

Estado actual de la investigación en requisitos no funcionales¹

Current State of Research on Non-Functional Requirements²

Estado atual da pesquisa em requisitos não funcionais³

Código SICI: 0123-2126(201206)16:1<225:EAIRNF>2.3.TX;2-Q

Édgar Serna-Montoya⁴

¹ Fecha de recepción: 7 de mayo de 2011. Fecha de aceptación: 18 de noviembre de 2011. Este artículo se deriva de un proyecto de investigación denominado *Estructuración de una metodología genérica para la realización de pruebas de caja negra en los sistemas de información*, desarrollado por el grupo de investigación Sistemas de Información y Sociedad del Conocimiento (Sisco), de la Fundación Universitaria Luis Amigó, Medellín, Colombia.

² Submitted on May 7, 2011. Accepted on November 18, 2011. This article is the result of the research project *Development of a generic methodology for the execution of black box tests in information systems*. Developed by the research group Sistemas de Información y Sociedad del Conocimiento (Sisco), Fundación Universitaria Luis Amigó, Medellín, Colombia.

³ Data de recepção: 7 de maio de 2011. Data de aceitação: 18 de novembro de 2011. Este artigo deriva de um projeto de pesquisa denominado *Estruturação de uma metodologia genérica para a realização de testes de caixa preta nos sistemas de informação*, desenvolvido pelo grupo de pesquisa Sistemas de Informação e Sociedade do Conhecimento (Sisco), da Fundação Universitária Luis Amigó, Medellín, Colômbia.

⁴ Ingeniero de sistemas, Corporación Universitaria Remington, Medellín, Colombia. Ingeniero de sistemas, Universidad Nacional de Colombia, sede Medellín, Colombia. Profesor, Fundación Universitaria Luis Amigó, Medellín, Colombia. Correo electrónico: funlam@funlam.edu.co.

Resumen

Generalmente se reconoce que los requisitos no funcionales son una parte importante y difícil del proceso de la ingeniería de requisitos y que desempeñan un papel fundamental en el desarrollo de productos de *software*. El objetivo de este trabajo es identificar las investigaciones actuales alrededor de la administración de estos requisitos, para lo cual se realizó una revisión sistemática a la literatura y se identificaron los estudios empíricos disponibles acerca de la temática. Se hizo una búsqueda manual y en bases de datos y se identificaron 1560 artículos, de los cuales dieciocho resultaron ser de investigación empírica de alta calidad relevantes para las preguntas de esta revisión. También se investigó acerca del concepto que se tiene actualmente de los beneficios y las limitaciones de los métodos para administrar estos requisitos. Se presenta el estado de la investigación en cinco áreas: elicitación, dependencias, métricas, estimación de costos y priorización.

Palabras clave

Ingeniería de requisitos, revisión basada en la evidencia, ingeniería de *software*.

Abstract

It is generally recognized that non-functional requirements are an important and difficult part of the engineering of requirements; and that they play a key role in the development of software products. The objective of this paper is to identify current research on the administration of these requirements. To attain this goal, a systematic review of the literature was carried out. This study identified available empirical studies on the subject. The search was carried out manually and in databases. It identified 1560 articles; 18 of them were empirical high-quality research studies that are relevant for the research question of this review. There was also research on the current concept concerning the benefits and limitations of the methods used in managing these requirements. The review presents research status in five areas: elicitation, dependencies, metrics, cost estimation and prioritization.

Keywords

Requirements engineering, evidence-based review, software engineering.

Resumo

Geralmente reconhece-se que os requisitos não funcionais são uma parte importante e difícil do processo da engenharia de requisitos e que desempenham um papel fundamental no desenvolvimento de produtos de *software*. O objetivo deste trabalho é identificar as pesquisas atuais em torno da administração destes requisitos, para isso realizou-se uma revisão sistemática da literatura e identificaram-se os estudos empíricos disponíveis sobre a temática. Foi feita uma busca manual nos bancos de dados e identificaram-se 1560 artigos, dos quais dezoito resultaram ser de pesquisa empírica de alta qualidade relevantes para as perguntas desta revisão. Também se pesquisou sobre o conceito que existe atualmente dos benefícios e as limitações dos métodos para administrar estes requisitos. Apresenta-se o estado da pesquisa em cinco áreas: elicitacão, dependências, métricas, estimativa de custos e priorização.

