

REVISIÓN SISTEMÁTICA PARA EL DISEÑO DE UNA METODOLOGÍA DE REQUERIMIENTOS REFERENCIALES EN DOMINIOS COMPLEJOS.

Moyano Ezequiel UNTDF, Urciuolo Adriana UNTDF, Czelada Alejandro UNTDF, Izarra Emilio UNTDF, Rigoni Brian UNTDF
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Fuegia Basket 251, (9410) Ushuaia. Tierra del Fuego. Tel: ++54-2901-432403

moyano@untdf.edu.ar, urciuolo@untdf.edu.ar, eaczelada@untdf.edu.ar, eizarra@untdf.edu.ar, brian.rigoni1@gmail.com

Resumen

Los sistemas que modelan el comportamiento de la naturaleza se caracterizan por su complejidad, la mayoría desarrollados por expertos del dominio, sin considerar el uso de métodos y herramientas de Ingeniería de software; así el conocimiento adquirido queda solo en ese ámbito. Para el desarrollo de nuevos sistemas el conocimiento debe obtenerse nuevamente. El reuso del conocimiento del dominio resulta esencial.

Para lograr el reuso del conocimiento es necesario contar con métodos de especificación de requerimientos a nivel de dominio (Requerimientos Referenciales), que puedan ser utilizados en futuros sistemas del dominio, reduciendo tiempos y esfuerzos.

A pesar de la importancia de la etapa de requerimientos en el desarrollo de software, existen pocos antecedentes sobre métodos para la formulación de Requerimientos Referenciales como paso previo al Análisis de Dominio.

El presente trabajo presenta una revisión sistemática, en base al protocolo recomendado por Kitchenham, respecto al estado del arte en el uso de una metodología para la especificación de Requerimientos Referenciales en dominios complejos.

Palabras clave: Revisión Sistemática, Requerimientos Referenciales, Ingeniería de Requerimientos, Análisis de Dominio, Reuso,

Contexto

La línea de investigación se desarrolla en el Instituto de Desarrollo Económico e Innovación (IDEI) de la Universidad Nacional de Tierra del Fuego, por parte de un Grupo de docentes-investigadores que lleva adelante proyectos en la temática de Sistemas de Información Ambiental – Hidroinformática desde hace más de 10 años.

El proyecto se presentó a la convocatoria realizada por la UNTDF en septiembre de 2018 y resultó aprobado luego de haber sido sometido a evaluación externa. (Resolución Rectoral 514/2019), para el período de. (Abr/2019-Mar/2021).

1. Introducción

El estudio de los sistemas que modelan la naturaleza adquiere cada vez mayor importancia, dada la función preponderante que los mismos cumplen en la toma de decisiones para el manejo sustentable de los recursos naturales y el medio ambiente. Los modelos ambientales son abstracciones de la naturaleza y de su comportamiento, que permiten la toma de decisión a partir de predicciones[1].

La complejidad de estos sistemas está dada por la necesidad de manejar grandes volúmenes de datos, objetos espacio / temporales, factores ambientales, políticas de gestión, normativas legales, etc[2].

El reuso de conocimiento del dominio en tales situaciones resulta esencial.

Estos sistemas en gran parte fueron desarrollados por expertos del dominio en forma aislada, sin considerar el uso de métodos y herramientas de Ingeniería de software; con lo cual el conocimiento adquirido queda solo en el ámbito en el cual se obtuvo (organismo o institución).

Para el desarrollo de nuevos sistemas (en el mismo dominio), el conocimiento debe obtenerse nuevamente, lo cual involucra un gran esfuerzo. De allí la importancia del reuso de conocimiento en dominios complejos.

Los sistemas de información ambiental, como los de modelado de cuencas, se caracterizan por su complejidad. Rozan una gran cantidad de subdominios diversos de conocimiento, generalmente heterogéneos[3], tanto técnica como semánticamente; se relacionan con el manejo de los datos de distintos componentes interactuantes del ambiente: el suelo, el agua, el aire y las especies existentes.

La mayoría de los fenómenos ambientales se debe modelar considerando cinco dimensiones: su localización en el espacio (latitud, longitud y altitud), el tiempo en que se obtiene y los fenómenos particulares que son analizados.

