

Marine ecology service reuse through taxonomy-oriented SPL development



Agustina Buccella^{a,b,*}, Alejandra Cechich^a, Matias Pol'la^{a,b}, Maximiliano Arias^{a,b},
Maria del Socorro Doldan^c, Enrique Morsan^c

^a GIISCO Research Group, Departamento de Ingeniería de Sistemas – Facultad de Informática, Universidad Nacional del Comahue, Buenos Aires 1400, Neuquen 8300, Argentina

^b Consejo Nacional de Investigaciones Científicas y Técnicas – CONICET, Argentina

^c Instituto de Biología Marina y Pesquera “Almirante Storni”, Universidad Nacional del Comahue – Ministerio de Producción de Río Negro, San Antonio Oeste, Argentina

ARTICLE INFO

Article history:

Received 8 July 2013

Received in revised form

10 September 2014

Accepted 12 September 2014

Available online 20 September 2014

Keywords:

Domain-specific taxonomies

Software reuse

Domain engineering

Geographic information systems

ISO 19100 standards

ABSTRACT

Nowadays, reusing software applications encourages researchers and industrials to collaborate in order to increase software quality and to reduce software development costs. However, effective reuse is not easy and only a limited portion of reusable models actually offers effective evidence regarding their appropriateness, usability and/or effectiveness. Focusing reuse on a particular domain, such as marine ecology, allows us to narrow the scope; and along with a systematic approach such as software product line development, helps us to potentially improving reuse. From our experiences developing a subdomain-oriented software product line (SPL for the marine ecology subdomain), in this paper we describe semantic resources created for assisting this development and thus promoting systematic software reuse. The main contributions of our work are focused on the definition of a standard conceptual model for marine ecology applications together with a set of services and guides which assist the process of product derivation. The services are structured in a service taxonomy (as a specialization of the ISO 19119 std) in which we create a new set of categories and services built over a conceptual model for marine ecology applications. We also define and exemplify a set of guides for composing the services of the taxonomy in order to fulfill different functionalities of particular systems in the subdomain.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Over the past several years, the software engineering field has been aimed at improving processes for generating software products. Thus, new approaches have emerged evoking similar goals and trying to generate the same benefits; all of them seek to improve the development time and costs of the stages of the software development life cycle, and ensure at the same time, a suitable time-to-market without affecting the quality of the final product. A remarkable particularity of all these new approaches is that they have increasingly focused efforts on taking advantage of

reusing techniques at different levels to achieve the goals. Software product line engineering (SPL) (Clements and Northrop, 2001; Pohl et al., 2005; van der Linden et al., 2007), component-based software development (CBSD) (Szyperski, 1998), and service-oriented software engineering (SOSE) (Papazoglou et al., 2007) are some of the most important areas in which reusing is the key for generating better software products. In general terms, the approaches are focused on the development of software systems by combining reusable components and/or services perhaps developed at different times, by different people, and possibly with different uses in mind. At the same time, the reuse can be applied on different stages of the software development; we can reuse requirements, designs, objects, functions, or complete components or frameworks as black or white boxes. However, as in other industry fields, reuse needs a systematic methodology in order to ensure its effectiveness. Systematic software reuse (Frakes and Isoda, 1994) comprises a set of techniques and tools to guide developers on developing reusable software artifacts carefully planned and managed.

* Corresponding author at: GIISCO Research Group, Departamento de Ingeniería de Sistemas – Facultad de Informática, Universidad Nacional del Comahue, Buenos Aires 1400, Neuquen 8300, Argentina. Tel.: +54 299 4490300x435; fax: +54 299 4490313.

E-mail addresses: agustina.buccella@fi.uncoma.edu.ar (A. Buccella), alejandra.cechich@fi.uncoma.edu.ar (A. Cechich), matias.polla@fi.uncoma.edu.ar (M. Pol'la), ariasmaxi89@gmail.com (M. Arias), msdoldan@gmail.com (M. del Socorro Doldan), qmorsan@gmail.com (E. Morsan).

In particular, in SPLE, systematic reuse is defined by a set of design and implementation techniques which support the development and management of reusable software artifacts. Depending on the SPL methodology followed,¹ these techniques apply different mechanisms according to the two phases involved here – *domain engineering* and *application engineering*. In the first phase, the domain analysis is aimed at identifying, capturing, and organizing all source information gathered from existing systems in the domain, domain experts, textbooks, prototyping, experiments, already known requirements on future systems, etc. (Czarnecki, 2002). As a result, a set of reusable and configurable components, implementing commonalities and variabilities of a particular problem domain, is defined as part of a reference architecture. In the second phase, *application engineering*, this reference architecture must be instantiated, by binding the variability of the reusable components, in order to generate the application's architecture specific of an organization. Finally, the last activity of this phase returns a particular software product. As we can observe from these two phases, the success of an SPL development depends on the identification, use, and management of the reusable artifacts. Thus, the application of systematic reuse is crucial here. However, achieving software reuse is not an easy task and special efforts must be invested. Main technical and non-technical challenges of systematic reuse include (Schäfer et al., 1993):

- *Non-technical*: The extra-effort invested on achieving software reuse, with respect to economical, organizational, and managerial issues should be then retrieved when new products are generated.
- *Technical*: The selection and application of an SPL methodology should guarantee an effective reuse by simplifying the process of product generation.

In SPLE, these two challenges have been analyzed and studied in several proposals in the literature (Pohl et al., 2005; Matinlassi, 2004; Bosch, 2000; van der Linden et al., 2007), without concluding on any standard way to manage them. However, several efforts in the academy and industry have been addressed on defining standard information and activities (of any type) to support the process of designing and implementing the reusable components. Some of these efforts are focused on creating domain taxonomies and conceptual models to be shared during component development. These taxonomies and models try to improve the communication and provide a common vocabulary among participants (Sujatha et al., 2011). At the same time, the taxonomies and models, as they will be used in an SPL development, are specialized according to a specific domain. Among them, the geographic domain presents different particularities that make it attractive to analyze. Geographic information systems (GIS) are considered as members of an area emerging from general-purpose information systems but taking aspects from other areas such as cartography and topology. In addition, the geographic domain includes a group of more specific domains or branches, each of them focused on its own particularities.

We can find a first classification in which the geographic area is divided into three main branches² (Bonnett, 2008): *human geography*, focused on the study of patterns and processes of the human society; *physical geography*, focused on the productions and interactions of organisms, climate, soil, water, and landforms, over the nature environment; and *environmental geography*,

combining the physical and human phenomenon to analyze interactions between the environment and humans. Also, within the physical geography we can find other areas or domains including oceanography and climatology; and at the same time, the oceanography domain includes other subdomains such as marine geology, marine ecology and marine fishery.

Terms like “classification”, “ontology” and “taxonomy” are used abundantly when modeling GIS; but, as mentioned by Rees (2003), the distinction between these terms is often blurred. How to develop a taxonomy in a global way is described by many authors (Bruno and Richmond, 2003; Choksy, 2006; Cisco and Jackson, 2005; Sujatha et al., 2011); however, these taxonomy development methods stop at the organizational level, and are not formally described, still not allowing engineering a complete domain-specific taxonomy. Therefore, developing domain-specific taxonomies that help support domain modeling remains a challenge.

Looking at organizations for standardization, such as the International Organization for Standardization³ (ISO), and more specifically, the Open Geospatial Consortium⁴ (OGC), we can find rules for guiding the process of representing any geographic domain. Examples of that are the ISO 19109 std (*Rules for Application Schema*)⁵ for the construction of application schemas (based on spatial and temporal conceptual models), the ISO 19107 std (*Spatial Schema*)⁶ for the definition of spatial representations, or the ISO 19119 std (*Services*)⁷ for the definition of a geographic service taxonomy. These abstract and generic standards have been created to be applied to any subdomain included in the geographic domain. Therefore, they are useful as a starting point for helping us define a domain-specific taxonomy.

In a previous work (Buccella et al., 2013), we have described our experiences in an SPL development in the marine ecology subdomain focusing on the improvements (of time and cost) in the product derivation phase. In the paper presented here, we fully describe the semantic resources created for assisting each of the activities in this SPL development and thus promoting the systematic software reuse in this subdomain. Therefore, the main contributions of this work are focused on the definition of a standard conceptual model in this subdomain together with a set of services and guides which assist the process of SPL development. The services are structured in a service taxonomy (as a specialization of the ISO 19119 std) in which we create a new set of categories and services built over the conceptual model for the marine ecology subdomain. Then, we define a set of guides for composing the services of the taxonomy in order to fulfill different functionalities of particular systems in the subdomain. Finally, we show how these two resources are useful in the domain and application phases of an SPL development. In addition, we also show how they give stakeholders a common vocabulary, and more importantly, a common service structure as a basis for trade-offing.

This paper is organized as follows. The next section presents related work in the literature taking into account proposals for defining service taxonomies in reusable approaches, SPLE and CBSD, and for GIS domains particularly. Section 3 provides background and describes our previous experiences on SPL development on the marine ecology subdomain. Section 4 describes our service taxonomy for this subdomain and generic guides for

³ <http://www.iso.org>

⁴ <http://www.opengeospatial.org/>

⁵ Geographic Information. Rules for Application Schema. Draft International Standard 19109, ISO/IEC 2005.

⁶ Geographic Information. Spatial Schema International standard 19107, ISO/IEC 2003.

⁷ Geographic information. Services International Standard 19119, ISO/IEC, 2005.

¹ In the literature there exist several different methodologies to follow an SPL development (Pohl et al., 2005; Matinlassi, 2004; Bosch, 2000).