Palavras chave

Engenharia de requisitos, revisão baseada na evidência, engenharia de *software*.

Introducción

La complejidad de los sistemas de *software* está determinada por la funcionalidad y por los aspectos de calidad: rendimiento, fiabilidad, exactitud, seguridad y usabilidad (Doerr et ál., 2005). Estos aspectos se conocen como requisitos no funcionales del *software*, y comúnmente se acepta que su manejo y equilibrio es una parte importante y difícil del proceso de la ingeniería de requisitos (Berntsson et ál., 2008) y que desempeñan un papel fundamental en el desarrollo del *software* (Chung et ál., 2000).

Una de las características de estos requisitos es la especificación de ciertos niveles de calidad y, por consiguiente, en muchos casos, es posible cuantificarlos (Boehm e In, 1996). Esto es importante no solo para su comprensión (Berntsson et ál., 2008), sino también para su planificación (Cleland-Huang et ál., 2005). Tratarlos ineficazmente o no tratarlos puede dar lugar a un producto más costoso y, posiblemente, que se demore su salida al mercado (Cysneiros y Leite, 2004) o, en el peor de los casos, a errores en el desarrollo del producto (Breitman et ál., 1999; Finkelstein y Dowell, 1996).

Varios estudios (Brooks, 1987; Cysneiros y Leite, 1999) han demostrado que elicitación de estos requisitos es costoso y difícil de manejar, y de acuerdo con Chung et ál. (2000), a menudo, son mal comprendidos, en comparación con aspectos menos críticos del desarrollo del *software*. Generalmente se reconoce que las decisiones acerca de cuáles criterios de calidad deben precisarse en un producto tienen grandes efectos en su desarrollo y en la elección de la arquitectura. Esto significa que el área de los requisitos no funcionales es importante para comprender con más detalle qué dependencia existe entre su calidad y otros componentes del sistema que se va a desarrollar.

No se encuentran esfuerzos sistemáticos significativos en la literatura para identificar, sintetizar y reportar los estudios acerca de los requisitos no funcionales. Para cubrir este vacío, esta revisión sistemática de la literatura tiene por objeto recopilar y comparar la evidencia empírica existente en relación con los requisitos no funcionales, con la idea de ofrecerles a los investigadores un

camino para futuras investigaciones, y a los profesionales, una guía para adoptar tecnologías adecuadas. En esta revisión solo se consideran los artículos centrados en los métodos y la administración de estos requisitos, es decir, los trabajos sobre priorización de requisitos, en general, están por fuera del alcance de esta revisión.

El artículo se estructura de la siguiente forma: en la primera sección se describe el método de revisión utilizado; en la segunda se describen los resultados obtenidos; en la tercera sección se discuten los beneficios y las limitaciones de los métodos de elicitación de los requisitos no funcionales; en la cuarta sección se resumen las principales conclusiones, y en la última se hacen recomendaciones para investigaciones futuras.

1. Método de investigación

Esta investigación se llevó a cabo, como una revisión sistemática a la literatura (*Systematic Literature Review* [SLR]) (Kitchenham, 2010), una forma metódica de identificar, evaluar y analizar la investigación disponible relevante para una pregunta de investigación particular; es decir, se presentan los elementos de un estudio de mapeo sistemático. Además, incluye actividades como planeación de la revisión, identificación de las preguntas de investigación, estrategia de búsqueda y registro, selección de los estudios, evaluación de la calidad y extracción y síntesis de los datos.

Se identificaron cinco áreas importantes para la administración de los requisitos no funcionales: elicitación, métricas, dependencias, estimación y priorización de costos. Debido a que la elicitación está incluida desde el primer paso, identificar los requisitos no funcionales es importante para cuantificar, es decir, para hacer medibles los criterios de calidad-métricas. Para identificar las interdependencias (dependencias) entre requisitos no funcionales es importante antes estimar el costo (estimación de costos).

La implementación de un determinado requisito no funcional desde otro puede tener un efecto negativo o positivo en el costo de la implementación de un segundo requisito no funcional. Por último, se priorizan los requisitos no funcionales. El objetivo de esta RSL es responder a las siguientes preguntas de investigación:

- PI1: ¿qué se conoce actualmente de las investigaciones empíricas acerca de los requisitos no funcionales en relación con la elicitación, la priorización, la estimación de costos, las dependencias y las métricas?
- PI2: ¿qué métodos de investigación empírica se han utilizado para evaluar los requisitos no funcionales?

