

PROPUESTA DE METODOLOGÍA PARA AUTOMATIZACIÓN DEL PROCESO DE ELICITACIÓN DE REQUISITOS

Noemí Medina, Pablo Pytel, Ma. Florencia Pollo-Cattaneo
Grupo de Estudio en Metodologías de Ingeniería de Software (GEMIS)
Programa de Maestría en Ingeniería en Sistemas de Información.
Universidad Tecnológica Nacional. Facultad Regional Buenos Aires
Medrano 951 (C1179AAQ) Ciudad Autónoma de Argentina. Buenos Aires Tel +54 11 4867-7511
{noemi.medina, ppytel, flo.pollo}@gmail.com

Resumen

Durante el ciclo de vida del desarrollo de un sistema software suele ser necesario analizar una gran cantidad de documentación escrita en lenguaje natural (es decir, en el lenguaje propio de los usuarios del sistema de información) a fin de descubrir y especificar los requisitos del sistema. Esta etapa del proceso de elicitación consume gran cantidad de tiempo y es propensa a muchos errores cuando se realiza manualmente. Esto lleva a pensar en formas de automatizar la extracción de conocimiento a partir de dicha documentación disponible. La Ingeniería de Software tradicionalmente ha apoyado las iniciativas de modelado con las denominadas herramientas Computer-Aided Software Engineering (CASE), además de la generación parcial de código fuente. Varios autores han propuesto la utilización de herramientas y metodologías que automaticen algunas etapas del proceso de elicitación de requisitos. Sin embargo, todavía no se ha logrado definir una metodología para encarar todo el proceso de elicitación en forma coherente y completa. En este contexto, el presente trabajo de investigación tiene como objetivo desarrollar una metodología que asista al ingeniero de requisitos en el proceso de elicitación utilizando herramientas que automaticen la tarea de procesamiento y análisis de documentación preexistente.

Palabras clave: Ingeniería de requisitos, Elicitación de Requisitos, Herramientas, Automatización.

Contexto

Esta investigación se desarrolla dentro de las líneas de trabajo del Grupo de Estudio de Metodologías para Ingeniería en Software y Sistemas de Información (GEMIS), el cual se encuentra conformado por un equipo de docentes y alumnos dentro del ámbito de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional. Este grupo busca la sistematización de cuerpos de conocimientos y promoción sobre el campo de la Ingeniería en Sistemas de Información y la Ingeniería en Software, sus aplicaciones y abordajes metodológicos en todo tipo de escenarios (convencionales y no convencionales).

Introducción

Tal como sucede en otras ingenierías, el proceso de construcción de un sistema software en el marco de la Ingeniería de Software comienza con la especificación del producto que pretende ser el resultado final del proceso [Pressman, 2005].

Sin embargo, los problemas que a menudo tienen que resolver los ingenieros de software son extremadamente complejos [Sommerville, 2005], esto genera que sea difícil establecer exactamente lo que el sistema debe hacer. Las descripciones de los servicios y restricciones se consideran los requisitos para el sistema, y al proceso de descubrir, analizar, documentar y verificar estos servicios y restricciones se lo

denomina Ingeniería de Requisitos (IR). De esta manera, la IR se ocupa de la descripción completa del comportamiento a implementar en el software, incluyendo la descripción de las funcionalidades que el sistema software deberá realizar, el detalle de las interacciones que los usuarios tendrán con el software y los requisitos no funcionales que imponen restricciones al diseño o funcionamiento del sistema software [Pohl, 1994]. Esto se formaliza luego en un documento de Especificación de Requisitos Software [IEEE, 1998] que debe ser completo y aceptado por los interesados.

Normalmente, en los proyectos de desarrollo de software se dispone de una gran cantidad de documentos no estructurados (por ejemplo entrevistas de usuarios, documentación de sistemas preexistentes, manuales de procedimientos, formularios, entre otros) provenientes de diferentes fuentes organizacionales que son necesarios analizar y transformar en requisitos estructurados [Meth et al., 2013]. Debido a que se sabe que esta etapa del proceso de elicitación consume mucho tiempo y es propenso a errores cuando se realiza manualmente, esto lleva a pensar en formas de automatizar la extracción de conocimiento a partir de toda esa información escrita existente.