² <http://en.wikipedia.org/wiki/Geography>

representing particular functionalities of possible systems. Section 5 analyzes usability and functionality of the taxonomy and guides by considering their application for building a software product line, and two products derived. Future work and conclusions are discussed afterwards.

2. Related work

In the software modeling literature there exist different proposals about taxonomies but pursuing different goals. For example, looking at the specific areas we are interested in, such as SPLE or CBSD, the proposed taxonomies are defined as conceptual frameworks; i.e they define a set of categories useful for classifying different methodologies in some aspects, such as variability techniques (Svahnberg et al., 2005), non-functional properties (Noorian et al., 2012) and evolution concerns (Svahnberg and Bosch, 1999). As far as we know, there is no previous work about domain-oriented service taxonomies used to assist the process of developing SPLs or components. However, there are several works defining different types of services that could be used in any software development process. For instance the *Arizona Dictionary and Taxonomy of Human Services*⁸ is being developed by members of the Arizona Taxonomy Committee (composed of representatives of departments, cities and towns of the Arizona State) for defining a common language by means of a set of specific services.

Other set of works proposes methodologies for creating taxonomies that will be useful for information system development. For instance, the works presented by Hunink et al. (2010), Bruno and Richmond (2003), Choksy (2006) and Nickerson et al. (2009) propose methodologies to create taxonomies that assist different activities within and across organizations. In addition, these methodologies are applied in specific domains in order to show the viability of the proposals. For example, in Hunink et al. (2010) the methodology is applied to support the development of a platform for an European software ecosystem.

With respect to service taxonomies, in the geographic information area there are some interesting works that should be considered. Some of them, presented in ESPRIT/ESSI (1998), Albrecht (1996), Sklar and Constanza (1991) and Bai et al. (2009), define approaches for categorizing basic operations or features that can be used by any GIS application. For example in Sklar and Constanza (1991), authors describe a wide set of geographic activities focused on the environmental domain such as inventory, assessment, management, and prediction. These activities are used as references for the work presented in Albrecht (1996). It proposes a set of twenty operations that allow building environmental applications using GIS. These operations are independent of data structures and they are oriented to user's tasks rather than software engineers' tasks. Another work has been presented by the BEST-GIS project (Best Practice in Software Engineering and methodologies for developing GIS applications) in ESPRIT/ESSI (1998), whose main goal is to define a list of key GIS operations based on the experience of selected users and the contribution of key field experts. The list defines nine main categories and several other generic subcategories, which should be used for creating a checklist for selecting and defining user requirements when a new GIS application is being implemented. Another work (Reed et al., 2009), presented on the Web as the Ocean Data Portal (ODP)⁹ as part of the International Oceanographic Data and Information Exchange (IODE)¹⁰ programme, is developing a set of standards to promote the exchange and

dissemination of marine data and services. The project has developed a set of software components to mainly allow different marine organizations to upload geographic data sets (according to specific metadata defined by the project) and to use a set of ODP services for discovery, view, analysis, and download of data sources. Finally, the work presented in Bai et al. (2009) defines a geospatial taxonomy for service discovery and interoperability. It is focused on improving the identification of services by providing more description (by means of layers) of each published service.

The next efforts to define basic operations in GIS domains have been addressed by organizations for standardization such as the International Organization for Standardization¹¹ (ISO) and the Open Geospatial Consortium¹² (OGC). The ISO, by means of the ISO Technical Committee 211¹³ (ISO/TC 211), and the OGC have proposed a series of standards to provide rules and methods for creating interoperable geographic systems. Both organizations published a reference model^{14,15} to understand the complex set of specifications for conceptual models, services, exchange languages and many other components of a GIS architecture.

Among the wide range of defined standards, we are particularly interested in the Service Architecture standard (defined in Open-GIS Service Architecture) and the ISO 19119 std. They define a taxonomy of geographic services based on the Extended Open Systems Environment (EOSE) model for geographic information (ISO 19101 std (*Reference model*)¹⁶). The taxonomy is created as part of the information viewpoint document,¹⁷ in order to achieve semantic interoperability by improving common knowledge of the information processed by the systems. This taxonomy consists of five main categories¹⁸:

- (HI) *Human interaction services* for management of user interfaces, graphics, and multimedia; and for presentation of compound documents.
- (MMS) *Model/information management services* for management of the development, manipulation, and storage of metadata, conceptual schemas, and datasets.
- (WTS) *Workflow/Task services* for support of specific tasks or work-related activities conducted by humans.
- (PS) *Processing services* for large-scale computations involving substantial amount of data. It contains four subcategories based on the General Feature Model (ISO 19109 (*rules for application schema*)¹⁹ std):
 - (PS-S) *Geographic processing services – spatial*
 - (PS-T) *Geographic processing services – thematic*
 - (PS-Te) *Geographic processing services – temporal*
 - (PS-M) *Geographic processing services – metadata*
- (CS) *Communication services* for encoding and transfer of data across communications networks.

For each of these types of services, the ISO 19119 std provides examples that can be specialized according to requirements of specific products. For example, the *geographic viewer* (HI-GV)

¹¹ <http://www.iso.org>

¹² <http://www.opengeospatial.org/>

¹³ <http://www.isotc211.org/>

¹⁴ ISO 19101:2002 - Geographic information - Reference model.

¹⁵ OGC Reference Model - Open Geospatial Consortium Inc. - Reference number: OGC 08-062r7, 2011.

¹⁶ Geographic Information. Reference model International standard 19101, ISO 2002.

¹⁷ Geographic Service Architecture Viewpoints defined in Information technology- Open Distributed Processing -Reference model: Overview - ISO/IEC 10746-1:1998.

¹⁸ The acronyms before each category will be then used as prefixes in our specialized services.

¹⁹ Geographic Information. Rules for application schema. Draft International Standard 19109, ISO/IEC 2005.

⁸ <https://www.azdes.gov/taxonomy/>

⁹ <http://www.oceandataportal.org>

¹⁰ <http://www.iode.org>

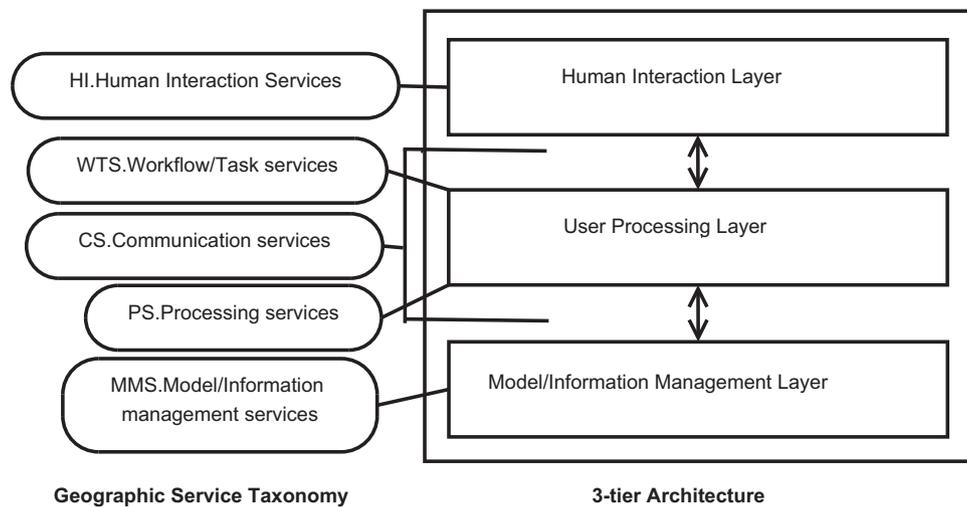


Fig. 1. Geographic services categories defined within the 3-tier architecture.

service is defined as part of the HI category to allow users to interact with map data. Another example is the *route determination service* (PS-S1) defined within the PS category to determine the optimal path between two specified points.

In addition, the ISO 19119 std defines examples of reference architectures using a multi-tier architecture model. One of these examples is a reference model based on a three-tier architecture containing the *human interaction tier*, responsible for the interaction with the user; the *user processing tier*, responsible for the functionality required by the user; and the *model/information management tier*, responsible for physical data storage and data management. The main advantage of this architecture is the required separation of the functionality of the system into three different independent layers that interact only through their well-defined interfaces.

Finally, the ISO 19119 std defines the way the geographic service taxonomy fits in with this three-tier architecture. Fig. 1 shows the reference model proposed in the standard.

Although we can see that there exists a wide set of works in the literature generating standard information to define services and assist the development of new products, there is a lack of focused effort to support the systematic software reuse. In this work, we propose continue working on a standard definition of services within the geographic domain, specializing them for the marine ecology subdomain. The goal of our specialization is to promote an enabling environment of software product line development in this subdomain by generating standard operations and guides towards an assisted development and thus promote the effective software reuse.

3. Background

In this section we present the contextual and methodological basis that support our work. We firstly describe particularities of the marine ecology subdomain and the conceptual model created as a standard structure of information for defining services in this subdomain. Then we describe the main activities of an SPL development in which our semantic resources are involved.

3.1. Contextual information: The marine ecology subdomain

In general terms, the marine ecology subdomain involves scientific studies of the interdependence of all organisms living in the marine-life habitat, and their interactions with each other and the surrounding environment. In particular, three main

aspects of these organisms are encompassed by the investigations in this subdomain, *distribution and classification of organisms*, *identification of fishing zones together with their impacts*, and *relations among organisms*. In order to study and analyze these aspects, scientists must collect a large amount of data which conforms the useful information of the subdomain.

