

Una Plataforma para el Desarrollo de LPS en el Dominio Geográfico: Caso de Estudio en el Subdominio de Paleontología

Fiorella Pesce¹ *, Sofia Caballero¹, Agustina Buccella^{1,2}, and Alejandra Cechich¹

¹ GIISCO Research Group

Departamento de Ingeniería de Sistemas - Facultad de Informática

Universidad Nacional del Comahue

Neuquen, Argentina

{fiorella.pesce,sofia.caballero,agustina.buccella,alejandra.cechich}@fi.uncoma.edu.ar

² CONICET

Abstract. La Líneas de Productos de Software (LPS) son un paradigma orientado al reuso de software dentro de dominios particulares. Aspectos claves dentro de este paradigma son las particularidades inherentes de estos dominios y las técnicas aplicadas para sistematizar las formas de maximizar el reuso dentro de los mismos. En este artículo describimos un proceso para la creación de líneas de productos orientado al reuso de jerarquías de dominios, partiendo del dominio geográfico en general y considerando el subdominio paleontológico en particular. Nuestro proceso se basa en estandarizaciones y técnicas previas ya aplicadas para otro subdominio geográfico, el de ecología marina. Aquí mostramos cómo estas técnicas son aplicadas en el subdominio de paleontología, mejorando el reuso sistemático de artefactos software.

Keywords: Líneas de Productos de Software, Reuso de Dominios, Dominio Geográfico, Paleontología

1 Introducción

La Ingeniería de Líneas de Productos de Software (ILPS) [6, 13, 18] propone un proceso de desarrollo de software cuyo principal objetivo es maximizar el reuso de artefactos de manera de obtener aplicaciones más rápidas y de mejor calidad. Como el desarrollo de las LPSs propone un reuso dentro de dominios particulares, el éxito del desarrollo depende de la identificación, uso y administración de los artefactos dentro de cada uno de ellos; y por lo tanto, la aplicación de técnicas específicas para la sistematización del reuso es crucial.

El dominio geográfico [11] es amplio ya que incluye aspectos generales aplicables a todos los productos dentro de este dominio. Aquí, la creación de una LPS

*

puede resultar impracticable (la cantidad de variabilidad que habría que definir dentro de cada servicio se volvería inmanejable), por lo que es lógico pensar en la división en diferentes subdominios donde cada uno posea ciertas características del dominio geográfico común, como realizar zoom en un mapa, obtención de coordenadas, etc; y ciertas características especiales como conocer la profundidad del océano en determinadas zonas. De esta forma se construyen componentes pensados para formar parte de una jerarquía de dominios, es decir, se definen un conjunto de servicios generales a nivel del dominio genérico geográfico aplicables a cualquier línea a ser implementada dentro de los subdominios que incluya.

En la literatura las LPSs son un área de investigación muy activa en donde existen muchas propuestas de técnicas o métodos que mejoren las actividades involucradas en su desarrollo. Esto se puede ver en la cantidad de revisiones literarias o sistemáticas que existen hoy en día enfocadas en técnicas de modelado, implementación, validación, etc. [4, 5, 12, 16]. Enfoques similares a nuestra propuesta orientada a una jerarquía de dominios pueden ser aquellos relacionados con el desarrollo de ecosistemas de software o multi-líneas de software [7, 15]. En general, estos trabajos se enfocan en la problemática de haber desarrollado componentes para dominios diferentes que deben luego trabajar juntos al ser integrados dentro de un mismo desarrollo. En nuestro caso, estos problemas son minimizados ya que los componentes creados son diseñados para formar parte de una jerarquía dentro del dominio geográfico compartiendo estandarizaciones ya definidas.

Una primera aproximación a nuestro enfoque ha sido presentado en [9, 10] en donde hemos construido una LPS dentro del subdominio de ecología marina, pero tomando servicios comunes de dominios de más alto nivel como el oceanográfico y el geográfico. En esos trabajos hemos definido una metodología para la creación de LPSs que diseña e implementa una serie de *artefactos de software* necesarios para comunicar y modelar el dominio basado en estandarizaciones y técnicas bien definidas. En este artículo mostramos el proceso de desarrollo para la creación de LPSs, basado en la metodología previamente presentada [9, 10] para el subdominio de ecología marina, pero enfocándonos en el reuso de los artefactos creados y adaptados al subdominio paleontológico. En particular nos centramos en tres de los artefactos de software creados como parte de la ingeniería de dominio.

