2400 NV, Denmark

Grupo Empresarial GeoCUBA, Calle 6 No. 301 Esq. 3ra, Miramar, La Habana 11300, Cuba

Department of Geography, Ohio State University 43210 USA

Centre for Logistics and Decision Support, CSIR, PO Box 395, Pretoria, 0001, South Africa

Environmental Systems Research Institute, Inc, 8620 Westwood Center Drive, Vienna, VA 22182-2214, USA

International Hydrographic Bureau, 4 quai Antoine 1er, Monaco

Dept. Burgerlijke bouwkunde, Kasteelpark Arenberg 40, B-3001 Heverlee, Belgium

Institute of Informatics Problems, Russian Academy of Sciences, 44, Vavilova st. Moscow, 117333, Russia

ABSTRACT

Theme: Spatial Data Infrastructure, Spatial Data, Analytical Cartography, Spatial Theory

The Commission on Spatial Data Standards of the International Cartographic Association (ICA) is working on defining spatial models and technical characteristics of a Spatial Data Infrastructure (SDI). To date, this work has been restricted to the Enterprise and Information Viewpoints from the ISO Open Distributed Processing standard. The Commission has developed models for these two viewpoints. The models describe how these two different parts of an SDI fit together in the viewpoints in question. The models cannot be seen as a final result, but more as a small step towards a model that defines the previously mentioned overall model of the SDI and its technical characteristics.

During the model development process, the role of the different actors in an SDI have also been identified in Use Case diagrams of an SDI. The model is developed using the Unified Modelling Language (UML).

1. INTRODUCTION

In order to get an overview of the area of modelling a spatial data infrastructure (SDI), one of our first tasks was to review the different reference models applicable to the SDI. The architecture reference model used by the International Standards Organization (ISO) Technical Committee for Geographic Information/Geomatics, ISO/TC 211 (ISO 19101, 2002), the OpenGIS Reference Model (ORM) (OGC, 2003) and the Geospatial Interoperability Reference Model (GIRM), were the main reference models reviewed and discussed. The base used in the majority of these reference models is the Reference Model of Open Distributed Processing (RM-ODP) (ISO/IEC 10746, 1995), which defines a framework comprising five viewpoints: Enterprise, Information, Computation, Engineering and Technology. RM-ODP allows describing complex distributed systems giving a framework of different levels of abstraction (Delgado 2004).

This work is based on the SDI as a type of complex version of a Virtual Map 3, as defined by Moellering (1980, 1984). It should be clear to the reader that these sorts of spatial databases are usually easily transformable into the other forms of maps. Real Maps (Hard Copy), Virtual Map 1 (Screen Display), and sometimes into a Virtual Map 2 (example: CD-ROM or DVD). It should be realized that as spatial data, two of these classes of maps are visualizations, as defined by Nyerges (1980) as Surface Structure, while the non-visual classes 2 and 3 are defined by Nyerges (1991) as Deep Structure. For an overall review of these spatial concepts and many others from Analytical Cartography, please see Moellering (2000).

The motivation to undertake this work is to obtain a multi-perspective description of the Spatial Data Infrastructure (SDI) as part of the terms of reference for 2003-2007 for the Spatial Data Standards Commission of the International Cartographic Association (ICA). These are enunciated as follows:

- To develop a conceptual model of the Spatial Data Infrastructure (SDI) using the Unified Modeling Language (UML) and associated modeling concepts, working in the areas of science, technology and standards, at the global, regional and national levels.
- To define the technical characteristics of the SDI, and concepts for appropriate data sets for the SDI.

The work described here is aimed to approach the above objectives. Our work began by developing a preliminary high-level model of an SDI, identifying some of the actors, uses cases and classes, described using UML (Cooper et al 2003). Subsequently, the ICA Commission on Spatial Data Standards has exploited the first two RM-ODP perspectives to describe an SDI: Enterprise and Information viewpoints. Future work will be to define the computation perspective contained in the SDI Reference Model according to the action field defined by our ICA Commission.