Our work, along with two expert organizations in this subdomain – the Institute of Marine Ecology and Fishery “Almirante Storni”²⁰ (IBMPAS) and the Patagonian National Center²¹ (CENPAT-CONICET) – has allowed us to define an interesting set of activities and goals, and to abstract them in order to be applied by the whole marine ecology community. Both organizations are responsible for storing and analyzing information about checklists of species (i.e. censuses) in three gulfs of the Argentinean Patagonia (*San Matías, San Jorge and Nuevo Gulfs*). Each census, performed once a year, collects information about the population of benthic species living in this area. This information is then used for spatial processing in order to obtain information about spatial distribution of the species, population variation patterns in different scales, etc. In particular, bivalve species, such as the native oyster *Ostrea puelchana* are studied at the IBMPAS and *echinoderms* species are studied at CENPAT-CONICET.

In this context, we must firstly highlight our first semantic resource (a conceptual model) needed to allow scientists to store their data and share the same data structure to create and manipulate information. This conceptual model was also defined to include information about more-general domains such as the oceanographic and geographic domains and thus allow some data to be reused on those domains. In addition, in order to improve the understandability of data, we defined the model by following the rules of the ISO 19109 and ISO 19107 (*Spatial Schema*)²² standards. Fig. 2 shows part of this conceptual model by using a graphical notation in which we merge the MADS (modeling of application data with spatio-temporal features) (Parent et al., 1999) and the UML notation. It helps us to show specific spatial information represented in the model.

The figure shows geographic features²³ such as censuses, stations, zones, fishing zones and species according to the particularities of the subdomain. For example, we can see in the model

²⁰ Instituto de Biología Marina y Pesquera – <http://ibmpas.org/>.

²¹ Centro Nacional Patagónico – <http://www.cenpat.edu.ar/>.

²² Geographic Information. Spatial Schema International standard 19107, ISO/IEC 2003.

²³ The terms “feature” and “geographic feature” have different meanings. In the ISO 19109 std a “geographic feature” is defined as an abstraction of real world phenomena.

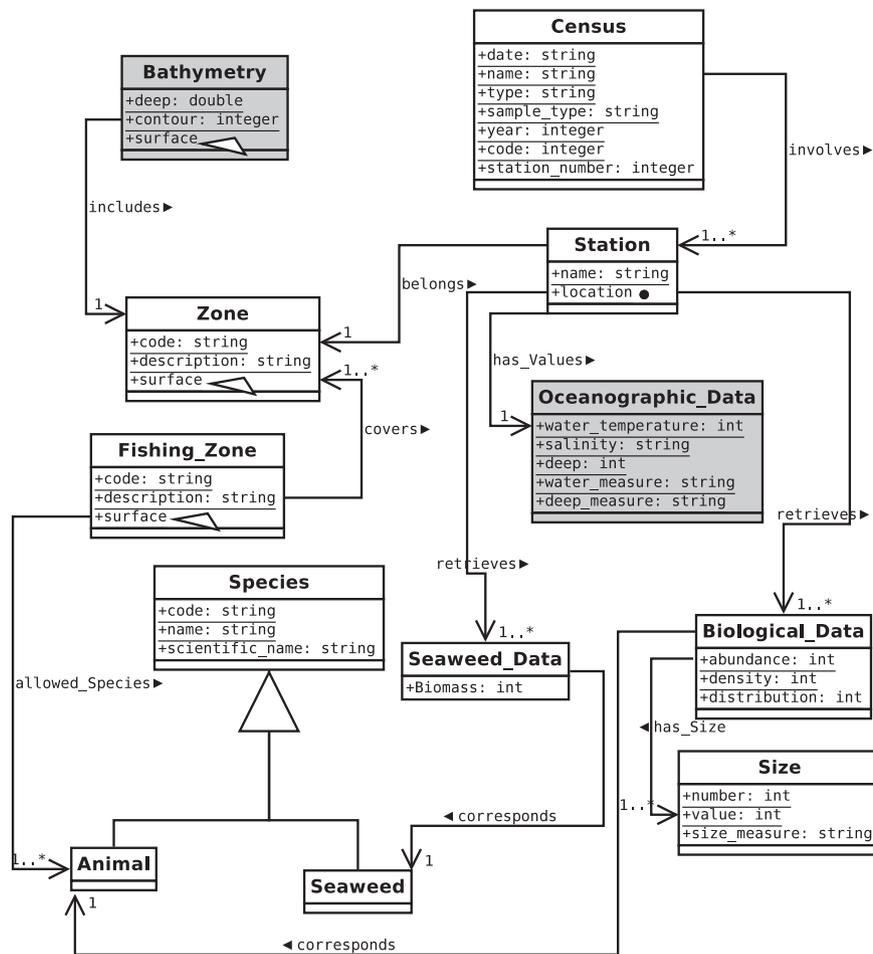


Fig. 2. Part of the conceptual model for the Marine Ecology Subdomain.

that the *Species* class represents a generalization of animals and seaweeds (defined by the specializations *Animal* and *Seaweed* subclasses). These specializations are associated to biological or seaweed data respectively, storing specific information about the context and number of species which have been found (represented by attributes of *Biological_Data* and *Seaweed_Data* classes). All the information is collected from a station (by the *Station* class) which represents a latitude and longitude point in the ocean in which a census was performed. In addition, we defined some classes (in gray in the diagram) that are particular to more general domains such as the *Bathymetry* class representing measurements of ocean depth by depth contours, and the *Oceanographic_Data* class representing specific information about the ocean conditions. These two classes belong to the oceanographic domain.

This conceptual model conforms the basis for defining the set of services used to manipulate information.

3.2. Methodological context: The product line development process

We have defined and applied a domain-level oriented SPL methodology (extended from methodologies proposed in Bosch, 2000; Pohl et al., 2005; Czarnecki et al., 2005; Kang et al., 1990) to develop an SPL in the marine ecology subdomain (Buccella et al., 2013; Pernich et al., 2010, 2012). The SPL was developed to provide a solution for the two organizations working in the same specific subdomain: IBMPAS and CENPAT-CONICET. The applied methodology (Buccella et al., 2013) defined two main analysis as part of the domain engineering phase: *subdomain* and *organizational*. In the

first analysis we performed several domain-related tasks such as determining the information sources to be used (standards, existing applications, domain experts, etc.), analyzing and organizing features or services that the domain should offer, and identifying the set of reusable components that could be used to implement the features defined. In the second analysis (organizational), which is more related to the construction of the software product line than to the domain itself, we defined the range of products to be implemented together with the set of services (with their variabilities) that are part of the line and those that are product specific. Here, instead of describing all the activities performed in these two analysis (fully explained in Buccella et al., 2013), we identify those activities in which the use of our semantic resources (conceptual model, service taxonomy and/or guides) was specially useful. In addition, we indicate how stakeholders were involved. According to the two analysis of the domain engineering phase, these specific activities and tasks are the following:

- **Domain analysis**

- *Information source analysis (ISA)*: This activity analyzes sources of information that can support the domain analysis in order to obtain a first set of requirements. Participants: biologists and software engineers.
- *Subdomain analysis and conceptualization (SAC)*: The information recovered in the previous process is used to analyze and organize the features or services that the subdomain should offer together with the general features derived from the upper domains. Also, in this process the subdomain

must be conceptualized by different software artifacts (such as class models and process models) when it is possible. Participants: biologists and software engineers.

- *Reusable component analysis (RCA)*: This process identifies the set of reusable components that could be used to implement the features defined in the last process. It returns a preliminary reference architecture. Participants: software engineers.
- *Organizational analysis*
 - *Reuse and boundary analysis (RBA)*: This activity defines the organizational boundary, commonality, and variability features. Thus, by considering the features specified in the *subdomain analysis and conceptualization* process and the information from domain experts, the scope of the product line must be defined. Participants: software engineers.
 - *Platform analysis and design (PAD)*: This activity builds the reference architecture based on the features defined in the previous activities and processes. The preliminary structure of reusable components defined in the *reusable component analysis* process is reorganized and refined. Here, each component with their common and variable parts (when necessary) is fully designed. Participants: software engineers and developers.

Then, in the application engineering phase (Pohl et al., 2005), in which we perform the specific activities for developing new products from the line, we must consider the following three main activities in which the service taxonomy and guides might help:

- *Application requirements engineering (ARE)*: This activity must retrieve the specific requirements for a particular organization or application by considering the reusable domain requirements. Participants: biologists and software engineers.
- *Application design (AD)*: By taking into account the reference architecture and the specific requirements of an organization, the activity must define the application's architecture. It selects and configures the reusable components of the reference architecture and adds specific adaptations. In this activity, the variabilities defined for the reusable components must be bounded in order to fix the specific functionality of the resultant product. Participants: software engineers.
- *Application testing (AT)*: This activity must validate and verify an application against its specification. Participants: all members of the project team.

Notice that, in Section 5, this set of activities of each phase will be used to evaluate the impact of the semantic resources generated.

4. A taxonomy to assist an SPL development in the marine ecology subdomain

In this section we firstly introduce the methodology used for building the service taxonomy and then, we focus on the definition of the taxonomy itself and guides.