Este artículo está organizado de la siguiente manera. En la siguiente sección describimos nuestro proceso de desarrollo de LPSs orientado al dominio geográfico y sus subdominios. En la Sección 3 mostramos un caso de estudio en donde aplicamos dicho proceso al subdominio de paleontología en lo relevante a la información asociada a las excavaciones para encontrar piezas paleontológicas. Luego, realizamos una validación preliminar en donde analizamos mejoras debido al reuso efectuado. Por último describimos las conclusiones y trabajos futuros.

2 Proceso de Desarrollo para la creación de Líneas de Productos de Software Orientada a Subdominios

La Figura 1 muestra los cuatro pasos de nuestro proceso para la ingeniería de dominios en el desarrollo de una LPS. Recordemos que esta ingeniería es la que se encarga de identificar, capturar y organizar toda la información fuente recopilada. Como resultado, genera una plataforma de software con un conjunto de artefactos reusables y configurables que proveen un entorno en común [18].

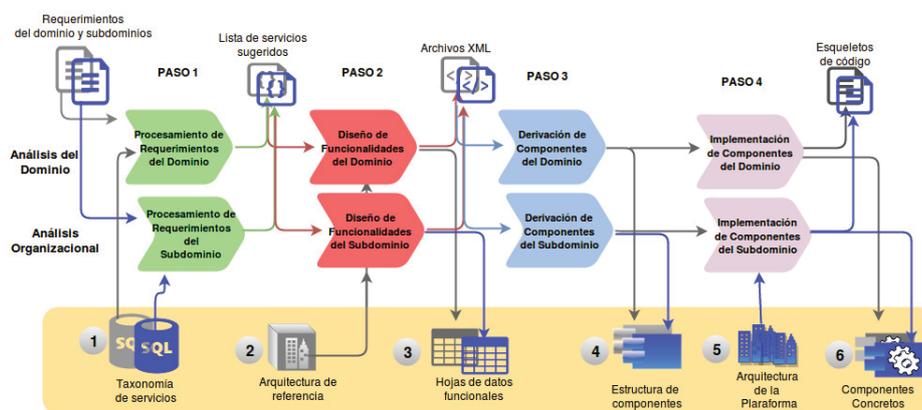


Fig. 1. Proceso de desarrollo como parte de la ingeniería de dominio de las LPSs junto con los artefactos de software creados y utilizados

En nuestra metodología de desarrollo, la ingeniería de dominios se divide en dos tipos de análisis: de *dominio* y *organizacional*. El primero involucra el análisis y diseño de la información dentro de un dominio específico pero enfocándose en una vista general. Luego, el análisis organizacional utiliza la información anterior para adaptarla al contexto de la LPS que está siendo desarrollada. En nuestro caso de aplicación, dentro del análisis de dominio se debe analizar el dominio geográfico general obteniendo servicios comunes al conjunto de subdominios que podrían ser de interés. Luego, en el análisis organizacional, las actividades están enfocadas en el subdominio específico, en nuestro caso, en el subdominio paleontológico.

En la Figura 1 podemos observar los seis artefactos de software que deben ser desarrollados o utilizados en los cuatro pasos involucrados en la ingeniería de dominio. En general cada artefacto es construido durante el análisis del dominio (artefactos en color gris) y refinados y completados en el análisis organizacional (artefactos en color azul). El primero de ellos, el cual es una de las entradas del paso 1, es una *taxonomía de servicios* del dominio analizado construida como una estructura jerárquica, con las categorías por las cuales clasificar los servicios. En trabajos previos, esta taxonomía ha sido creada para el dominio geográfico

considerando el estándar de Servicios de Arquitectura (definido en el OpenGIS Service Architecture)³, y la ISO/DIS 19119⁴ como base para definir los servicios y sus categorías [10]. La taxonomía es creada en base a los requerimientos de los usuarios expertos, en forma general (en el análisis de dominio), y refinada y especializada en el análisis organizacional para contener aquellos servicios particulares del subdominio paleontológico. Luego, la salida de este paso es la lista de servicios requeridos en el subdominio para el diseño de las funcionalidades.