Different notations may be chosen as appropriate to reflect the requirements of the viewpoint into the RM-ODP. These notations may be natural, formal, textual or graphical. In this work that follows, UML was used as the main notation language to express the two viewpoints.

2. THE PROCESS OF DEVELOPING THE MODELS

The process of developing the models followed a simple and well-proven path as planned by the ICA Standards Commission. When developing the conceptual models for describing the two viewpoints on an SDI, we have been using the following approach:

- Scientific discussion to define the concepts, and describe how they might fit together;
- Organise the concepts;
- Develop a graphic expression of the resulting SDI model as currently envisioned.

The heterogeneous composition of the membership of the ICA Commission on Spatial Data Standards encouraged contributions from different perspectives to the ICA SDI model.

During this discussion of the entire SDI modelling process, whenever there has been a need for it, different terms and definitions have been discussed in order to have the same basis and understanding throughout the discussion. Ideally, we have tried to keep the same use of terms and definitions as used in the ISO/IEC 10746 and ISO/TC211 standards.

3. INTRODUCTION TO FIVE SDI MODELING VIEWPOINTS

The architectural reference model provided by RM-ODP consists of five different viewpoints:

- Enterprise View;
- Information View;
- Computation View;
- Engineering View;
- Technology View.

The Enterprise Viewpoint, (the first) describes the purpose, scope and policies for a Spatial Data Infrastructure (SDI). The Information Viewpoint, the second view, describes the semantics of information and information processing incorporated into an SDI. The Computational Viewpoint, the third view, is a functional decomposition of the SDI into objects and services that interact at interfaces. It is in this viewpoint that one will find the service-oriented architecture (SOA). The Engineering Viewpoint, the fourth, contains the mechanisms and functions required to support distributed interaction between the objects within an SDI. The Technology Viewpoint, the fifth and last viewpoint, contains the specific technology(ies) chosen for the implementation of an SDI. However, it is only the first two viewpoints that we will take into consideration in this paper: i.e. the Enterprise and the Information viewpoints. These are essentially components from the Real World and Information Structure, as the top two of the six defined Nyerges (1980) Data Levels.

Why have the SDI in the first place?
 Verbal description:
 To facilitate provision, management and use of geospatial data and services.
 Base collection of technologies, policies and institutional arrangements to facilitate availability of, and access to, spatial data. Data discovery, evaluation, and application. [GSDI Cookbook v1.1, 2001].