4.1. Steps for developing a service taxonomy

Firstly, we defined the set of services needed for storing, querying and showing the information of the conceptual model. To do so, we applied a deductive approach (Nickerson et al., 2009) starting from the abstract services of the ISO 19119 std. Obviously, as the universe of geographic services is too extensive, the standard defines general and abstract services. Therefore, in our

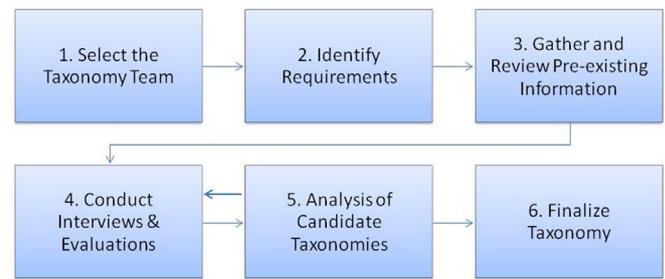


Fig. 3. Steps for taxonomy development.

work, we used them to classify new specific services depending on the particular branch or subdomain in which the system is being implemented, and independently from any particular application.

We considered typical steps proposed for the success of a taxonomy development (Hunink et al., 2010; Choksy, 2006), and we adapted them into six steps (Fig. 3) according to our requirements.

In the first step, *select the taxonomy team*, we defined the taxonomy team involving five expert users (biologists in general) and six software engineers.²⁴ An important thing to highlight here is that these participants reside in different cities (software engineers live in Neuquen and biologists in San Antonio Oeste and Puerto Madryn), so we had to implement some mechanisms to guarantee a fluent communication among them, such as shared files and videoconferencing technologies. In the second step, *identify requirements*, we specified the main requirements and objectives of the taxonomy. We focused on the possible uses considering expert users' and software engineers' interests. The third step, *gather and review pre-existing information*, involved the compilation of information considering two main aspects. On one hand, the ISO 19119 std was deeply reviewed and analyzed according to the five main types of services provided for classifying specific features²⁵ or services in geographic systems. On the other hand, we compiled all the digital and on-paper information related to daily activities of biologists in their respective works. The fourth and fifth steps, *conduct interviews and evaluation* and *analysis of candidate taxonomies*, represented an iterative process in which progressive interviews allowed us to define possible services to be included in candidate taxonomies. Thus, by working together (software engineers and biologists), we firstly defined a set of activities needed to perform the analyses required by each of the main aspects (Section 3.1) described in the subdomain. We focused on defining which of these activities might be supported by computer systems to arrive to better and faster conclusions. Thus, we analyzed the set of defined activities and transform them as a set of standard services of any GIS implemented in the marine ecology domain. In addition, we applied our domain-level approach (Buccella et al., 2013; Pernich et al., 2010) by considering the branches or domains in which the geographic information can be divided. For this work we considered three main domains: *geography*, *oceanography*, and *marine ecology*. Fig. 4 shows these domains in which we can observe that services defined for the marine ecology domain are included in those defined by the oceanography domain.

Specifically, these two steps were managed by the software engineers by guiding biologists on describing their tasks. In particular, we have generated one document for each service category of the standard to focus on specific similar services on

²⁴ Software developers were not part of this team.

²⁵ A feature describes the functional and quality characteristics of a system (Bosch, 2000; Pohl et al., 2005).

each of them. We did not use here a formal way to complete these documents, but we tried to guarantee that all the participants could opine about each new change on these documents during the iterative process. Finally, in the sixth step, *finalize taxonomy*, a review of the taxonomy was performed and final changes were submitted. This step took several rounds of reviews (more than six) in order to improve and clarify taxonomy descriptions and metadata. In particular, the time between each round, that is, once a set of changes were submitted and accepted for beginning the next round, was about no more than two months.

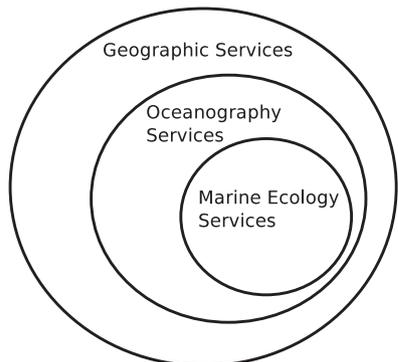


Fig. 4. Services included in more general domains according to the level in which they are defined.

4.2. A service taxonomy for the marine ecology subdomain

Other semantic resource needed to represent the marine ecology subdomain is the set of services which interact to each other to manipulate the stored information in the conceptual model (Section 3). Together with this model, the taxonomy is intended for providing a standard classification of different services the domain can contain and thus provide a framework for allowing reuse. In this way, we built the taxonomy by specializing each category of the ISO 19119 std (see Fig. 1).

In Figs. 5–7 we can see a subset of the services defined for the *geographic human interaction (HI)* category (see Fig. 1).

The services written in italic are part of the ISO 19119 std. The first figure (Fig. 5) shows services of the geographic domain defined as generic services which allow users to manipulate maps and layers in an interactive way. Then, we defined the services HI-MM (*Map manipulation*) and HI-RM (*Raster manipulation*) as specializations of HI-GV (*Geographic viewer*) and HI-CVI (*Geographic viewer – imagery*) respectively. For example, services within the HI-MM category were defined for manipulating general aspects to interact with maps such as zoom in or zoom out, and panning.

Then, in Figs. 6 and 7 we specialize these general services to fulfill the requirements of the oceanography and marine ecology subdomain. In Fig. 6 we can see that for the general HI-LM5.1 service (*Show/hide layers according to specific scales*) we defined six new services to show layers such as zones, stations and fishing, according to the conceptual model (Fig. 2). In addition, as we

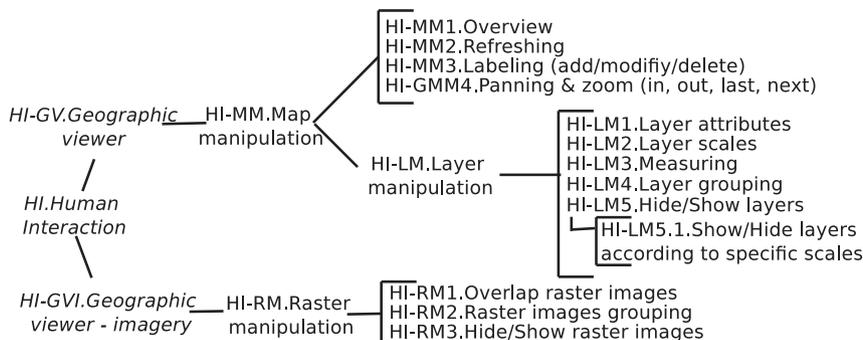


Fig. 5. Services of the geographic domain for the Geographic Human Interaction category.

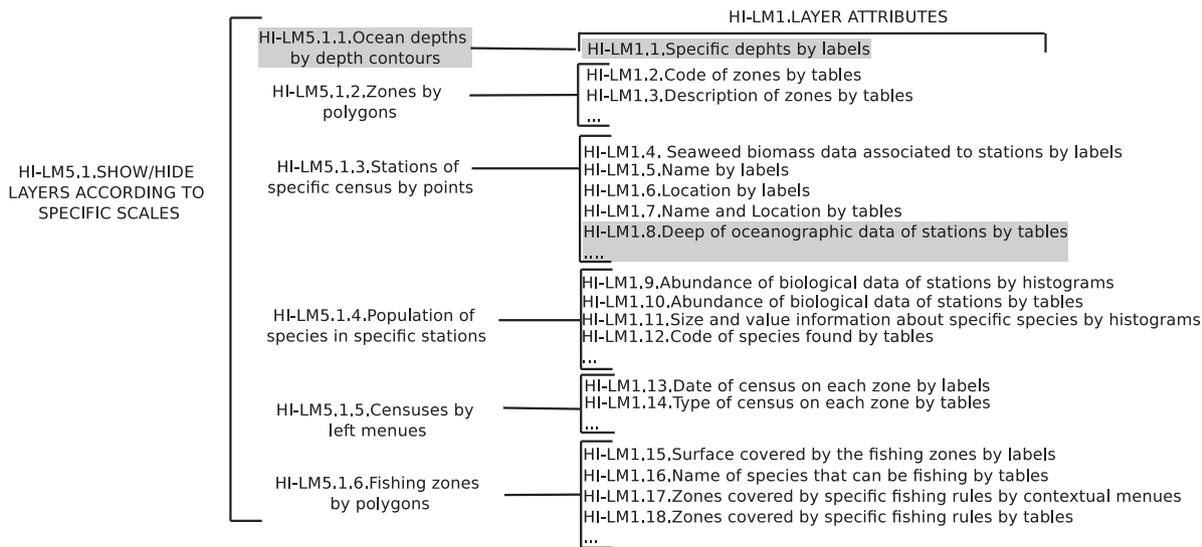


Fig. 6. Services of the marine ecology and oceanography domains for the Geographic Human Interaction category.

