Computers & Geosciences 73 (2014) 108–121

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

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.

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 (SPLE) (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).

http://dx.doi.org/10.1016/j.cageo.2014.09.004
0098-3004/

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 a 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 branches[2] (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 Standardization[3] (ISO), and more specifically, the Open Geospatial Consortium[4] (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

[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

[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.