El segundo artefacto es una *arquitectura de referencia* y es utilizada como entrada en el paso 2 de *diseño de funcionalidades*. Dicha arquitectura debe especificar una estructura preliminar para la interacción de los servicios definidos en la taxonomía. Como podemos observar, sólo existe una instancia de la misma en color gris debido a que es extraída y reusada de la especificada en la norma ISO 19119. A su vez, en este paso se deben crear las *hojas de datos funcionales* que poseen cada una de las funcionalidades del dominio geográfico general (en gris) y particular (en azul). Las mismas se diseñan mediante la interacción de los servicios de la taxonomía necesarios. También, cada hoja funcional se representa mediante un conjunto de archivos XML que permiten analizar automáticamente estos modelos en busca de inconsistencias o incompatibilidades a la hora de especificar la variabilidad [8, 20]. Luego, en el paso 3 se debe realizar la *derivación de componentes abstractos* para crear componentes reusables en base a la información definida en las hojas de datos funcionales [1, 19]. Como parte de este paso y el siguiente (*Implementación de componentes*) hemos definido en trabajos previos mecanismos iniciales que asisten a la creación del cuarto artefacto la (*estructura de componentes*) [2].

Luego, los artefactos refinados determinan la estructura del software que conforma el quinto artefacto, la *arquitectura de la plataforma*. Ésta contiene la forma en que cada hoja de datos funcional refinada (en base al refinamiento de la taxonomía) se implementa como componentes de software. Esta plataforma es luego utilizada en la configuración de los productos para crear la arquitectura de una aplicación específica.

3 Un Caso de Estudio en el Subdominio Paleontológico

Primero es necesario introducir algunos aspectos particulares de este subdominio. En la Figura 2 mostramos una parte del modelo conceptual en donde nos enfocamos en las clases que nos sirven para comprender luego las funcionalidades definidas. Para construir el mismo, además de la información elicitada de los usuarios expertos en el subdominio de paleontología, hemos extraído información de las ISO 19109 y 19107 para datos geográficos, y las normas ISO 21127:2014⁵,

³ The OpenGIS Abstract Specification: Service Architecture, 2002.

⁴ Geographic information. Services International Standard 19119, ISO/IEC, 2005.

⁵ Information and documentation – A reference ontology for the interchange of cultural heritage information

LIDO⁶ y CIDOC⁷. Estas tres últimas normas son específicas del subdominio paleontológico ya que definen y clasifican servicios y protocolos requeridos para la gestión de colecciones dentro de los museos y la representación de las piezas paleontológicas encontradas.