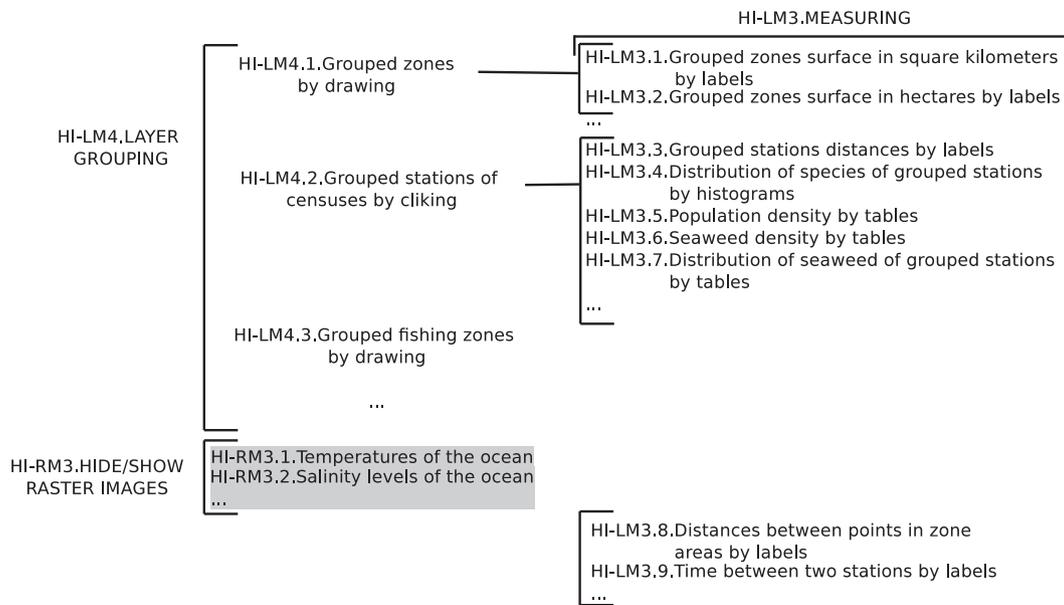


Fig. 7. Services of the marine ecology and oceanography domains for the Geographic Human Interaction category.

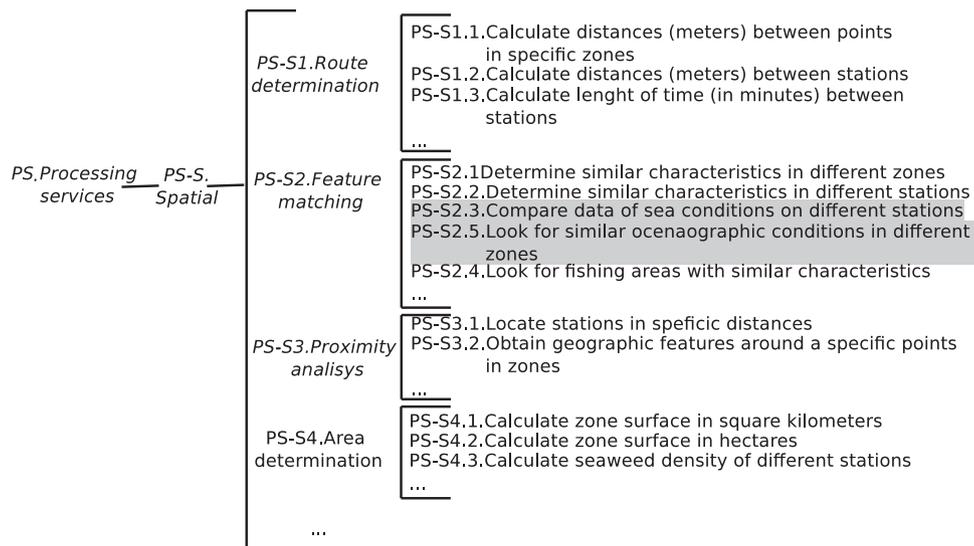


Fig. 8. Services of the marine ecology subdomain for the Geographic Processing – Spatial category.

defined the set of services that interact directly with the users, we explicitly determined the way in which these layers are shown (such as polygons and points). Also, Fig. 6 shows the HI-LM1 (*Layer attributes*) service denoting the specific attributes that must be contained (associated to each layer). For example, the layer showed by the HI-LM5.1.3 (*Stations of specific census by points*) service can contain attributes defined by HI-LM1.4 to HI-LM1.8 services.

Following, in Fig. 7 we show several services created for measuring, grouping and rastering visualization aspects. For example, the HI-LM4 (*Layer grouping*) service was defined to allow users to group different geographic features, such as fishing zones and stations, when processing services, which are described in the next paragraph. These grouped geographic features are used to calculate measures that are shown by services of the HI-LM3 (*Measuring*) category by using labels, tables, or histograms. Another important remark about the HI services is that some of them are denoting the same functionality but with different visual representations. This happens because user experts (perhaps be at different organizations) have different visual requirements

according to their specific work. For instance, the HI-LM1.9 and HI-LM1.10 services (Fig. 6) show biological data of stations by using histograms or tables respectively. Finally, the services of the figures that are highlighted with gray color are representing oceanographic services. For example, the HI-LM1.1 service (*Show specific depth by labels*) of Fig. 6 applies to any subdomain included in the oceanography domain.

Then, in Fig. 8 we show part of the services for the marine ecology subdomain included in the Geographic Processing – Spatial category.

The four services (from PS-S1 to PS-S4)²⁶ represent processes used to analyze data spatially. For example, the PS-S1.1 service is defined to calculate the distance in meters between two points (indicated by a user) within specific zones (represented by a layer) of the map, and the PS-S4.1 and PS-S4.2 services calculate the surface of a zone or group of zones in different formats (square

²⁶ The services written in italic are part of the ISO 19119 std.

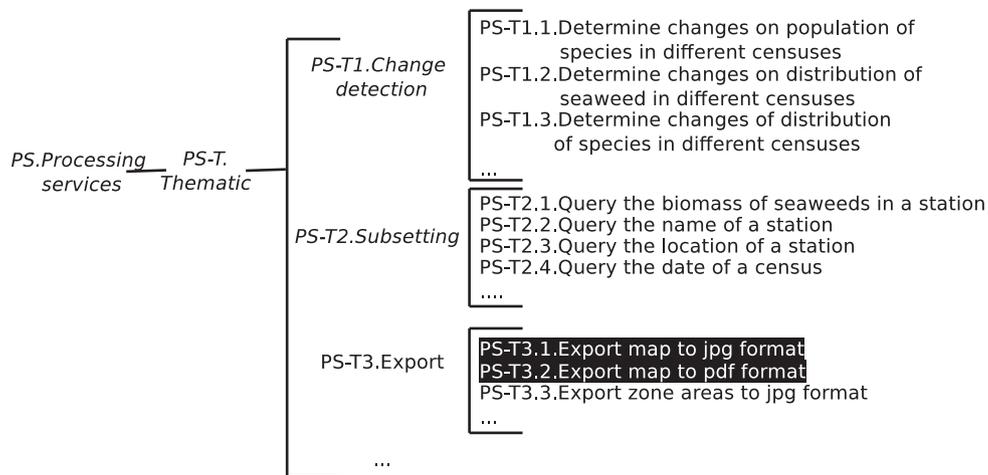


Fig. 9. Services of the marine ecology subdomains for the Geographic Processing – Thematic category.

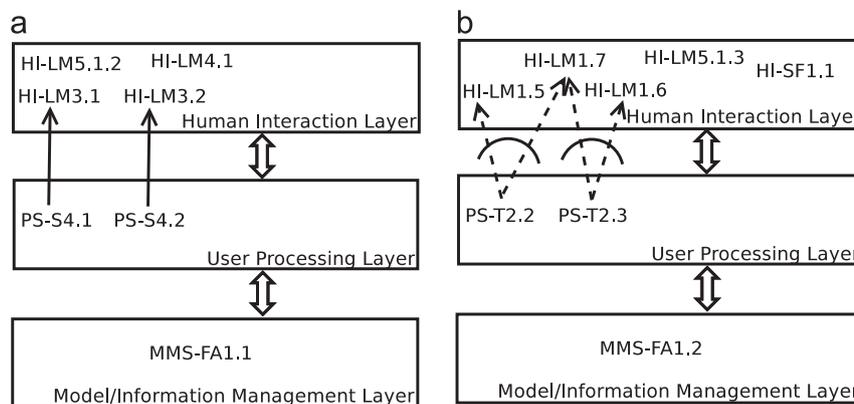


Fig. 10. Set of services within the architecture to implement two generic functionalities. (a) Calculate zone surface in different measures and (b) query station attributes.

kilometers or hectares). The PS-S2 service (*Feature matching*) includes services defined to compare geographic and thematic data in order to find similarities among them. For example the PS-S2.5 service (*Look for oceanography conditions in different zones*) is used to analyze different aspects in similar zones of a region such as the population or distribution of species (animals or seaweeds).

Finally, in Fig. 9 we show part of the services for the marine ecology subdomain included in the Geographic Processing – Thematic category. The two first categories (PS-T1 and PS-T2) are part of the ISO 19119 std and have been specialized to include services for analyzing changes among populations, distribution and characteristics of animals, seaweeds, etc., and query thematic attributes of the main classes of the conceptual model (Fig. 2). The last category showed in the figure includes services to export maps or parts of a map into an image in different formats such as jpg and pdf. Services in black denote services belonging to the geographic domain.

Similarly, we have also specialized the categories defined by the Geographic processing-temporal (PS-E) and Geographic model/information management (MMS) services into particular services of the marine ecology subdomain. For example, for this last category we included the set of services needed to store and query geographic features by specializing the *feature* (MMS-FA) and *map access* (MMS-MA) services.

4.3. Guides for composing services during SPL design and product implementation

The service taxonomy defines a set of services by describing specific tasks according to the nature of the category in which they

are defined. These services are semantically described, which allow software engineers and user experts (such as biologists) to understand the precise task that they must perform. However, it is necessary to know the way in which these services can be combined for implementing specific functionalities of a GIS. Thus, we define the last semantic resource consisting of guides showing interactions among the services by considering the three-tier reference architecture (Fig. 1), in which the first layer corresponds to the HI-Human Interaction services (Figs. 5–7), the second one to the PS-User Processing services (Figs. 8 and 9), and the third one to the MMS-Model/information Management services.