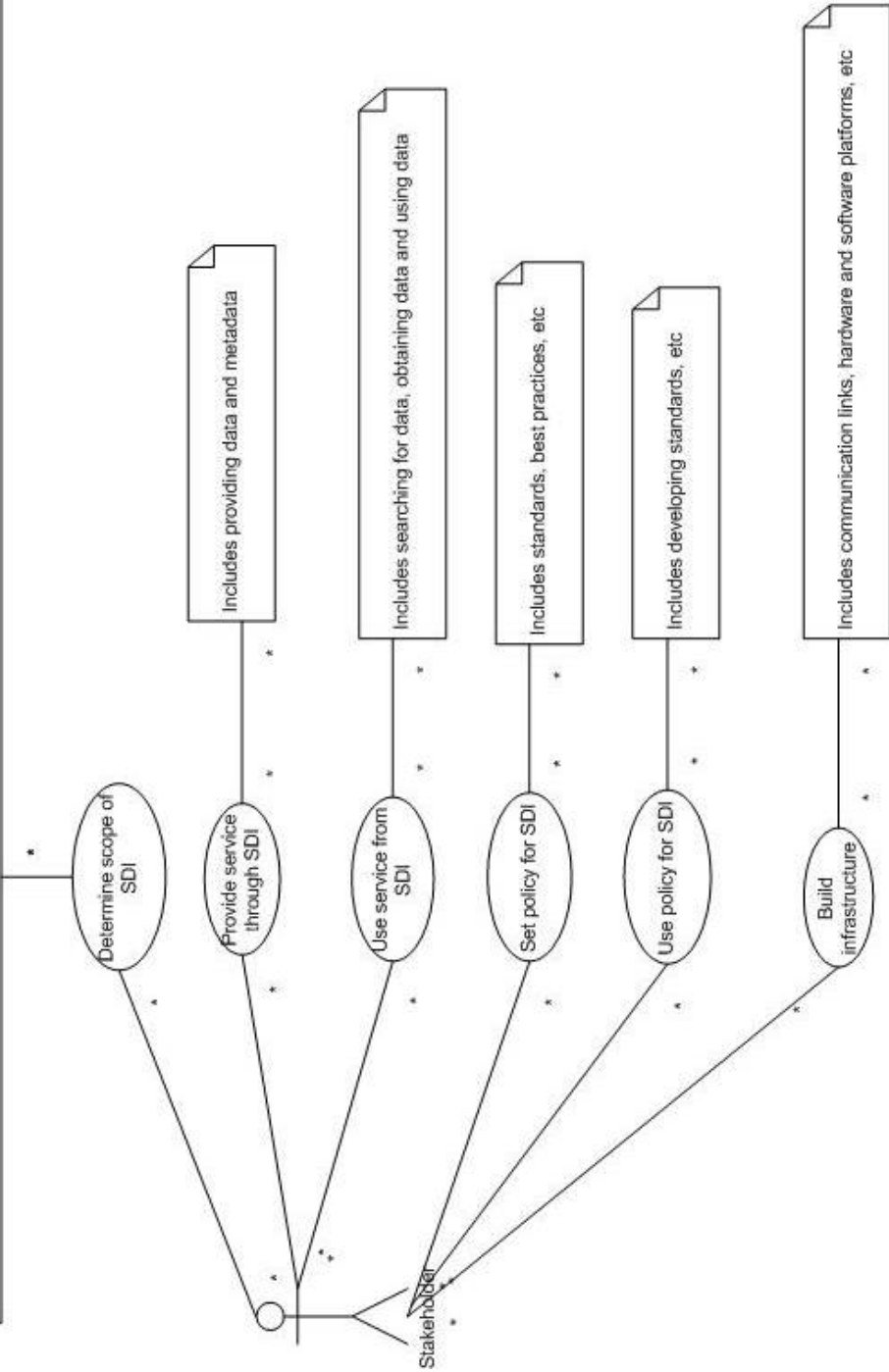


Figure 1. Use Case Diagram for the Enterprise Viewpoint

4. THE ENTERPRISE VIEWPOINT

As mentioned above, this viewpoint consists of four different elements, i.e. the purpose, scope and policies for an SDI. However, these are only head topics, and it takes more than the headlines to build an SDI. As can be seen in the use case diagram (Figure 1), each stakeholder within an SDI can be part of different use cases.

For example, here the same stakeholder could determine the scope of an SDI, use services from an SDI (such as searching for, obtaining, and using data), and/or build the infrastructure used by the SDI (whether it be the networks, computers, software or whatever else). Each one of these interactions then comprises a separate use case.

As can be seen in Figure 1 above, the scope and policies of an SDI can be separated into the stakeholder that uses either the service or sets the service. The same can be said in a way for the policy. The reason for this division of labor is that the groups responsible for developing and for maintaining the two parts of the use case have totally different interests, and points of view from each other, even though on a high level their general interest must be said to be mutual.

In Figure 1 the overall actor has been generalised into a stakeholder. However, this actor can be divided into five different actors (see Figure 2) all having a role to play in one or other of the use cases in Figure 1. In the diagram, each of these five actors *extends* the stakeholder actor.