1.1. Estrategia de búsqueda y registro

La estrategia de búsqueda incluyó la revisión en bases de datos electrónicas y búsquedas manuales en actas de congresos. Se utilizaron las siguientes bases de datos electrónicas: ACM Digital Library, Engineering Village, IEEE Xplore y Wiley Inter Science Journal Finder.

La cadena de búsqueda se aplicó a dos bases de datos electrónicas más: ScienceDirect-Elsevier y SpringerLink. Sin embargo, la cadena era demasiado larga y compleja para ser utilizada en estas bases de datos, por lo que se utilizó un subconjunto de ella en ambas bases de datos y se compararon los resultados con las cuatro bases de datos iniciales. Todos los artículos encontrados del subgrupo en estas dos últimas también se encontraron en las bases de datos primarias.

La International Requirements Engineering Conference, el principal evento en la ingeniería de requisitos en el mundo, está indexada en el Compendex y en la bases de datos Inspec; por lo tanto, no se realizaron búsquedas manuales en ella. Sin embargo, se realizaron búsquedas manuales en los siguientes informes de conferencias internacionales: International Workshop on Software Product Management (IWSPM), Measuring Requirements for Project and Product Success (MeReP) y Requirements Engineering Foundation for Software Quality (REFSQ).

Se realizó una búsqueda manual en la conferencia REFSQ, debido a que es la segunda conferencia más grande dirigida a la ingeniería de requisitos; en IWSPM, debido a que su enfoque es la gestión de productos de *software*, que incluye la priorización de requisitos; además, se buscó en MeReP, porque su enfoque son las métricas y mediciones. El proceso de selección de los estudios se realizó en cuatro fases. En la fase 1 se realizó la búsqueda en las bases de datos utilizando los términos de búsqueda que se enumeran en la tabla 1.

Tabla 1. Términos de búsqueda identificados

Categoría	Palabras buscadas	Conector
C1: QR	Non functional requirements Nonfunctional requirements Non-functional requirements Non-functional software requirements Qualities Quality attributes Quality characteristics Quality factors Quality requirements	OR
C2: <i>software</i>	Software	

Continúa

Categoría	Palabras buscadas	Conector
C3: elicitación	Elicitation Requirements gathering Requirements acquisition	OR
C4: dependencia	Change impact Conflict Dependence Interdependence Inter-dependencies Relationships Traceability Trade off Trade offs Tradeoff Trade-off Tradeoffs Trade-offs	OR
C5: métricas	Metrics Measurement	OR
C6: costos	Software development effort Cost estimation	OR
C7: priorización	Prioritisation Prioritise Prioritising Prioritization Prioritize Prioritizing	OR

Fuente: presentación propia de los autores.

Chung et ál. (2000) definen alrededor de 160 términos diferentes para requisitos no funcionales. Además de varios estándares que también los definen. Para incluir todos esos términos sería necesario crear una cadena de búsqueda muy larga y complicada, y aun así no se garantizaría que todos fueran cubiertos, por lo que se excluyeron términos específicos acerca de requisitos no funcionales, como requisitos de desempeño y usabilidad. Por lo tanto, la categoría C1 quedó compuesta por diferentes términos generales, las categorías C3 y C6 se basaron en revisiones sistemáticas previas (Finkelstein y Dowell, 1987; Jørgensen y Shepperd, 2007, respectivamente). La cadena de búsqueda final, que arrojó una población de 2647 artículos, de los cuales 1560 eran relevantes para la investigación, se construyó de la siguiente manera: C1 AND C2 AND (C3 OR C4 OR C5 OR C6 OR C7).

1.2. Selección de los estudios

Los 1560 estudios relevantes de la fase 1 se introdujeron en una herramienta de base de datos de referencia. Solo se incluyeron trabajos escritos en inglés. En la fase 2 se clasificaron todos los trabajos según el primer autor, con el objetivo de determinar su pertinencia para la RSL. En esta fase se excluyeron los artículos con títulos que indicaban claramente que estaban por fuera del alcance de la RSL.