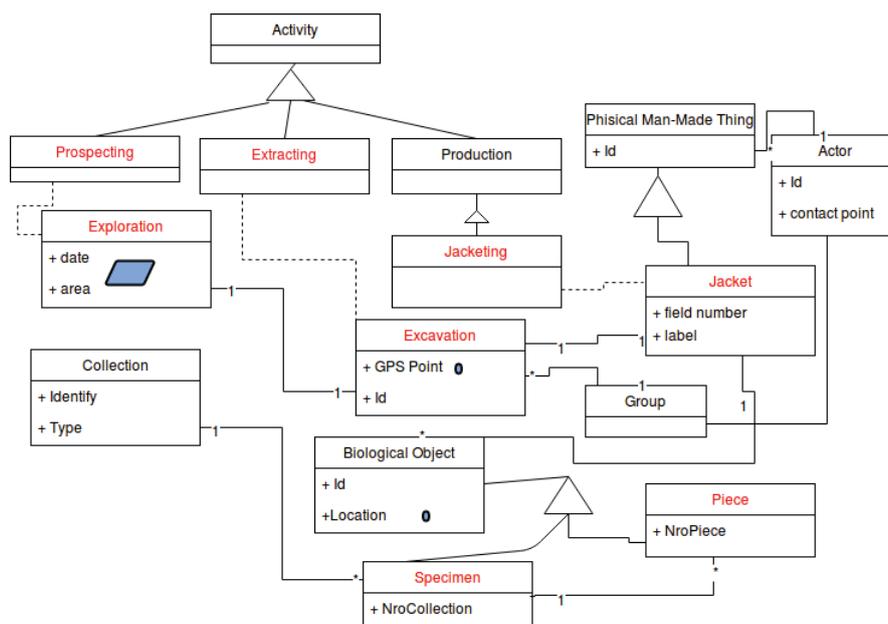


Fig. 2. Parte del Modelo de Datos para el Subdominio Paleontológico

La Figura 2 muestra en color rojo aquellas clases que han sido definidas por los usuarios expertos y en color negro las extraídas de las tres últimas normas. Según la ISO 21127 cualquier acción que se realice sobre un objeto físico, se representa como una actividad. En particular, la actividad de *exploración* (*prospecting*) es la que releva zonas en las que puedan realizarse las *excavaciones* (*extracting*) a cargo de un grupo de profesionales. En el caso de encontrar algún *especimen*, se inicia la actividad de armado de un *bochón* (*jacket*) (mediante la actividad de *jacketing*), el cual puede contener *objetos biológicos* – *especimen* o *pieza*. A éstos se les aplican procedimientos mecánicos y/o químicos para retirar correctamente el objeto encontrado. A su vez, la información que debe ser geográfica, como las áreas y puntos, ha sido representada utilizando el enfoque MADS (Modeling of Application Data with Spatio-temporal features) [17].

⁶ LIDO - Lightweight Information Describing Objects Version 1.0 - <http://network.icom.museum/cidoc/working-groups/lido/what-is-lido/>

⁷ CIDOC Conceptual Reference Model Version 6.0 - <http://www.cidoc-crm.org>

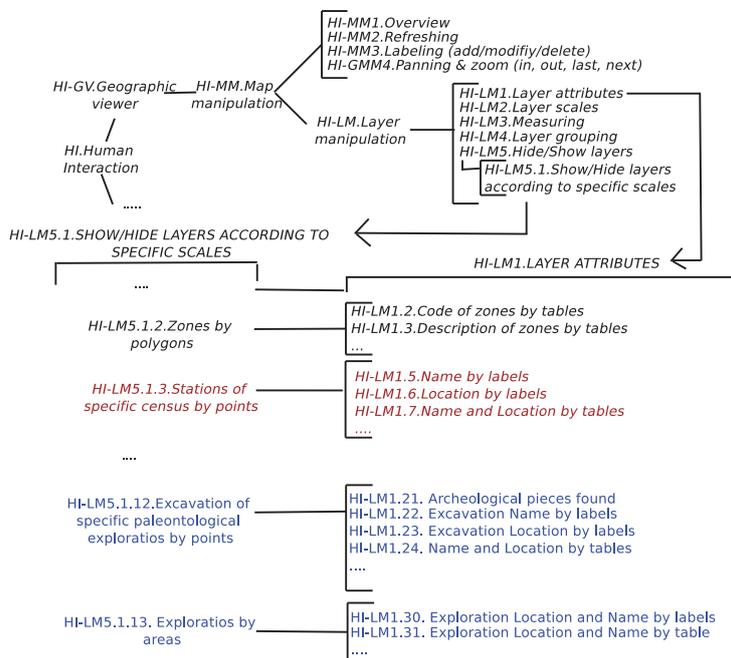


Fig. 3. Algunos servicios de la categoría *interacción humana* (HI) definidos en la jerarquía de dominios

Luego, en base a este modelo conceptual y a los requerimientos funcionales del subdominio creamos la *taxonomía de servicios paleontológicos*. Ésta ha sido definida siguiendo las bases descriptas en [10] mediante la realización de un proceso iterativo de trabajo entre todos los participantes del subdominio, usuarios expertos, ingenieros de software y desarrolladores, y utilizando como base la taxonomía geográfica definida en la ISO 19119. También aquí se utilizaron las normas específicas del subdominio (ya mencionadas previamente) y la arquitectura de referencia de la ISO 19119 ya aplicada para la LPS en el subdominio de ecología marina [9].