La ingeniería de dominio tiene como objetivo lograr la reutilización del conocimiento y de las especificaciones de un dominio específico, de manera de poder ser utilizadas en diversas aplicaciones del mismo ámbito.[4]

Las técnicas de análisis de dominio están asociadas a la reutilización, su principal característica es capturar información relacionada con el dominio, y determinar qué y cómo este conocimiento se reutilizará en el desarrollo de futuras aplicaciones[5].

No obstante existen escasas técnicas de especificación de requerimientos a nivel de dominio (Requerimientos Referenciales) que permitan, para sistemas de tal complejidad, obtener un modelo de dominio completo y consistente.

Así como en la ingeniería de software tradicional la primer etapa es la ingeniería de requerimientos (donde existen un gran desarrollo)[6], en la ingeniería de dominio, como paso previo al análisis de dominio, se deben obtener requerimientos referenciales[7].

Para lograr el reuso del conocimiento es necesario contar con métodos de especificación de requerimientos a nivel de dominio (*requerimientos referenciales*), con el propósito de que la información (identificada, capturada y organizada) sea reutilizable en la creación de nuevas aplicaciones de software en ese dominio, en otros subdominios similares, reduciendo tiempos y esfuerzos de considerable magnitud.

A pesar de la importancia de la etapa de requerimientos en el desarrollo de software[8], existen pocos antecedentes sobre métodos para la formulación de requerimientos referenciales como punto de partida para el Análisis de Dominio, especialmente en dominios complejos donde adquiere más trascendencia para el reuso.

El presente trabajo tiene como objetivo fundamental realizar una revisión sistemática [9] respecto al desarrollo de metodologías de especificación de requerimientos inferenciales en dominios complejos, y sobre todo en dominios ambientales.

Los resultados obtenidos se enmarcan dentro de las actividades preliminares relacionadas con el proyecto de investigación “Diseño de una Metodología para Requerimientos Referenciales en Dominios Complejos. Caso De Estudio: Aplicaciones De Modelado De Cuencas” (PIDUNTDF) que se encuentra actualmente en desarrollo y será parte del marco teórico y del estado del arte de la temática a desarrollar.

2. Línea de Investigación

Como parte de la línea de investigación que el equipo de investigadores lleva adelante en temáticas vinculadas a los Sistemas de Información Ambiental – Hidroinformática, se

plantean los siguientes objetivos a alcanzar en el presente proyecto:

Promover el reuso de conocimiento en dominios complejos a través del desarrollo de una metodología de análisis y especificación de Requerimientos Referenciales, que permita reducir tiempo y esfuerzo en la construcción de aplicaciones de un dominio de características complejas.

3. Revisión Sistemática.

Para realizar la revisión sistemática respecto al diseño de metodologías para el uso de requerimientos referenciales en dominios complejos se utilizó como base el protocolo recomendado por Kitchenham[9], definido como una guía para conducir el proceso de revisión sistemática de la literatura en ingeniería de software.

PASO 1: El primer paso de protocolo es establecer la formulación de las preguntas sobre investigación, el objetivo que se persigue es poder determinar de la literatura que se obtenga, la existencia de estudios en el campo o temática a fin del proyecto. No solo aspectos teóricos sino también estudios que definan estrategias sobre casos reales.

Se definen una pregunta central y subpreguntas para profundizar el tema de investigación. Estas preguntas nos ayudarán posteriormente (sobre la muestra final de artículos) determinar el grado de interés respecto al proyecto.

Preg. 1: ¿Existen estudios o evidencias relacionadas a la obtención y/o especificación de requerimientos referenciales en sistemas complejos?

SubPreg. 1.1: “¿Cuáles son los principales aportes de los trabajos relacionados al objetivo del proyecto?”

SubPreg. 1.2: “¿Que característica demuestra el tipo de estudio de la literatura?”

PASO 2: A los efectos de realizar la búsqueda de literatura se deben determinar dos aspectos importantes:

El Conjunto de librerías digitales como fuentes de información y Cadenas de búsqueda.