Para minimizar la amenaza de excluir artículos relevantes, de los excluidos inicialmente se seleccionaron al azar dos conjuntos muestrales, uno organizado por el primer autor (con diferentes artículos) y otro por el segundo y el tercer autores. Dos artículos fueron objeto de debate; sin embargo, ninguno se añadió a los previamente seleccionados. Después de la fase 2, se identificaron 727 artículos como estudios relevantes. Durante la siguiente fase se revisaron los resúmenes de todos los trabajos excluidos por el primer autor, y el 20% de estos por el segundo y el tercero. Cuatro artículos fueron objeto de debate en esta fase, de los cuales se añadió uno a los trabajos seleccionados. Al final de la fase 3 quedaron 229 artículos para el proceso de selección de la fase 4.

1.3. Evaluación de la calidad

Se utilizaron los siguientes criterios de selección (Dybå y Dingsyr, 2008) para garantizar la calidad de los trabajos y para excluir aquellos trabajos de investigación no empírica:

- CS1: el estudio se basa en una investigación empírica.
- CS2: están bien definidos las preguntas, los objetivos y los propósitos de investigación del estudio.
- CS3: está bien definido el contexto del estudio.

Cada uno de estos criterios fue calificado en una escala dicotómica (1 = sí o 0 = no). De los 229 artículos se seleccionaron dieciocho con el objetivo de evaluar la calidad con base en los tres criterios de selección. Para aceptar un trabajo se definió que debía estar calificado con *sí* en el CS1 y con *sí* en CS2 o CS3.

La definición de lo que constituye un estudio de caso no siempre es evidente. Por ejemplo, la expresión *estudio de caso* aparece en los artículos de investigación en ingeniería de *software*, aunque se utiliza para ejemplos demasiado pequeños (Runeson y Höst, 2009). Por lo tanto para CS1, en esta RSL se utilizó la siguiente definición para estudio de caso: “Es un método empírico destinado a investigar fenómenos contemporáneos en su contexto” (Kitchenham, 2010, p. vi).

Cada uno de los dieciocho estudios fue evaluado de acuerdo con siete criterios de calidad (tabla 2). Estos siete criterios se basan en la lista de verificación para estudio caso propuesta por Runeson y Höst (2009).

Tabla 2. Criterios de evaluación para calidad

ID	Criterio de calidad
CC1	El diseño de la investigación es adecuado para hacer frente a las preguntas de investigación
CC2	Los procedimientos de recopilación de datos son suficientes para el propósito
CC3	Los procedimientos de análisis son suficientes para el propósito
CC4	Los resultados están claramente definidos y son creíbles, y las conclusiones están justificadas
CC5	Se adoptan diferentes puntos de vista en el caso
CC6	Las amenazas a la validez de los análisis se abordan de manera sistemática
CC7	Las conclusiones son implicaciones para la práctica

Fuente: presentación propia de los autores.

Los siete criterios de calidad de la tabla 2, junto con los tres criterios de selección, proporcionan una medida de la calidad de los estudios incluidos y una medida del grado de confianza en los resultados de cada estudio. En la tabla 3 se muestran todos los criterios de calidad clasificados en la misma escala dicotómica de los tres criterios de selección; las áreas se refieren a las cinco áreas identificadas en relación con la pregunta de investigación PI1.

Tabla 3. Valoración de la calidad de los trabajos incluidos

Estudios incluidos		Área	CS1	CS2	CS3	CC1	CC2	CC3	CC4	CC5	CC6	CC7	Total
1	Sibisi y Waveren 2007)	Priorización	1	1	1	1	1	1	1	1	1	1	10
2	Berntsson et ál. (2008)	Métricas	1	1	1	1	1	1	1	1	1	1	10
3	Cysneiros y Leite (1999)	Elicitación	1	1	1	1	1	1	1	1	0	1	9
4	Regnell et ál. (2007)	Estimación de costos	1	1	1	1	1	1	1	1	0	1	9
5	Hassenzahl et ál. (2001)	Elicitación	1	1	1	1	1	1	1	1	0	1	9
6	Kusters et ál. (1999-1)	Elicitación	1	1	1	1	1	1	1	1	0	1	9
7	Kusters et ál. (1999-2)	Elicitación	1	1	1	1	1	1	1	1	0	1	9