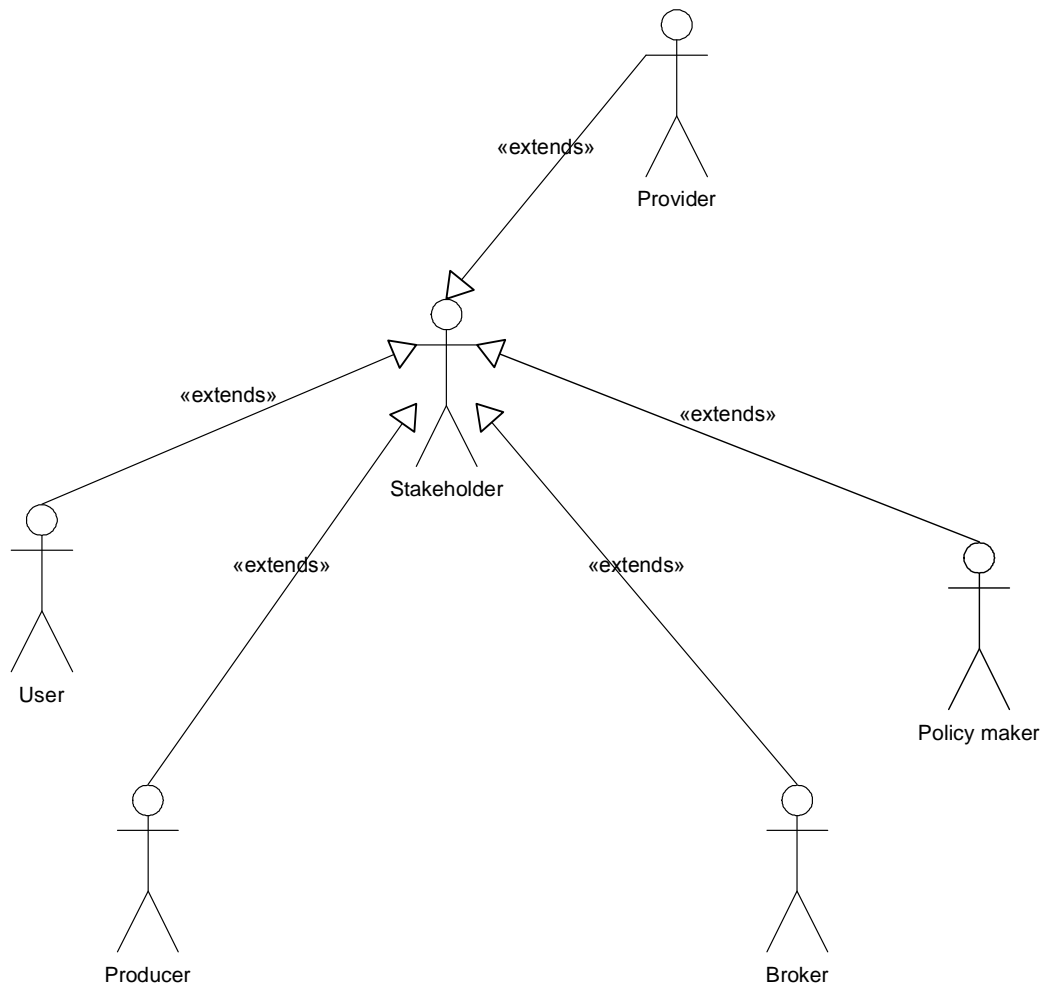


Figure 2. The Actors in the Enterprise Viewpoint

In order to see now the different parts of the use cases fit together, we have developed an initial view object model, as shown in Figure 3. As can be seen in the Figure, the policy part only consists of a single object class. However, this class can be divided into several other classes through inheritance, as shown in Figure 4. The class “Policies” in Figure 4 will be treated as an abstract UML class because this class can never be instantiated, whereas the subclasses can be instantiated with the attributes from this class, along with attributes from their own class.