The guides represent generic functionalities that can be used as a basis for implementing a specific function of a particular system. Fig. 10²⁷ shows the set of services of an architecture involved in two generic functionalities. The first one (Fig. 10a) shows the *calculate zone surface in different measures* function in which seven services of the service taxonomy are used. The functionality uses PS-S4.1 (*calculate zone surface in square kilometers*) and PS-S4.2 (*calculate zone surface in hectares*) services to coordinate the process of allowing users to group a set of zones (HI-LM4.1) and show the surface in two different measures: square kilometers (HI-LM3.1) and hectares (HI-LM3.2) by using labels in the map. The arrows from the services of the user processing layer to two services of the human interaction layer represent direct dependencies between services; that is, PS-S4.2 *requires* HI-LM3.2

²⁷ Some of the generic functionalities have been simplified in order to use only the services of the service taxonomy described in this paper.

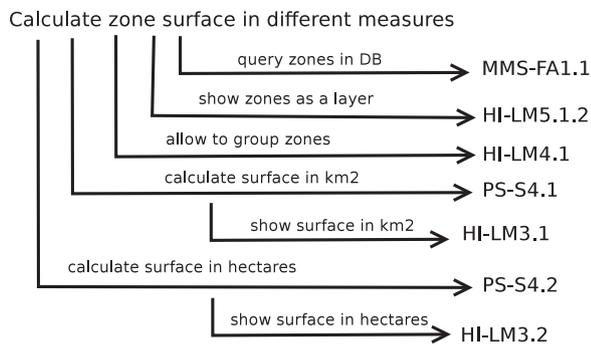


Fig. 11. Service interaction to implement the function *calculate zone surface in different measures*.

to show the result as a label.²⁸ Another functionality (*query station attributes*) is shown in Fig. 10b involving eight different services of different categories of the service taxonomy. The dotted lines represent alternative choices for showing station attributes. For example, the PS-T2.2 (*query the name of a station*) service can require HI-LM1.5 (name by labels) or HI-LM1.7 (name and location by tables) services depending on specific user requirements. Thus, these generic functionalities can be reduced to implement specific functions of a system. For example, if a particular system must show the zone surface in only square kilometers, the PS-S4.2 and HI-LM3.2 services will not be used; or if the station name must be shown as a label, the HI-LM1.7 service will not be used either. In both functionalities modeled in Fig. 10, MMS-FA1.1 and MMS-FA1.2 are specialized services of the geographic model/information management category representing access to the database for zones and stations respectively. HI-SF1.1 (Fig. 10b) is a service of the human interaction category allowing users to select specific geographic features, in this case, a station.

Fig. 11 shows the sequence of service interactions to fulfill the functionality *calculate zone surface in different measures*. As we can see, the services are the same as in Fig. 10a, and here we show in which order and how they can be used. The figure shows all the processes involving the services to show the zones as a layer in a map, group, and finally calculate and visualize the results. Thus, firstly the function calls the MMS-FA1.1 service to find all zones in the database and then show them as a layer (HI-LM5.1.2). This allows a user to group specific zones by using the mouse over the layer (HI-LM4.1). Then, the calculus must be performed by calling the PS-S4.1 and PS-S4.2 services and the results are shown by calling the HI-LM3.1 and HI-LM3.2 services respectively.

As we can see from these two examples, the guides have the main goal of showing how services in the taxonomy are able to interact with each other. We have defined them as general as possible by using every service that could be useful or necessary to implement feasible functionalities. The processing services (PS) have been used as starting points to define the guides even though it does not mean that one processing service must derive in only one guide. For example the functionality *calculate zone surface in different measures* (Fig. 11) involves two PS services. Therefore, any system implemented either by an SPL or CBSD could use these guides by removing (when appropriate) services which are not necessary to implement a specific functionality.

The service taxonomy and the guides were used during an SPL development described in Buccella et al. (2013). From this experience, we briefly describe our validation in the next section.

Table 1
The assessment framework.

Dimensions	Categories	Attributes
Usability	Knowability	Clarity of E&S
		Consistency of E&S
	Operability	Helpfulness
Completeness		
Functionality	Efficiency	Configurability
		Efficiency HE
	Suitability	Efficiency TET
		Completeness
		Bound variability

5. Evaluating the service taxonomy and guides

After building the SPL, we analyzed usability and functionality dimensions of the semantic resources defined in the last section (service taxonomy and guides) in order to measure subjectively how these resources can improve activities of a software development process and help create suitable products. To do so, we divided the project team into three groups, biologists, software engineers and developers, in order to obtain different perspectives of these dimensions. Specifically, the project team was made up of six software engineers (the same as those who participate in the taxonomy team), six developers and five experts users (three of them also participated in the taxonomy team).

The applied methodology (Buccella et al., 2013) defined two main analyses as part of the domain engineering phase: *domain* and *organizational*, with the following specific activities and tasks (as we have described in Section 3):

- *Domain analysis*: Information source analysis (ISA), subdomain analysis and conceptualization (SAC) and reusable component analysis (RCA).
- *Organizational analysis*: Reuse and boundary analysis (RBA) and platform analysis and design (PAD).

Then, during the application engineering phase (Pohl et al., 2005), we considered the following activities: *Application requirements engineering (ARE)*; *application design (AD)*; and *application testing (AT)*.

Our assessment framework analyzes the service taxonomy and guides with respect to their usability and functionality. With respect to assessing *usability*, there exist in the literature a wide range of approaches such as those based on the main standards in this area: the ISO 9241-11²⁹ and ISO 9126-1.³⁰ An exhaustive study of these approaches is provided in Alonso-Ríos et al. (2010), in which authors define a classification of usability attributes to be taken into account in the development of usable systems. In our work, we extracted the usability attributes needed for evaluating the service taxonomy and guides as support for the construction of the SPL and its derived products in the marine ecology subdomain. These attributes conform our usability framework, as shown in the first dimension of Table 1, in which we define three categories. These categories and their attributes are defined according to the use of the service

²⁸ We use a notation similar to the orthogonal variability models proposed in Pohl et al. (2005), to represent variable services depending on the future implementations of particular systems.

²⁹ ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability.

³⁰ ISO/IEC 9126-1:2001 Software engineering – Product quality – Part 1: Quality model.

taxonomy and guides by the users.³¹ The categories are the following:

- **Knowability:** Property by means of which the user can understand, learn, and remember how to use the *service taxonomy and guides* within the marine ecology subdomain.
 - *Clarity of elements and structure:* Are the elements (specific services and guides) and the structure of the *service taxonomy and guides* organized in a way that enables their meaning to be easily understood?
 - *Consistency of elements and structure:* Are the elements (specific services and guides) and the structure of the *service taxonomy and guides* defined with enough uniformity and coherence?
 - *Helpfulness:* Can the *guides* help users when they cannot infer or remember how to use and combine the services?
- **Operability:** Capacity of the *service taxonomy and guides* to provide users with the necessary descriptions within the marine ecology subdomain.
 - *Completeness:* Are the elements of the *service taxonomy* and the *guides* enough to describe the subdomain and their tasks?
 - *Configurability:* Are the *guides* capable of allowing users to personalize the subdomain tasks?
- **Efficiency:** Capacity of the *service taxonomy and guides* to help produce appropriate results according to the invested effort.
 - *Efficiency in human efforts:* Are the *service taxonomy and guides* capable of helping produce appropriate results in return for the physical or mental effort that users invest?
 - *Efficiency in task execution time:* Is time invested by users appropriate to determine subdomain tasks by using the *service taxonomy and guides*?

Finally, with respect to *functionality* (second dimension of Table 1), the assessment framework analyzes the suitability (as defined in ISO 9126-1) in order to determine the appropriateness of the software design (based on domain requirements) with respect to the functionality provided by products derived from the SPL. This category is divided into *suitable completeness*, to measure if all guides defined and personalized in the design activities are implemented in the product; and *suitable bound variability* to measure if the variability is correctly bound. In this two measures, we must evaluate the resultant functionality of the generated product in order to analyze whether the instantiated variability resulted in the implementation of the specific requirement of an organization.

It is important to highlight the scope of each attribute of the assessment framework with respect to the resources (service taxonomy and guides). For example, *completeness* is defined as a characteristic of usability for assessing both resources. Otherwise, *helpfulness* applies only to the guides because it evaluates combined services which are implemented by the guides. The same happens with the suitability category because only the guides show the service interactions to provide the correct functionality. As a summary, Table 2 shows which resources are scoped by which attributes.

5.1. Evaluating usability and functionality in the development process

The activities of creating the SPL and the derived products (see Section 3) in the marine ecology subdomain (by using the service taxonomy and guides) were analyzed according to our assessment

Table 2
Attributes applied to resources.

Dimensions	Categories	Attributes	Resources	
			Service taxonomy	Guides
Usability	Knowability	Clarity of E&S	X	X
		Consistency of E&S	X	X
		Helpfulness		X
	Operability	Completeness	X	X
		Configurability		X
		Efficiency HE	X	X
Functionality	Suitability	Efficiency TET	X	X
		Completeness		X
		Bound variability		X

framework. Table 3 shows which attributes were evaluated for each activity. For example, *clarity and consistency of elements and structure* and *efficiency in human efforts and task execution time* are evaluated on the same five activities because they involved analysis and design tasks of the development process. The only difference is that the *efficiency in human efforts* attribute is not evaluated on the *information source analysis (ISA)* activity because it only obtained a preliminary set of requirements based on the source information. Thus, the results of the use of the taxonomy cannot be evaluated in this stage. Another example can be seen with the *reuse and boundary analysis (RBA)* and *platform analysis and design (PAD)* activities in which the helpfulness, completeness, and configurability attributes are applied. These attributes evaluate the usefulness of guides in their tasks.