Para el primer caso se utilizaron[10]:

1. *Google Scholar.*
2. *ACM Digital Library.*
3. *IEEEExplore.*
4. *Springer Link.*
5. *Science Direct.*
6. *SEDICI.*

Para el el segundo caso se definieron 4 criterios de búsqueda (concatenados según el criterio de cada librería)

- A. *Requerimientos + referenciales + sistemas + complejos*
- B. *Ingeniería + Requerimientos + análisis o ingeniería + dominios*
- C. *requirements + referential + systems + complex.*
- D. *engineering + requirements + analysis or engineering + domain.*

Los resultados obtenidos fueron ordenados por relevancia, según el criterio de cada librería digital.

PASO 3: Es importante considerar, más allá de los criterios de búsqueda utilizados, que varios de los estudios seleccionados no aporten evidencia relacionada a las preguntas de investigación propuestas. Para obtener un subconjunto más efectivo se definieron un conjunto de criterios[11][12], de inclusión (CI) y exclusión (CE).

Es importante aclarar los artículos son incluidos si cumple al menos uno de los CI y serán excluidos los que verifiquen al menos de uno de los CE[10].

CI1: Trabajos actuales, comprendidos entre 2005 y 2020.

CI2: Contengan conceptos relevantes del proyecto en su título, abstract y/o palabras claves.

CI3: Demuestren ser de interés y/o aporte del uso de requerimientos referenciales en sistemas complejos.

CE1: Sean revisiones sistemáticas que no generan contribuciones de relevancia.

CE2: Trabajos duplicados o similares en toda su concepción.

CE3: Trabajos que si bien puedan ser sobre requerimientos no están aplicados a sistemas complejos.

CE4: Trabajos no sean artículos científicos publicados.

PASO 4: Por último se debe determinar la calidad de los artículos incluidos en la revisión sistemática, para limitar posibles sesgos y guiar la interpretación de los resultados [13].

A efectos de obtener una valoración de calidad de los artículos seleccionados se utilizó una evaluación cuantitativa basada en cuatro preguntas recomendadas en Kitchenham y Charters[14].

Las preguntas de verificación definidas en la presente revisión sistemática son:

PC1: ¿El artículo contempla adecuadamente los objetivos planteados en el proyecto?

PC2: ¿El artículo presenta una clara relación entre la ingeniería de requerimientos aplicada a dominios complejos o su uso como etapa en el análisis de dominio?

PC3: ¿Se describieron adecuadamente los datos / estudios básicos?

PC4: ¿Los autores evaluaron la validez de los resultados obtenidos y/o el beneficio de especificar los requerimientos del dominio?

4. Resultados.

A continuación se presentan los resultados obtenidos obtenidos luego de aplicar el protocolo de Kitchenham.

En la tabla 1 se muestran los resultados del procedimiento, la segunda columna indica la cantidad total de artículos encontrados, la tercera los artículos seleccionados para su

análisis (las búsquedas se ordenaron por relevancia y se seleccionaron, para cada una de las cuatro, 15 en español y 25 en inglés)

La última columna indica la cantidad de artículos que nos quedan luego de aplicar los criterios de inclusión y exclusión (ver paso 3).

Librería	Artículos Totales	1er Filtro	2do Filtro
IEEE	3.753	75	37
Google Schooll	502.500	80	24
ACM	1.060.300	80	10
Springer	28.969	80	24
Science Direct	210.850	80	12
Sedici	3.056	80	27
Total	1.809.428	475	134

Tabla 1. búsqueda y selección de artículos.

Para obtener los artículos de real calidad para el estudio del estado del arte del proyecto, se aplica una evaluación cuantitativa (ver paso 4 de la revisión sistemática).

Para la valoración se utilizó el criterio del Centro para Revisiones y Difusión (CDR) Base de datos de resúmenes de revisiones de efectos (DARE)[15] de la Universidad de York. donde a cada pregunta se asigna un valor: Y (total) 1, P (parcial) 0,5 y N (Nulo) 0.

Y se define un umbral del 60 % o más, es decir aquellos que sumen una calificación de 2.5 o más. La tabla 2 refleja la cantidad de artículos que fueron considerados.