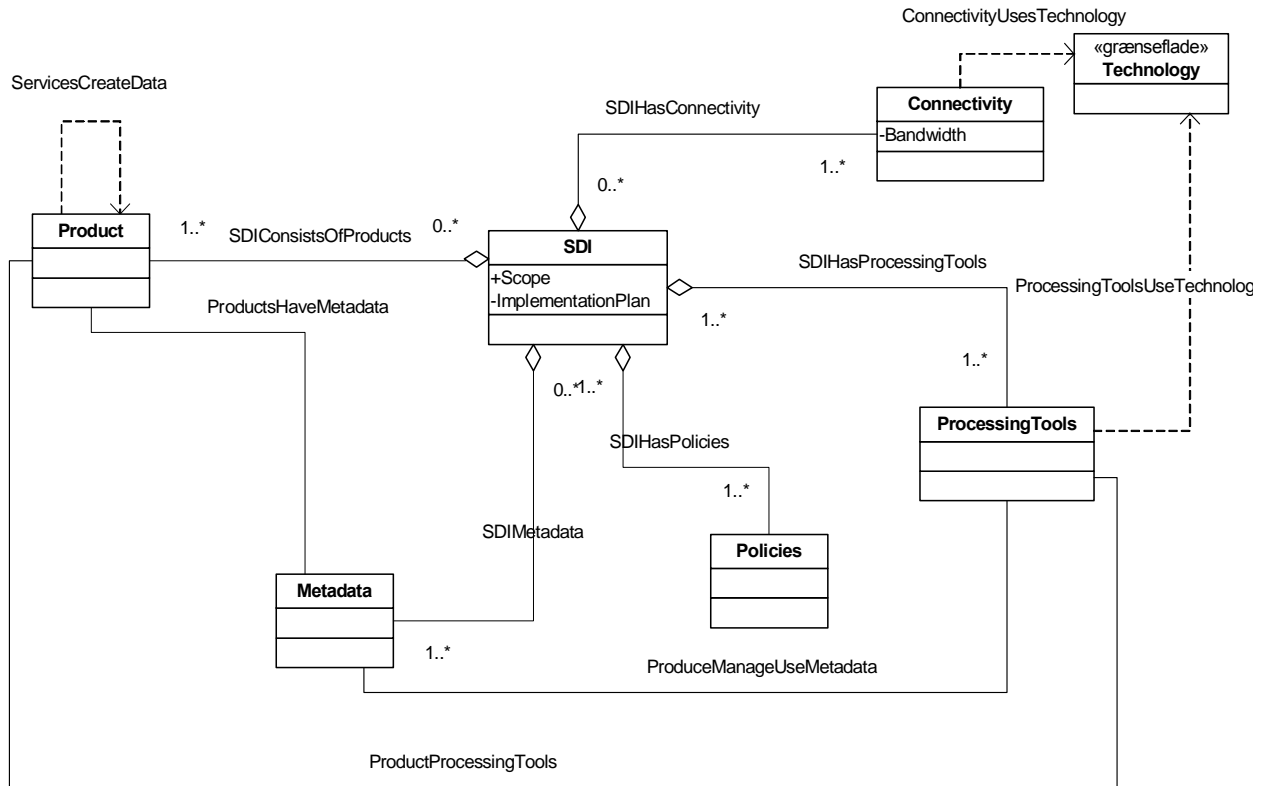


Figure 3. Object Diagram of the Enterprise Viewpoint

The classes Constraints, Standards and Best Practices shown in Figure 4 might have some association between each other, because Standards might impose some Constraints on the Policies and vice versa. The mentioned associations between the classes are not included in the figure due to uncertainty regarding the definition of the classes. When it comes to an association between the classes Standards and Best Practices, it should be recognized that the implementation of standards might end up with some implementation specification (e.g. ISO/PDTS 19139, 2005).