To perform the assessment, we elaborated a series of surveys, one for each activity and targeted group of the project team (biologists, software engineers, and developers), including questions about different aspects of the tasks related to the attributes of the framework. The number of participants was 136 considering the eight surveys all together. From them, almost 100% answered the surveys. They were delivered at the end of each activity and had to be answered before starting a new one. Each of them consisted of simple multiple choice questionnaire in which participants should select one of the options which in general evaluated the scale of agreement (such as “strongly disagree”, “moderately disagree”, “neutral”, “moderately agree” and “strongly agree”) of specific tasks or results obtained on each activity of the SPL development. From these answers we projected the subjective degree of acceptance of each usability and functionality aspect.

For example, in the *subdomain analysis and conceptualization (SAC)* activity (of the domain analysis), biologists and software engineers worked together to define the domain model. It was developed by using the conceptual model defined in the taxonomy (Fig. 2) as a reference model from which software engineers and domain experts could discuss possible changes, reductions, or extensions to represent data that they manipulate frequently. Here, also services of the MMS category were discussed in order to determine which specific data were needed. At the same time, by analyzing the service taxonomy, the participants defined the preliminary list of services according to the domain level analyzed, and modeled them in service templates (see Buccella et al., 2013). These templates were used as a design tool to describe each required functionality for the line. They contain the id of the service, its name, the type (inherited, when it is a general feature belonging to upper domains; or specific, when it is specific of the domain-level in which it is defined), the set of open source tools that can implement it, and the software artifact used to represent its functionality. In this last item, the guides were used to understand the interactions of different services and to model them into some software artifacts such as sequence or components diagrams. For instance, the *calculate zone surface in different*

³¹ In our analysis, the “users” are any group of members of the project team, biologists, software engineers and/or developers.

Table 3
Evaluated attributes on activities.

Dimensions	Categories	Attributes	Domain analysis			Organizational analysis		Product development		
			ISA	SAC	RCA	RBA	PAD	ARE	AD	AT
Usability	Knowability	Clarity of E&S	X	X	X			X	X	
		Consis. of E&S	X	X	X			X	X	
		Helpfulness			X	X	X	X	X	
	Operability	Completeness		X	X	X	X	X	X	
		Configurability				X	X		X	
Efficiency	Efficiency HE		X	X			X	X		
	Efficiency TET		X	X	X		X	X		
Functionality	Suitability	Completeness								X
		Bound variab.								X

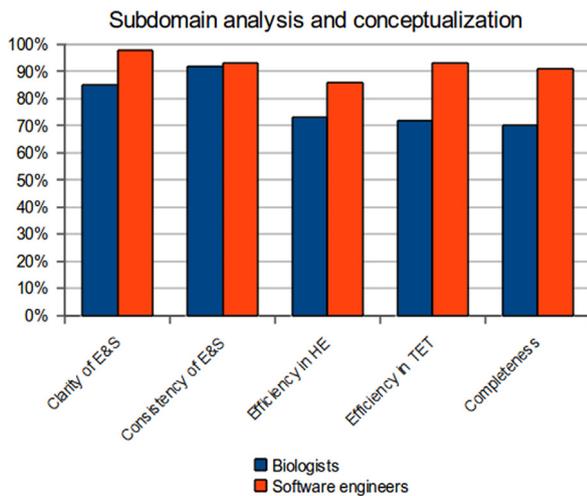


Fig. 12. Usability analysis of the *subdomain analysis and conceptualization* activity.

measures function presented in Fig. 10a was instantiated by the SPL to calculate and visualize the information only in square kilometers (PS-S4.1 and HI-LM3.1 services). Then, the collaboration diagram showing the service interactions to perform the generic functionality (Fig. 11) was simplified to contain only the required services and was used to represent the final functionality of the SPL.

At the end of the *subdomain analysis and conceptualization* (SAC) activity, a questionnaire was answered by each participant generating the results shown in Fig. 12. In this case we evaluated five usability attributes with respect to the use of the service taxonomy and guides (Table 3). In the figure we can see that the evaluation generated good results; all values are above a 70%. However, software engineers' analysis produced upper results than biologists' analysis. Lowest values correspond to the analysis of the efficiency attributes scored by the biologists. This happens because they exhibited difficulties in understanding graphical notations of the guides and needed more time to become familiar with them. Biologists understood the particular services and the structure of the taxonomy but they exhibited special difficulties in combining them together.

In the *reuse and boundary analysis* (RBA) activity (of the organizational level) the most important task was the addition of the variability to the services modeled in previous activities. As an example of the result of this activity, we can see in Fig. 13³²

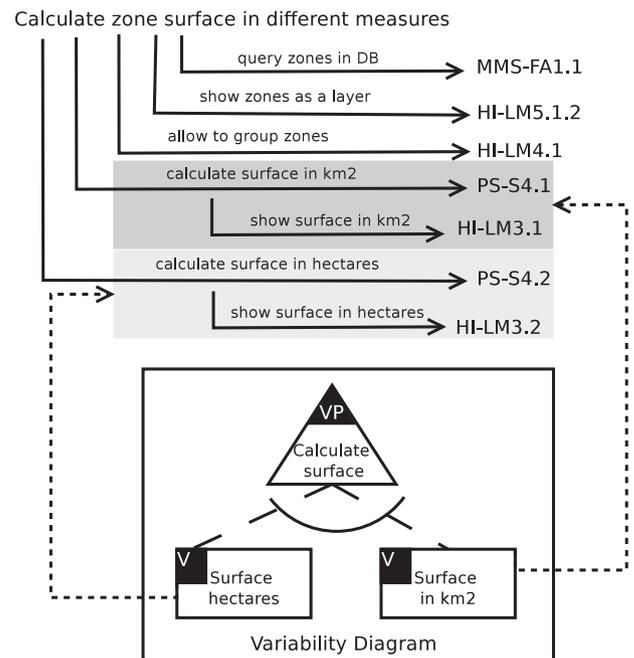


Fig. 13. Collaboration diagram with a variability model item to implement the *calculate zone surface in different measures* service.

the collaboration diagram of the same service showed in section 4.3 (in Figs. 10a and 11), *calculate zone surface in different measures* service. The figure shows graphically the variability defined for zone surface calculus. Then, in the application engineering process (Pohl et al., 2005), this variability will be instantiated according to the specific requirements of products (in square kilometers or hectares).

In this activity, as we can see in Table 3, we analyzed the helpfulness, completeness, and configurability attributes only for software engineers participants (biologists did not participate here). We got satisfactory results (all of them above 90%) showing that participants could understand and apply the guides, by adapting them to the design of variabilities in the SPL development.

In the *platform analysis and design* (PAD) activity (of the organizational analysis), software engineers and developers refined the reference architecture created in the *reusable component analysis* activity. The structures of the services proposed in the guides, for the use of the service taxonomy (Fig. 10), were useful to define and place the reusable components. Thus, the result of this activity was a new reference architecture with explicit mappings between services in the taxonomy and components of the architecture.

³² We used again the notation of variability models proposed in Pohl et al. (2005) due to its clarity on defining variability over UML software artifacts.

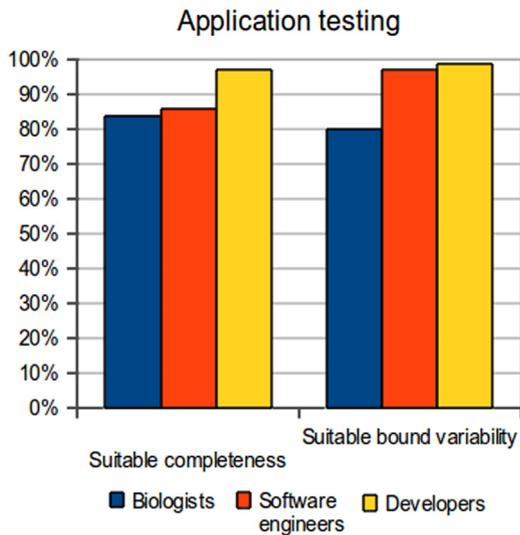


Fig. 14. Functionality analysis of the *application testing* activity.

The *platform analysis and design (PAD)* activity was analyzed with respect to helpfulness, completeness, and configurability attributes by considering the software engineer's and developer's point of view (Table 3). Here, the analysis showed lower results from the developers' than software engineers' because they had not previous knowledge (at this point) of the use of the guides.

Finally, we can show the *application testing (AT)* activity (of the application engineering phase) which represents the last step in the development of a new product. In this case, we analyzed quality attributes for the creation of one product (for the IBMPAS organization). Fig. 14 shows this analysis in which only the functionality attribute was considered. This analysis emerged from the use of the product in execution-time when participants had to analyze the suitability with respect to completeness and bound variability.³³ As we can see, developers returned better results because they implemented the product according to their understanding of the analysis and design specifications. On the other hand, biologists found that some services were not implemented as they expected in design time and some variabilities were bound incorrectly or were missing.

6. Conclusion and future work

In this paper, we have described three semantic resources which contribute for achieving an effective software reuse in an SPL development. Firstly, we briefly introduced a standard conceptual model, specially for the marine ecology domain, and shared by all product generated in the line. Then, based on this conceptual model and in order to provide a common language and information structure, we have defined a standard service taxonomy which includes the possible computer services involved in this domain. This taxonomy helps stakeholders to bridge the gap among their different skills by reducing the wide spectrum of information sharing. Thirdly, for supporting the use of the service taxonomy, we have created a set of guides showing the interaction among services for fulfilling specific functionalities of products in this domain. These guides are defined as general as possible in order to make them flexible enough to be instantiated in the generation of a new product.

