

DOCUMENTACIÓN DE REQUISITOS EN PROYECTOS DE EXPLOTACIÓN DE INFORMACIÓN

Vegega, C., Mansilla, D., Pollo-Cattaneo, Ma. F.,
Pytel, P., Rodríguez, D., Diez, E., García-Martínez, R.

Grupo de Estudio en Metodologías
de Ingeniería de Software

Facultad Regional Buenos Aires
Universidad Tecnológica Nacional
Medrano 951 (C1179AAQ) Ciudad Autónoma de
Buenos Aires, Argentina. Tel +54 11 4867-7511
fpollo@posgrado.frba.utn.edu.ar

Grupo de Investigación
en Sistemas de Información

Dpto. Desarrollo Productivo y Tecnológico
Universidad Nacional de Lanús
29 de Septiembre 3901 (1826) Remedios de Escalada,
Lanús, Argentina. Tel +54 11 6322-9200 Ext. 194
rgarcia@unla.edu.ar

RESUMEN

El proceso de Educción de Requisitos para Proyectos de Explotación de Información, contempla la realización de un conjunto de plantillas que se utilizan para documentar los conceptos educidos durante las primeras fases del proyecto. Por otro lado, la Ingeniería del Conocimiento (INCO) se encarga de la construcción de Sistemas Basados en Conocimientos, intentando representar el conocimiento y comportamiento inteligente de los seres humanos en un sistema artificial. Los Ingenieros del Conocimiento deben obtener el conocimiento de expertos humanos en una determinada área o dominio y luego modelizar estos conocimientos, de manera tal de que sean comprensibles por un sistema software. Para ello, se utilizan técnicas y herramientas provistas por la INCO. En este contexto, este proyecto busca identificar las técnicas de modelado de conocimientos, propias de la Ingeniería del Conocimiento, que puedan ser utilizadas para la documentación de requisitos en proyectos de explotación de información.

Palabras clave: Educción de Requisitos, Documentación de Requisitos, Ingeniería del Conocimiento, Explotación de Información.

CONTEXTO

Este proyecto de investigación articula líneas de trabajo de los proyectos

“Metodología para la Especificación de Requisitos en Proyectos de Explotación de Información” (Universidad Tecnológica Nacional) e “Ingeniería de Proyectos de Explotación de Información para PyMES” (Universidad Nacional de Lanús), los cuales se desarrollan en el marco de la cooperación existente entre el Grupo de Estudio de Metodologías de Ingeniería de Software (GEMIS) de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional (UTN-FRBA) y el Grupo de Investigación en Sistemas de Información (GISI) del Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús (DDPyT-UNLa).

INTRODUCCION

En [Pollo-Cattaneo et al. 2009] se señala la necesidad de adaptar el proceso tradicional de especificación de requerimientos de sistemas software para Proyectos de Explotación de Información (EdI) y se propone un proceso de educación de requisitos, mediante el cual se completan un conjunto de plantillas definidas en [Britos et al. 2008] a realizar, para documentar los conceptos que se educen, durante la fase de comprensión del negocio y del dominio del proyecto.

Cada plantilla se asocia a un concepto educido y tiene una descripción detallada del mismo, permitiendo la evolución del concepto a lo largo de todo el proceso de

educación. Luego, a partir de la información contenida en las plantillas, el Ingeniero de EdI puede realizar las actividades del proyecto como, entre otras, determinar las características de los recursos humanos necesarios, identificar los repositorios de datos, determinar el modelo y seleccionar la herramienta a utilizar.

A continuación, se enuncian los conceptos a ser educidos y se describe brevemente cada uno [Britos et al. 2008]:

- *Definiciones, Acrónimos y Abreviaturas:* Se identifican las definiciones, acrónimos y abreviaturas, de forma de establecer el vocabulario a utilizar por todas las personas relacionadas al proyecto de explotación de información.
- *Objetivos del Proyecto:* Se identifican los objetivos del proyecto, según las necesidades del cliente.
- *Criterios de Éxito del Proyecto:* Se identifican los logros esperados para convertir el proyecto en un éxito.
- *Expectativas del Proyecto:* Se identifica lo que se espera lograr mediante la ejecución del proyecto.
- *Supuestos del Proyecto:* Se identifican los supuestos que deben ser tenidos en cuenta respecto de los objetivos del proyecto.
- *Restricciones del Proyecto:* Se identifican los límites del proyecto.
- *Riesgos del Proyecto:* Se identifican los riesgos que se asumen dentro del proyecto.
- *Plan de Contingencias:* Se definen los planes de contingencia que se aplicarán sobre los riesgos que se identificaron.
- *Recursos Humanos Involucrados:* Se identifican los recursos humanos involucrados en el proyecto, junto con su respectivo rol.
- *Requisitos:* Los objetivos del proyecto, definidos con anterioridad, se descomponen en requisitos.
- *Fuente de Datos:* Se establecen las fuentes de datos para cada uno de los requisitos, así como la ubicación y acceso a las mismas.
- *Supuestos de los Requisitos:* Se identifican los supuestos a tener en cuenta por cada requisito.
- *Restricciones de los Requisitos:* Se identifican los límites de cada requisito.
- *Atributos relacionados con los Requisitos:* Se establecen los atributos que se utilizarán para cumplir cada requisito. Los atributos son las características de los elementos del dominio identificados.
- *Riesgos de los Requisitos:* Se identifican los riesgos que involucran cada requisito.
- *Plan de Contingencia de los Requisitos:* Se definen los planes de contingencia que se vinculan a los riesgos identificados por cada requisito.
- *Evaluación de Herramientas:* Se evalúan las herramientas disponibles para seleccionar la más conveniente para el cumplimiento de los objetivos del proyecto.

Por otra parte, en [Pytel et al. 2011] se indica que la Ingeniería del Conocimiento (INCO) utiliza técnicas específicas que permiten la representación del conocimiento. Dentro de las técnicas de representación del conocimiento, se destacan las siguientes [García-Martínez, 1994; García-Martínez & Britos 2004]:

- *Las Tripletas Concepto-Atributo-Valor* sirven para identificar los atributos que describen los conceptos que se utilizan en el desarrollo del proyecto, especificando los valores que estos atributos poseen.

- Las *Redes Semánticas* se utilizan para expresar relaciones entre los conceptos del dominio. En una red semántica, la información se representa en un grafo orientado, formado por nodos que representan a los conceptos y por arcos unidireccionales que representan las relaciones entre ellos.
- Los *Marcos* son una técnica de representación que organiza el conocimiento del dominio en árboles jerárquicos, en forma similar a la representación de información que se realiza en la programación orientada a objetos. Se definen marcos (similares a las clases de la programación orientada a objetos) que representan objetos con características similares, a partir de los cuales, se generan instancias particulares de los mismos. Cada marco posee un conjunto de propiedades y facetas para representar los valores de cada propiedad. Esta técnica permite representar relaciones del dominio, a través de las relaciones entre los marcos.
- Los *Guiones* se utilizan para representar secuencias estereotipadas de sucesos. En cada guión se define un conjunto de acciones (o escenas) que realizarán los actores y objetos involucrados, un conjunto de condiciones que deben cumplirse para que el guión sea ejecutado y un conjunto de resultados que se realizarán una vez que las acciones enunciadas en el guión se hayan finalizado.
- Las *Tablas de Decisión* permiten documentar, en forma ordenada, un conjunto de decisiones. Estas decisiones consisten en la selección de acciones a ejecutar, si las condiciones que rodean esa decisión se cumplen.
- Las *Seudoreglas* son una técnica que representa el conocimiento a través de la definición de un conjunto de condiciones y sus consecuentes acciones, que de dichas condiciones se desprenden. Se representan de la forma: “SI (condiciones) ENTONCES (acciones)”. Están relacionadas con las tablas de decisión, ya que toda regla definida en la tabla de decisión, debe encontrarse definida como seudoregla.
- Los *Hechos Difusos* permiten la representación de conocimiento impreciso o ambiguo, a través de la definición de funciones de pertenencia que asignan un valor numérico (que varía entre 0 y 1). Este valor representa el grado de pertenencia de un elemento en un conjunto.