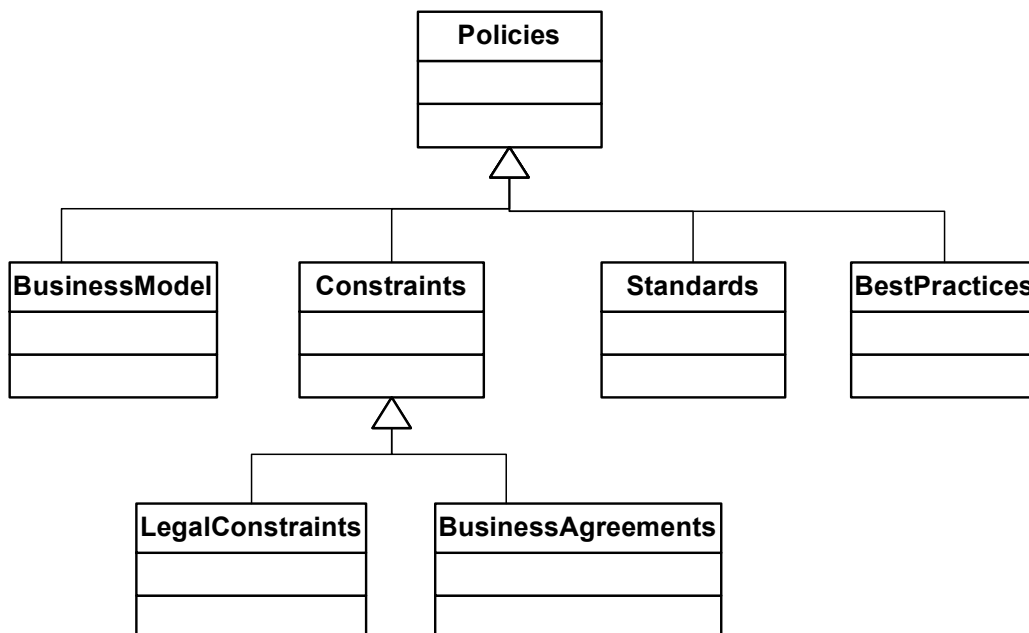


Figure 4. The Policy Class in the Enterprise Viewpoint

5. THE INFORMATION VIEWPOINT

Where the Enterprise Viewpoint had its main focus on the administrative setup for an SDI, the Information Viewpoint deals with the data and the semantics of the data (ISO 19101, 2002). As it can be seen on Figure 5, the center on which everything turns, is the product. The product will in this context be defined as services and/or data that forms the SDI.