La taxonomía, que se define en base a las capas de la arquitectura, posee un conjunto de categorías principales en las cuales deben estar incluidos los servicios correspondientes. Así en la Figura 3 podemos observar algunos servicios de la categoría *interacción humana* (HI) ya definidos previamente [10]. Estos servicios pertenecen al dominio geográfico (en color negro y cursiva), y al subdominio de ecología marina (en color rojo y cursiva). A su vez también están definidos los servicios del subdominio paleontológico (en color azul), en donde podemos observar algunos orientados a las formas de visualización de la información. En este caso vemos los atributos de exploraciones y excavaciones que pueden ser mostrados en forma de tablas o por medio de etiquetas directamente sobre un mapa.

De esta manera, como resultado para cada categoría, se obtiene la taxonomía completa dentro del subdominio paleontológico con servicios reusados y heredados del dominio geográfico.

El siguiente paso es la creación de las hojas de datos funcionales (tercer artefacto) con las funcionalidades del subdominio en base a los servicios de la taxonomía paleontológica y la arquitectura de referencia. La confección de estas hojas de datos sigue los mismos lineamientos definidos en trabajos previos [1, 14] y utiliza nuestra herramienta de soporte llamada *Datasheet Modeler* [14]. Ésta permite crear las hojas mediante la utilización de dependencias de variabilidad [1] y servicios de la taxonomía, traduciendo a archivos XML procesables por la computadora. La Figura 4 muestra la funcionalidad *Registrar excavaciones*, la cual permite registrar una nueva excavación dentro de una exploración ya existente en el sistema. Como podemos observar, se realiza una interacción entre los servicios de la taxonomía ya definida de manera de poder llevar cabo la funcionalidad. La variabilidad representada define en este caso que la carga de datos debe contemplar que pueda realizarse en lugares sin conexiones a internet o dispositivos GPS. Por lo tanto, datos de las excavaciones (tales como coordenadas geográficas y objetos encontrados) deben permitirse mediante el ingreso manual de información. A su vez, en caso se requiera en algún producto, también se contemplan las opciones de cargar los datos que provienen de planillas de cálculo, y/o a través de dispositivos que sí contengan GPS.

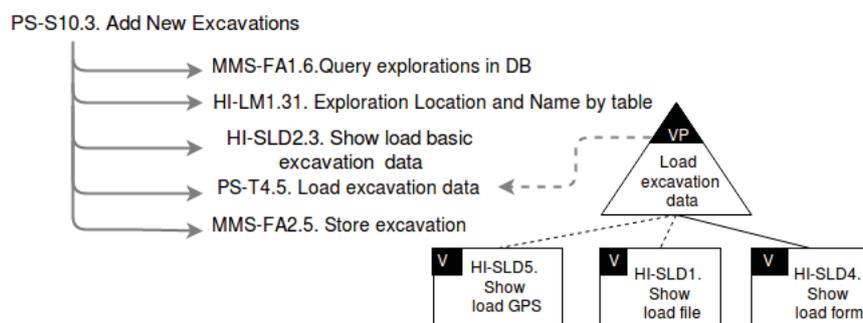


Fig. 4. Representación de la funcionalidad *Registrar Excavaciones* en una Hoja de Datos Funcional

Luego, como salida de la herramienta se crean automáticamente los archivos XML de cada una de estas hojas de datos. Estos archivos nos permiten procesarlas automáticamente para validarlas [19, 20] y crear una estructura preliminar de componentes reusables (artefacto 4) [2, 3], que luego formarán parte de la arquitectura de la plataforma (artefactos 5). Éstos deben ser componentes implementados (artefacto 6) en el último paso.

4 Una Evaluación Preliminar del Reuso Alcanzado

Para realizar una evaluación, hemos realizado un análisis del grado de reuso logrado en dos de los artefactos creados y descritos en la sección anterior.