Finally, the use of the taxonomy and guides was subjectively assessed during an SPL development process. Results from this evaluation lead us to remark the following:

- Clarity, consistency, and efficiency of the service taxonomy and guides allowed software engineers and biologists to work in a structured and guided way for defining the first activities of an SPL development (analysis and design activities). Here, the service taxonomy in particular worked as a controlled vocabulary for all participants providing a common language (which result in better communication). Thus, time for learning domain particularities was reduced.
- Completeness, helpfulness, and configurability of the guides allowed software engineers, developers, and biologists to design functionalities of the SPL and products under a controlled range of service combinations restricting the universe of possibilities. Therefore, domain and product-specific functionalities were modeled, personalized, and implemented according to the particular needs of the products to be developed.
- Functionality of final products was tested by analyzing the service and guide specifications against functions of running products. In our work, software engineers, developers, and biologists participated in the early stages of product development by personalizing the guides according to their own requirements, and thus improving the identification of these functionalities in a working system.

As a future work we are defining process patterns that might support the process of combining the services defined in our taxonomy. These patterns will standardize interactions among services in order to support the development of specific functionalities of a system. At the same time, the standardization will allow the design of automatic tools to support the development of new platforms or products.

Acknowledgments

We would like to thank the two organizations that are collaborating in this project: IBMPAS and CENPAT-CONICET. This work is partially supported by the UNComa project "Reuso de Software orientado a Dominios" part of the program "Desarrollo de Software Basado en Reuso" and by the PAE-PICT 2007-02312 "Métodos y Herramientas para Software Masivamente Distribuidos".

References

- Albrecht, J., 1996. Universal gis operations for environmental modeling. In: Proceedings of the Third International Conference/Workshop on Integration GIS and Environmental Modeling. Santa Fe, NM.
- Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., Moret-Bonillo, V., 2010. Usability: a critical analysis and a taxonomy. *Int. J. Hum.-Comput. Interact.* 26 (1), 53–74.
- Bai, Y., Di, L., Wei, Y., 2009. A taxonomy of geospatial services for global service discovery and interoperability. *Comput. Geosci.* 35 (April (4)), 783–790, url: <http://dx.doi.org/10.1016/j.cageo.2007.12.018>.
- Bonnett, A., 2008. *What is Geography?*. SAGE Publications Limited, London, Great Britain.
- Bosch, J., 2000. *Design and Use of Software Architectures: Adopting and Evolving a Product-line Approach*. ACM Press/Addison-Wesley Publishing, Co., New York, NY, USA.
- Bruno, D., Richmond, H., 2003. The truth about taxonomies. *Inf. Manag. J.* 37 (2).
- Buccella, A., Cechich, A., Arias, M., Pol'la, M., Doldan, S., Morsan, E., 2013. Towards systematic software reuse of gis: insights from a case study. *Comput. Geosci.* 54 (0), 9–20, url: <http://www.sciencedirect.com/science/article/pii/S0098300412003913>.
- Choksy, C.E.B., 2006. 8 steps to develop a taxonomy. *Inf. Manag. J.* 40 (6), 30–41.
- Cisco, S.L., Jackson, W.K., 2005. Creating order out of chaos with taxonomies. *Inf. Manag. J.* 20, 45–50.
- Clements, P.C., Northrop, L., 2001. *Software Product Lines: Practices and Patterns*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

³³ Metrics with respect to functionality were used in this analysis.

- Czarnecki, K., 2002. Domain Engineering. In: John J., Marciniak (Ed.), *Encyclopedia of Software Engineering*. John Wiley & Sons, Inc., New York, NY, USA (Chapter 3). URL: (<http://dx.doi.org/10.1002/0471028959.sof095>).
- Czarnecki, K., Helsen, S., Eisenecker, U.W., 2005. Formalizing cardinality-based feature models and their specialization. *Softw. Process: Improv. Pract.* 10 (1), 7–29.
- ESPRIT/ESSI, 1998. Project No 21580. Guidelines for Best Practice in User Interface for GIS. Section 6 List of Key GIS Operations. (<http://www.gisig.it/best-gis/guides/chapter6/six.htm>).
- Frakes, W.B., Isoda, S., 1994. Success factors of systematic reuse. *IEEE Softw.* 11 (September (5)), 14–19. url: (<http://dx.doi.org/10.1109/52.311045>).
- Hunink, I., Rene, E., Jansen, S., Brinkkemper, S., 2010. Industry taxonomy engineering: the case of the european software ecosystem. In: *Proceedings of the Fourth European Conference on Software Architecture: Companion Volume (ECSA '10)*. ACM, New York, NY, USA, pp. 111–118. URL: (<http://doi.acm.org/10.1145/1842752.1842779>).
- Kang, K., Cohen, S., Hess, J., Nowak, W., Peterson, S., 1990. Feature-Oriented Domain Analysis (FODA) Feasibility Study. Technical Report CMU/SEI-90-TR-21, Software Engineering Institute, Carnegie Mellon University Pittsburgh, PA.
- Matinlassi, M., 2004. Comparison of software product line architecture design methods: Copa, fast, form, kobra and qada. In: *Proceedings of the 26th International Conference on Software Engineering (ICSE '04)*. IEEE Computer Society, Washington, DC, USA, pp. 127–136.
- Nickerson, R.C., Varshney, U., Muntermann, J., Isaac, H., 2009. Taxonomy development in information systems: developing a taxonomy of mobile applications. In: *17th European Conference on Information Systems (ECIS 2009)*, Italy, pp. 1138–1149.
- Noorian, M., Bagheri, E., Du, W., 2012. Non-functional properties in software product lines: a taxonomy for classification. In: *Proceedings of the 24th International Conference on Software Engineering and Knowledge Engineering (SEKE'2012)*. Knowledge Systems Institute Graduate School, Redwood City, San Francisco Bay, USA, pp. 663–667.
- Papazoglou, M.P., Traverso, P., Dustdar, S., Leymann, F., 2007. Service-oriented computing: state of the art and research challenges. *Computer* 40 (11), 38–45.
- Parent, C., Spaccapietra, S., Zimányi, E., 1999. Spatio-temporal conceptual models: data structures + space + time. In: *Proceedings of the 7th ACM International Symposium on Advances in Geographic Information Systems (GIS'99)*. ACM Press, New York, NY, USA, pp. 26–33.
- Pernich, P., Buccella, A., Cechich, A., Doldan, S., Morsan, E., 2010. Reusing geographic e-services: a case study in the marine ecological domain. In: Cellary, W., Estevez, E. (Eds.), *Software Services for e-World*. IFIP Advances in Information and Communication Technology, vol. 341. Springer, Boston, pp. 193–204 URL: (http://dx.doi.org/10.1007/978-3-642-16283-1_23).
- Pernich, P., Buccella, A., Cechich, A., Doldan, S., Morsan, E., Arias, M., Pol'la, M., 2012. Product-line instantiation guided by subdomain characterization: a case study. *J. Comput. Sci. Technol.* 12 (3), 116–122 (Special Issue).
- Pohl, K., Böckle, G., Linden, F.J.v.d., 2005. *Software Product Line Engineering: Foundations Principles and Techniques*. Springer-Verlag, New York, Inc., Secaucus, NJ, USA.
- Reed, G., Keeley, R., Belov, S., Mikhailov, N., 2009. Ocean data portal: a standards approach to data access and dissemination. In: *Proceedings of the Sustained Ocean Observations and Information for Society (OceanObs'09)*, vol. 2, ESA Publication WPP-306, Venice, Italy.
- Rees, R.V., 2003. Clarity in the usage of the terms ontology, taxonomy and classification. *Civil Eng.* 20, 1–8.
- Schäfer, W., Prieto-Diaz, R., Matsumoto, M. (Eds.), 1993. *Software Reusability*. Ellis Horwood, New York, NY, USA.
- Sklar, F., Constanza, R., 1991. The development of dynamic spatial models for landscape ecology: a review and prognosis. In: Turner, M., Gardner, R. (Eds.), *Quantitative Methods in Landscape Ecology*. Springer, New York, NY, USA, pp. 239–288.
- Sujatha, R., Bandaru, R., Rao, R., 2011. Taxonomy construction techniques – issues and challenges. *Civil Eng.* 2 (5), 1–8.
- Svahnberg, M., Bosch, J., 1999. Evolution in software product lines: two cases. *J. Softw. Maint.* 11 (November (6)), 391–422 URL: ([http://dx.doi.org/10.1002/\(SICI\)1096-908X\(199911/12\)11:6<391::AID-SMR199>3.0.CO;2-8](http://dx.doi.org/10.1002/(SICI)1096-908X(199911/12)11:6<391::AID-SMR199>3.0.CO;2-8)).
- Svahnberg, M., van Gurp, J., Bosch, J., 2005. A taxonomy of variability realization techniques: research articles. *Softw.-Pract. Exp.* 35 (July (8)), 705–754 URL: (<http://dx.doi.org/10.1002/spe.v35:8>).
- Szyperski, C., 1998. *Component Software – Beyond Object-oriented Programming*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA (ISBN: 0-201-17888-5).
- van der Linden, F., Schmid, K., Rommes, E., 2007. *Software Product Lines in Action: The Best Industrial Practice in Product Line Engineering*. Springer-Verlag, New York, Inc., Secaucus, NJ, USA.