Valor alcanzado	Cantidad de Artículos
4 Pts.	7
3.5 Pts.	11
3 Pts.	11
2.5 Pts.	10
2 Pts.	54
1.5 Pts.	15
1 Pts.	11
0.5 Pts.	5
0 Pts.	10
Total Considerados	39

Tabla 2. Validación de calidad.

5. Conclusión

El protocolo utilizado para la revisión sistemática de la bibliografía permitió obtener un conjunto concreto de publicaciones a fines a la línea de investigación del proyecto, e identificar las actuales iniciativas que se encuentran en la literatura disponible.

Se puede determinar que no existe una evidencia empírica concreta sobre una metodología de especificación de requerimientos referenciales. Existen pocos antecedentes (a nivel dominio) como paso previo al Análisis de Dominio, sobre todo en sistemas ambientales. Si bien en la última década se ha mostrado una preocupación más acabada sobre esta problemática.

Las actuales investigaciones dirigen su preocupación en la definición de y lenguajes específicos de dominio, elevando el nivel de abstracción mediante la especificación de los conceptos del dominio, sin considerar la necesidad de definir requerimientos referenciales que permitan el reuso del conocimiento.

6. Formación de Recursos Humanos

El Equipo de Trabajo está conformado por docentes investigadores de la UNTDF, licenciados en informática y alumnos de la carrera de Sistemas.

El Proyecto incluye docentes en etapa de formación de postgrado, el Director del proyecto está realizando su tesis de magister en Ingeniería de Software en la UNLP, denominada: "Reuso de Requerimientos Referenciales en Dominios Complejos. Caso de Estudio: Aplicaciones de Modelado de Cuencas." Dirigida por la Mgs. Adriana Urciuolo y co-dirigida por el Dr. Leandro Antonelli de la UNLP.

Participaran Asistentes que realizarán sus primeras en actividades de investigación para desarrollar sus estudios de postgrado en temáticas afines (A. Czelada y E. Izarra). Colaborará el alumno Brian Rigoni, con el objetivo formarlos en los aspectos generales de la

investigación y que le sirva para el desarrollo de su futura tesis de grado.

A través del proyecto se espera consolidar un equipo de trabajo de la UNTDF en la temática.

7. Referencias

1. Moyano E, Urciuolo A, Iturraspe R, Una Metodología para Requerimientos Referenciales en dominios complejos. WICC 2019, San Juan.
2. Urciuolo A. Sistemas de Hidroinformática Oo + Soluciones Evolutivas 2005.
3. Rol Of Information Technology In Environmental Information System, 2003
4. JanisOsis, Erika Asnina, Model-Driven Domain Analysis and Software Development Riga Technical University, Latvia. 2010.
5. Emeline, Regis; Gustavo, Tondello; Análise de Domínio, Universidad Santa Catarina, 2000
6. Olga De Troyer and Erik Janssens, A Feature Modeling Approach for Domain - Specific Requirement Elicitation. Universiteit Brussel.
7. Torres Pérez L., Delgado Dapena M., Entorno de ingeniería de requisitos aplicado para producir software en una universidad, 2014.
8. Rossi G, Antonelli L, Sampaio JC, Oliveros A, Buenas prácticas en la especificación del dominio de una aplicación, WER 2013.
9. Kitchenham, B., Procedures for performing systematic reviews. UK, Keele University, 2004.
10. Pinto N., Acuña C., Tomaselli G., Impacto de las emociones del usuario en la percepción de la calidad del software: Una revisión sistemática, UTN - Facultad Regional Resistencia, 2019.
11. Noruzi, A. The New Generation of Citation Indexes. <http://dx.doi.org/10.1515/LIBR.2005.170> , 2005.
12. Petticrew M, Helen R. Systematic reviews in the social sciences: A practical guide. Blackwell Pub; 2006.
13. Cascaes da Silva, F., Valdivia A., Gutierrez F., Escalas y listas de evaluación de la calidad de estudios científicos. Revista Cubana de Información en Ciencias de la Salud. 2013.
14. Kitchenham B., Charters S., Guidelines for

performing systematic literature reviews in software engineering, Keely University.

15. Centre for Reviews and Dissemination, What are the criteria for the inclusion of reviews? <http://www.york.ac.uk/inst/crd/faq4.htm>, 2007