Primero hemos analizado el reuso dentro de la taxonomía de servicios considerando la cantidad de servicios que hemos definido para el subdominio paleontológico. Para esta evaluación hemos analizado tres tipos de servicios:

- *Servicios completamente reusados (SCR)* que son aquellos que están dentro del dominio geográfico o de algún subdominio y son utilizados sin modificación dentro de las funcionalidades de la LPS paleontológica. Ejemplos de estos servicios son los que están en color negro y rojo en la Figura 3 y que son de incumbencia para este subdominio.
- *Servicios reusados por especialización (SRE)* que son servicios nuevos pero generados a partir de servicios ya definidos para otros dominios. Ejemplos de estos servicios son los de color azul en la Figura 3.
- *Servicios sin reuso (SSR)* que son aquellos creados completamente nuevos sin ninguna relación con los servicios de la taxonomía previamente desarrollada.

En base a esos tres tipos de servicios, y de un total de unos 120 servicios definidos como útiles, identificamos un 25% como SCR, un 58.3 % como SRE y un 16.6% como SSR. La Figura 5a) muestra estos porcentajes gráficamente.

Luego, para analizar las funcionalidades, hemos simulado la creación de un producto en donde elegimos un conjunto de diez funcionalidades (definidas en las hojas de datos funcionales) como parte de la plataforma LPS y las hemos instanciado aleatoriamente, de manera de generar un nuevo producto. Para ilustrar este proceso, en la Figura 6 mostramos la hoja de datos resultante de la funcionalidad *Registrar Excavaciones* en donde hemos instanciado la variabilidad para permitir además de la carga manual por formularios, también la carga por medio de dispositivos GPS. De esta manera, el producto resultante contendrá ambas opciones para la carga de una excavación.



Fig. 5. Análisis del reuso logrado en la taxonomía y las hojas de datos funcionales instanciadas

Este proceso se efectuó por cada una de las 10 funcionalidades que fueron utilizadas para la creación del producto. Para analizar el reuso hemos contado la

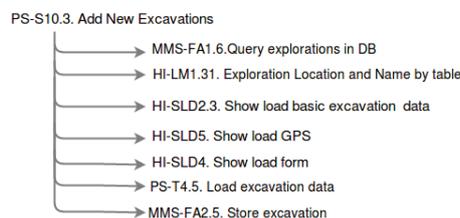


Fig. 6. Funcionalidad *Registrar Excavaciones* instanciada para un producto en particular

cantidad de *servicios de otros dominios (SOD)* utilizados en la funcionalidades y la cantidad de *servicios propios del dominio (SPD)* de la taxonomía. Los SOD son aquellos que hemos clasificado como la suma de los SCR y SRE y los SPD son los SSR. De este análisis, surgen los resultados mostrados en la Figura 5b) en donde podemos observar un porcentaje de aprox. un 65% de SOD y un 35% de SPD. Este resultado resulta razonable considerando que muchas funcionalidades utilizan servicios SRE (especializados para el subdominio).

5 Conclusiones y Trabajo Futuro

En este trabajo hemos mostrado un proceso para el desarrollo de LPS orientado a reuso de artefactos de software dentro de una jerarquía de dominios. A partir de trabajos previos dentro del dominio geográfico, hemos generado y reusado artefactos en un nuevo subdominio en este caso paleontológico. A su vez hemos realizado un análisis preliminar del reuso logrado en la construcción de alguno de esos artefactos mostrando resultados prometedores. Como trabajo futuro, entenderemos el proceso hacia los artefactos que están involucrados en la ingeniería de la aplicación de las LPSs.