La Ingeniería del Software tradicionalmente ha apoyado las iniciativas de modelado con las denominadas herramientas Computer-Aided Software Engineering (CASE), desarrolladas desde mediados de la década de los setenta y que en la actualidad permiten el trazado y edición de diferentes tipos de diagramas (incluida la gama de diagramas UML), además de la generación parcial de código fuente [Zapata, 2007]. Sin embargo, estas herramientas no brindan apoyo a los procesos previos a la realización de los modelos, y esta labor se ha dejado completamente en manos de los analistas, quienes deben interpretar el dominio del problema para plasmarlo posteriormente en los modelos.

Aunque algunos autores entre los que se destacan [Sawyer et. al., 2002; Haak, 2005; Sampaio et al., 2007] han propuesto la utilización de herramientas y metodologías que automaticen algunas etapas del proceso de

elicitación de requisitos; todavía no se ha logrado definir una metodología robusta que guíe al ingeniero de requisitos para encarar todo el proceso en forma coherente y completa.

Asimismo, al momento de seleccionar la herramienta adecuada según las necesidades del proyecto, se presenta un problema adicional debido a que las propuestas se encuentran dispersas en diferentes artículos, en general están escritas en diferentes idiomas, y hasta con variados estilos de redacción, lo cual dificulta la comparación de los beneficios y defectos de cada propuesta.

Líneas de Investigación, Desarrollo e Innovación

Considerando el contexto identificado en la sección anterior, este trabajo tiene como objetivo proponer una metodología que englobe los enfoques existentes para automatizar el proceso de elicitación de requisitos. De esta manera, se espera brindar con un mecanismo que permita al ingeniero de requisitos seleccionar el mejor enfoque y herramienta de acuerdo a las características del proyecto y de la organización.

Para llevar a cabo este trabajo, se distinguen los siguientes enfoques no tradicionales que pueden ser aplicables a la automatización de análisis de textos:

- *Resumen de textos:*

Una de las formas de análisis automático de textos es la técnica de resumen. Existen dos estrategias para su confección: la “Abstracción” y la “Extracción” [Demitrio et al., 2005]. Borko y Bernier [1975] describen las abstracciones indicando dos tipos, “Abstracciones Indicativas” y “Abstracciones Informativas” [Mani, I. 2001]. Las “Abstracciones Indicativas” proveen una función de referencia para seleccionar documentos a ser leídos con mayor profundidad. Se construyen para ayudar al usuario a decidir cuándo es conveniente leer la información del documento fuente o no. En cambio, las “Abstracciones Informativas” cubren toda la

información sobresaliente en el documento fuente al mismo nivel de detalle.

Por otra parte, la estrategia de extracción consiste en seleccionar partes del texto para armar el resumen, la consecuencia directa de esto es que el texto resultante puede no ser consistente o completamente comprensible, en particular si se realiza en forma automática, pero aun así se utiliza porque resulta más barata su implementación, ya que no requiere que se infieran contenidos, sino que se extraen directamente del texto de entrada [Demitrio et al., 2005].

o *Minería de Textos:*

La minería de texto se parece mucho a la minería de pepitas de oro en el sentido de que se trata de encontrar información que no se conocía previamente [Hearst, 1999]. Hearst la define como la aplicación de procesos sobre textos para obtener datos no conocidos, diferenciando éstos de los procesos de búsqueda por Recuperación de Información (Information Retrieval). En este sentido, la minería de texto puede ser considerada como el área de investigación más reciente del procesamiento automático de textos y se define como el proceso automático de descubrimiento de patrones interesantes en una colección de textos. Estos patrones no deben de existir explícitamente en ningún texto de la colección, y deben de surgir de relacionar el contenido de varios de ellos [Kodratoff, 1999].

o *Procesamiento de Lenguaje Natural:*

Por Procesamiento de Lenguaje Natural (PLN, denominado también NLP por sus siglas en inglés) se entiende la habilidad de la máquina para procesar la información comunicada, no simplemente las letras o los sonidos del lenguaje [Gelbukh, 2007].

Metodología de Trabajo

Las tareas a realizar para proponer la metodología serán las siguientes:

- a) Se realizará una investigación documental exploratoria para identificar los enfoques y

herramientas existentes para la automatización de la elicitación de requisitos.