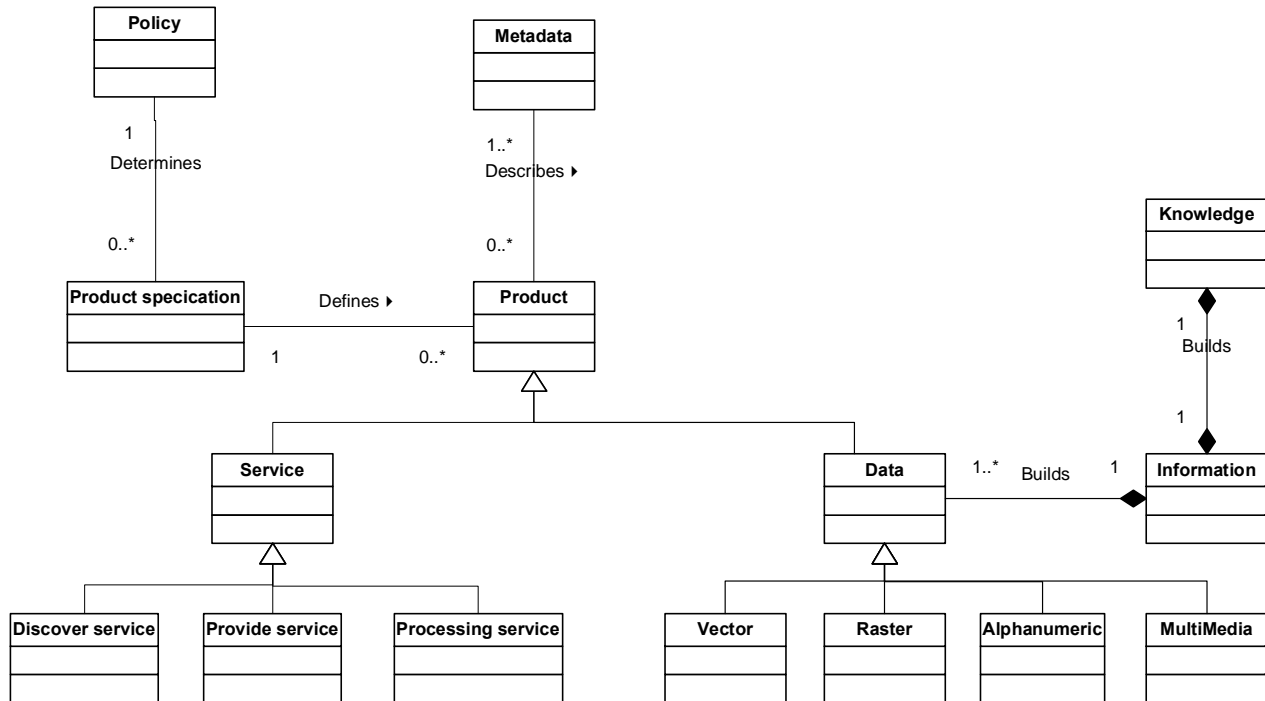


Figure 5. The SDI Information Viewpoint

If one focuses on some of the classes in Figure 5, e.g. Product Specification, Metadata, Policies and Product, and then further subdivides these classes into sub-elements, and then connects these elements with the stakeholders, as in Fig. 2, then it is possible to define the stakeholders' roles in connection with the classes and their sub-elements. As can be seen in Table 1, the stakeholder can have one of two different roles (active or passive) in relationship to a class. The *Active* stakeholder initiates or executes the class (for example), while the *Passive* stakeholder is the beneficiary of the class. The implication of this finding reveals that the classes must either be clearer in the definition, or alternatively, divided into classes according to the sub-elements.

Table 1 depicts the desired SDI levels of abstraction given by the ICA Commission on Spatial Data Standards. It means that not all the use cases are disaggregated until the maximum level regarding the purposes defined by the Commission. As it can be seen in the Table below, some rows have just one actor categorized as *Active* and only one as *Passive*; all these cases represent the maximum disaggregating level for the use case in question (eg. "Harvest Metadata"). Where there is more than one actor as *Active* or *Passive*, we can assume a generalized Use Case, that means this Use Case could be disaggregated into new ones, but for a better understanding of the whole phenomena, it has been modelled at this abstraction level. The most representative examples of this are the Use Cases associated to "Policies".

Analyzing the columns in Table 1, one can derive the roles for each actor. For instance, "Providers" in an SDI are in charge of the following activities:

- Obtaining and implementation of product specifications from users and policy makers
- Providing geospatial products (geospatial data and services)
- Managing geospatial products
- Assuring the quality of products (together with other actors)
- Publishing and providing metadata
- Assuring the quality of metadata (together with other actors)
- Applying policies established by policy makers.