References

1. A. Buccella, M. Pol'la, A.C., Arias, M.: A variability representation approach based on domain service taxonomies and their dependencies. In: International Conference of the Chilean Computer Science Society (SCCC). Talca, Chile (2014)
2. Arias, M., Buccella, A., Cechich, A.: Smooth transition from abstract to concrete spl components: a client-server implementation for the geographic domain. In: Proceedings of the IEEE ARGENCON'16. GRSS: IEEE GRSS, Buenos Aires, Argentina (2016)
3. Arias, M., Buccella, A., Cechich, A.: Un marco de trabajo para la gestión de requerimientos de software en líneas de productos de software. In: CLEI 2017 : XLIII Conferencia Latinoamericana en Informática. SLISW - Simposio Latinoamericano de Ingeniería de Software. To appear (2017)
4. Bashroush, R., Garba, M., Rabiser, R., Groher, I., Botterweck, G.: Case tool support for variability management in software product lines. *ACM Comput. Surv.* 50(1), 14:1–14:45 (Mar 2017), <http://doi.acm.org/10.1145/3034827>

5. Benavides, D., Segura, S., Ruiz-Cortés, A.: Automated analysis of feature models 20 years later: A literature review. *Information Systems* 35, 615–636 (September 2010)
6. Bosch, J.: Design and use of software architectures: adopting and evolving a product-line approach. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA (2000)
7. Bosch, J.: From software product lines to software ecosystems. In: Proceedings of the 13th International Software Product Line Conference. pp. 111–119. SPLC '09, Carnegie Mellon University, Pittsburgh, PA, USA (2009)
8. Braun, G., Pol'La, M., Cecchi, L., Buccella, A., Fillottrani, P., Cechich, A.: A dl semantics for reasoning over ovm-based variability models. In: 30th International Workshop on Description Logics (DL17). To appear, France (2017)
9. Buccella, A., Cechich, A., Arias, M., Pol'la, M., Doldan, S., Morsan, E.: Towards systematic software reuse of gis: Insights from a case study. *Computers & Geosciences* 54(0), 9 – 20 (2013), <http://www.sciencedirect.com/science/article/pii/S0098300412003913>
10. Buccella, A., Cechich, A., Pol'la, M., Arias, M., Doldan, S., Morsan, E.: Marine ecology service reuse through taxonomy-oriented SPL development. *Computers & Geosciences* 73(0), 108 – 121 (2014), <http://www.sciencedirect.com/science/article/pii/S0098300414002155>
11. Burrough, P., McDonnell, R.: Principles of Geographical Information Systems. Oxford University Press (1998)
12. Capilla, R., Snchez, A., Dueas, J.C.: An analysis of variability modeling and management tools for product line development. In: Proceedings of the Software and Services Variability Management Workshop: Concepts, Techniques, and Tools. pp. 32–47 (2007)
13. Clements, P.C., Northrop, L.: Software Product Lines : Practices and Patterns. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (2001)
14. Mancuso, M., Buccella, A., Cechich, A., Arias, M., Pol'la, M.: Datasheet modeler: Una herramienta de soporte para el desarrollo de funcionalidades en líneas de productos de software. In: Proceedings of the CACIC'15: XXI Argentine Congress of Computer Science. Junin, Buenos Aires, Argentina (2015)
15. Manikas, K., Hansen, K.M.: Software ecosystems a systematic literature review. *Journal of Systems and Software* 86(5), 1294 – 1306 (2013)
16. Munir, Q., Shahid, M.: Software Product Line: Survey of Tools. Master's thesis, Linkopings University, Linkopings, Sweden (2010)
17. Parent, C., Spaccapietra, S., Zimányi, E.: Spatio-temporal conceptual models: data structures + space + time. In: Proceedings of the GIS '99: 7th ACM International Symposium on Advances in Geographic Information Systems. pp. 26–33. ACM Press, New York, NY, USA (1999)
18. Pohl, K., Böckle, G., Linden, F.J.v.d.: Software Product Line Engineering: Foundations, Principles and Techniques. Springer-Verlag New York, Inc., Secaucus, NJ, USA (2005)
19. Pol'la, M., Arias, M., Buccella, A., Cechich, A.: Un sistema de anotaciones para la especificación de componentes de una línea de productos de software. *Revista Tecnología y Ciencia de la Universidad Tecnológica Nacional* 6(1), 116–122 (2015)
20. Pol'la, M., Buccella, A., Arias, A., Cechich, A.: Sevatax: Service taxonomy selection & validation process for spl development. In: XXXIV International Conference of the Chilean Society of Computer Science (SCCC'15). IEEE Computer Society Press, Santiago, Chile (2015)