- b) Se compararán dichos enfoques mediante un análisis teniendo en cuenta diferentes tipos de proyecto y de organización.
- c) Se seleccionarán los principales enfoques que se han procurado realizar para agilizar la comprensión del dominio del negocio.
- d) Se especificará una metodología para guiar el proceso de elicitación de requisitos usando estos enfoques.
- e) Se validará el comportamiento de la metodología propuesta mediante casos de estudio.
- f) Se realizará el análisis los resultados obtenidos.

Resultados y Objetivos

Como resultado de este proyecto, se espera desarrollar una metodología que asista al ingeniero de requisitos en el proceso de elicitación utilizando herramientas que automaticen la tarea de procesamiento y análisis de documentación preexistente.

Además, se proponen los siguientes objetivos específicos:

- Determinar los principales beneficios y limitaciones de los enfoques existentes para la automatización de la elicitación de requisitos.
- Proponer una guía para determinar el tipo de herramienta más convenientes a ser utilizados en la automatización de tareas según las características del proyecto y la organización.
- Definir un conjunto de tareas que permita al ingeniero de requisitos aplicar el enfoque y la herramienta correspondiente en forma exitosa de acuerdo a las características del proyecto y la organización.

Formación de Recursos Humanos

Este proyecto busca tanto la obtención de nuevos conocimientos como la motivación de

los implicados para que asciendan dentro del escalafón de la carrera de investigadores.

Los recursos humanos deben poseer una firme vocación de trabajar en el área de la ingeniería en software y la ingeniería de requerimientos de manera de canalizar y proveer una base sustentable de aporte al proyecto. De esta manera, se logra un doble beneficio, el proyecto obtiene e incorpora el conocimiento tecnológico de los recursos humanos en el área de la especialidad, a la vez que plantea un esquema de formación de especialistas de punta en el proceso de gestión.

El grupo de trabajo se encuentra formado por dos investigadores formados y un investigador en formación. Asimismo, este trabajo de investigación se realiza en el marco de un Trabajo de Final de Especialidad y de una Tesis de Magister en Ingeniería en Sistemas de Información.

oriented software development III (pp. 4-39). Springer Berlin Heidelberg.

Sawyer, P., Rayson, P., & Garside, R. (2002). REVERE: support for requirements synthesis from documents. *Information Systems Frontiers*, 4(3), 343-353.

Sommerville, I. (2005). *Ingeniería del software*. Pearson Educación.

Zapata, C. M. (2007). Un ambiente para la obtención automática de diagramas UML a partir de un lenguaje controlado. *Dyna*, 153, 225.

Referencias

Borko, H., & Bernier, C. L. (1975). *Abstracting Concepts and Methods*.

Demitrio, D. A. (2005). *Framework para elicitación automática de conocimiento*. Doctoral dissertation, Facultad de Informática. <http://hdl.handle.net/10915/4087>

Gelbukh, A. (2007). *Procesamiento de lenguaje natural y sus aplicaciones*. Artículo invitado. *Komputer Sapiens*, ISSN, 691, 6-11.

Haak, B. (2005) *Aspects Extractor: Identificación de Aspectos en la Ingeniería de Requerimientos*. Facultad de Ciencias Exactas, UNICEN.

IEEE Computer Society. *Software Engineering Standards Committee, & IEEE-SA Standards Board*. (1998). *IEEE Recommended Practice for Software Requirements Specifications*. Institute of Electrical and Electronics Engineers.

Kodratoff, Y. (1999). *Knowledge discovery in texts: a definition, and applications*. In *Foundations of Intelligent Systems* (pp. 16-29). Springer Berlin Heidelberg.

Mani, I. (2001). *Automatic summarization* (Vol. 3). John Benjamins Publishing.

Meth, H., Brhel, M., & Maedche, A. (2013). *The State of the Art in Automated Requirements Elicitation*, *Information and Software Technology*.

Pressman, R. S. (2005). *Ingeniería de Software*, Sexta Edición, Ed.

Pohl, K. (1994). *The three dimensions of requirements engineering: a framework and its applications*. *Information systems*, 19(3), 243-258.

Sampaio, A., Rashid, A., Chitchyan, R., & Rayson, P. (2007). *EA-Miner: towards automation in aspect-oriented requirements engineering*. In *Transactions on aspect-*