Entities	Use cases elements	User	Producer	Provider	Broker	Policy maker
Product spec's	Stipulate requirements	A			P	
	Translate into product specifications	P			A	
	Obtain product spec's		A	A	A	P
	Consult users	P				A
	Collect product spec's	P			A	
	Implement product spec's	P	A	A	A	P
Metadata (incl. Service capability)	Harvest metadata			P	A	
	Expose metadata	P			A	
	Publish metadata			A	P	
	Produce metadata		A	P	P	
	Provide metadata	P	A	A	P	
	Maintain metadata		A		P	
	Analyse metadata	A			P	
	Quality assure metadata	A	A	A	A	P
	Create metadata indices	P			A	P
Policies	Make policy	P	P	P	P	A
	Apply policy	A	A	A	A	P
	Business plan	A/P	A/P	A/P	A/P	A/P
Product (incl. data + services)	Produce product		A	P	P	
	Provide product	P		A	P	
	Provide query service	P			A	
	Obtain product	A			P	
	Data capture/creation (from source)		A	A		
	Manage product	P	A	A		
	Assure quality	A	A	P	P	
	Assure quality (accreditation of organisation)	P				A
	Assure quality (creation of product)	P	A	A		P

Table 1. Use Cases and Actors in an SDI. P: Passive, A: Active.

6. SUMMARY AND CONCLUSION

This work is a result of the activities undertaken by the ICA Commission on Spatial Data Standards during the last two years, according to its Terms of Reference for the period 2003-2007.

The use of the Open Distributed Processing Reference Model (RM-ODP) combined with the Unified Modelling Language (UML), resulted in a positive way to obtain a comprehensive model to describe a Spatial Data Infrastructure (SDI) in terms of scope, activities, actors (Enterprise Viewpoint) and the semantics of information and information processing (Information Viewpoint). In due course, the model of an SDI based on RM-ODP, as presented in this paper, should be refined and completed by the addition of the computation perspective, which comprises the services in an SDI, consistent with the concepts exposed in the Enterprise and Information viewpoints, to obtain a more comprehensive model to describe Spatial Data Infrastructures.

Future work will be also necessary to validate the model in specific user communities and at different levels of SDI (National, Regional and Global).

7. ACKNOWLEDGEMENTS

This is to recognize the contribution of Kevin Hawley in editing and polishing this manuscript for the ICA Spatial Data Standards Commission. Access and office facilities have been provided by the Department of Geography, Ohio State University, Columbus, Ohio, USA.

8. LITERATURE REFERENCES

- Cooper AK, Hjelmager J, Nielsen A and Rapant P (2003), A Description of Spatial Data Infrastructures (SDIs) Using the Unified Modelling Language (UML), 21st International Cartographic Conference, Durban, South Africa.
- Delgado T (2004), Analysis of Reference Models as Starting Point to Model the SDI in the ICA Spatial Data Standards Commission, prepared for the 2004 ICA Spatial Data Standards Commission meeting, Monaco.
- ICA Spatial Data Standards Commission (2003), Meetings held in Ithala, S. Africa, August, 2003.
- ICA Spatial Data Standards Commission (2004), Meetings held at IHB Headquarters, Monaco, July, 2004.
- ISO/IEC 10746 (1995), Open Distributed Processing - Reference model.
- ISO 19101 (2002), Geographic Information – Reference model.
- ISO/PDTS 19139 (2005), Geographic Information – Metadata – XML schema implementation.
- Moellering, H. (1980), "Strategies of Real-Time Cartography", Cartographic Journal, 17(1), June, pp. 12-15.
- Moellering, H., (1984), "Real Maps, Virtual Maps and Interactive Cartography", Chapter in Spatial Statistics and Models, G. L. Gaile and C. J. Wilmott, editors, Dordrecht: Reidel Publishing Co., pp. 109-131.
- Moellering, H., (2000), "The Scope and Content of Analytical Cartography", in Moellering (ed.), The Nature of Analytical Cartography, Special Issue on Analytical Cartography, Cartography and Geographic Information Science, 27(3), pp. 205 - 223.
- Nyerges, T., (1980), Modelling the Structure of Cartographic Information for Query Processing, Unpublished Ph.D. dissertation, Ohio State University, Columbus, Ohio. 203 pp.
- Nyerges, T., (1991), "Geographic Information Abstractions: Conceptual Clarity for Geographic Modeling", Environment and Planning A, 32, pp. 1483-1499.12