En este contexto, se busca reducir la complejidad del llenado de las plantillas, empleadas para la documentación de requisitos en proyectos de EdI, a través de la utilización de técnicas de representación de conocimiento provistas por la Ingeniería del Conocimiento.

OBJETIVOS DE LA INVESTIGACION

Se han identificado los conceptos del dominio que deben ser educidos dentro de un proyecto de EdI y definido un proceso que abarca la realización de plantillas para documentar éstos conceptos. Teniendo en cuenta este proceso de educación ya propuesto, este proyecto tiene como objetivo identificar las técnicas de representación de conocimientos de la Ingeniería del Conocimiento aplicables a la documentación de requisitos en proyectos de explotación de información, dentro del ciclo de vida de este tipo de proyectos.

FUNDAMENTACION E MPORTANCIA

Dentro del dominio de la EdI, durante el proceso de Educación de Requisitos, es necesaria la documentación y formalización de los requisitos a través del llenado de

plantillas que permitan tener una descripción detallada de cada requisito educido.

El problema que se ha identificado en este proyecto consiste en reducir la complejidad del llenado de las plantillas, estructurando un proceso que utilice las técnicas de representación de conocimientos de la Ingeniería del Conocimiento. De esta manera, se busca ofrecerles a los Ingenieros de EdI un conjunto de técnicas y herramientas que permitan reducir la complejidad del proceso de documentación.

METODOLOGÍA DE TRABAJO

En el desarrollo del presente proyecto de investigación, se seguirá un enfoque metodológico que consiste en el estudio y análisis de casos, de técnicas de análisis y de síntesis de comparaciones, realizándose las siguientes tareas:

- [a] Se realizará un relevamiento de las técnicas de representación de conocimientos, utilizadas en el marco de la Ingeniería del Conocimiento (INCO).
- [b] Se analizará la forma en la cual los requisitos educidos pueden ser representados en los formalismos propios de la INCO. A tal efecto se realizarán las siguientes actividades:
 - [b.1] Estudio de la utilización de las técnicas de representación de conocimientos provistas por la INCO.
 - [b.2] Identificación de casos de estudio para la educación y representación de requisitos.
 - [b.3] Formalización de los casos de estudio identificados, utilizando las técnicas previamente estudiadas de representación de conocimientos.
 - [b.4] Análisis de las ventajas y desventajas de documentar los requisitos educidos mediante técnicas de la INCO.

[c] Se formulará un modelo, basado en el análisis realizado, para la documentación de requisitos que se centrará en las técnicas de representación de conocimientos provistas por la INCO.

[d] Se identificarán las aportaciones del proyecto y se dará difusión mediante comunicaciones a congresos generales o del área específica.

FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo se encuentra formado por dos investigadores formados, tres investigadores en formación y un supervisor científico. En el marco de este proyecto se está desarrollando una tesis doctoral.

REFERENCIAS

- F. Pollo-Cattaneo, P. Britos, P. Pesado, R. García-Martínez (2009). Metodología para Especificación de Requisitos en Proyectos de Explotación de Información. Proceedings XI Workshop de Investigadores en Ciencias de la Computación. Pág. 467-469. ISBN 978-950-605-570-7.
- Britos, P., Dieste, O., García-Martínez, R. (2008). Requirements Elicitation in Data Mining for Business Intelligence Projects. IFIP Series, 274: 139-150.
- García-Martínez, R. y Britos, P. (2004). Ingeniería de Sistemas Expertos. Editorial Nueva Librería.
- García-Martínez, R. (1994). Adquisición de Conocimiento. En Abecasis, S. y Heras, C. Metodología de la Investigación. Nueva Librería. ISBN 950-9088-65-x.
- Pytel, P., Uhalde, C., Ramón, H. Castello, H., Tomasello, M., Pollo-Cattaneo, M., Britos, P., García-Martínez, R. (2011). Ingeniería de Requisitos Basada en Técnicas de Ingeniería del Conocimiento. Proceedings XIII Workshop de Investigadores en Ciencias de la Computación. Pág. 426-429. ISBN 978-950-673-892-1.
- Gomez, A., Juristo, N., Montes, C. y Pazos, J. (1997). Ingeniería del Conocimiento. Editorial R. Areces. Madrid.