Continúa

Estudios incluidos		Área	CS1	CS2	CS3	CC1	CC2	CC3	CC4	CC5	CC6	CC7	Total
8	Boehm e In (1996)	Dependencias	1	1	1	1	1	1	1	1	0	1	9
9	Cleland-Huang et ál. (2005)	Dependencias	1	1	1	1	1	1	1	1	0	1	9
10	In y Boehm (2005)	Dependencias	1	1	1	1	1	1	1	1	0	1	9
11	In et ál. (2001)	Dependencias	1	1	1	1	0	1	1	1	1	1	9
12	Zulzalil et ál. (2008)	Dependencias	1	1	1	1	1	1	1	1	0	1	9
13	Olsson et ál. (2007)	Métricas	1	1	1	1	1	1	1	0	1	1	9
14	Johansson et ál. (2001)	Priorización	1	1	1	1	1	1	1	1	0	1	9
15	Leung (2001)	Priorización	1	1	1	1	1	1	1	1	0	1	9
16	Doerr et ál. (2005)	Elicitación	1	1	1	1	1	0	1	1	0	1	8
17	Regnell et ál. (2008)	Métricas	1	0	1	1	1	0	1	1	0	1	7
18	Jacobs (1999)	Métricas	1	1	1	0	0	0	1	1	0	1	6

Fuente: presentación propia de los autores.

1.4. Extracción y síntesis de los datos

Se utilizó una forma de extracción de datos predefinida. De cada estudio se extrajeron los siguientes datos: identificación de la publicación, información bibliográfica —autor(es), año, título e información de la publicación—, objetivos y preguntas de investigación, metodología —diseño del estudio, descripción de la muestra, contexto del estudio, recopilación de datos y análisis de datos—, herramientas/enfoques/técnicas utilizados, resultados, conclusiones y limitaciones. Se sintetizaron los datos para identificar los temas que se derivan desde los hallazgos reportados en cada uno de los artículos revisados en esta RSL, una revisión descriptiva-no cuantitativa.

2. Resultados

2.1. Existen evidencias empíricas acerca de requisitos no funcionales (PI1)

A continuación se analiza la pregunta de investigación PI1: ¿qué se conoce actualmente de las investigaciones empíricas acerca de requisitos no funcionales en relación con la elicitación, la priorización, la estimación de costos, las dependencias y las métricas?

2.1.1. Elicitación

En la tabla 4 se listan las cinco técnicas de elicitación que se han identificado para elicitar requisitos no funcionales. En los estudios de la muestra se aprecia que no existe una visión clara acerca de cómo elicitar requisitos no funcionales. Todas las técnicas, a excepción de TE3, han sido evaluadas en más de un tipo de sistema.

Tabla 4. Técnicas de elicitación identificadas

Técnica	Origen	Descripción
TE1	Cysneiros y Leite (1999)	Uso de un léxico
TE2	Doerr et ál. (2005)	Uso de un modelo de calidad, lista de chequeo y un cuestionario de priorización
TE3	Hassenzahl et ál. (2001)	Técnica de la entrevista
TE4	Kusters et ál. (1999-1) Kusters et ál. (1999-2)	Estrategia basada en cuestionario soportado en una herramienta
TE5	Kusters et ál. (1999-1) Kusters et ál. (1999-2)	Entrevistas estructuradas

Fuente: presentación propia de los autores.

En los estudios de la muestra se encontraron algunas sugerencias acerca de cómo elicitar requisitos no funcionales. Por ejemplo, Cysneiros y Leite (2004) sostienen que estos requisitos no deberían tratarse con el mismo alcance de los requisitos funcionales, debido a que los no funcionales requieren un razonamiento más detallado. Doerr et ál. (2005) argumentan que la elicitación de los requisitos funcionales, los no funcionales y la arquitectura deben estar ligados, porque el refinamiento de los no funcionales es imposible sin el detalle de los funcionales y la arquitectura; además, Hassenzahl et ál. (2001) afirman que es importante reunir diferentes aspectos, como requisitos no funcionales, enfoque de diseño y las relaciones entre ellos para asegurar una comprensión básica del problema que se va a diseñar. El punto de vista de Hassenzahl et ál. (2001) acerca de la relación entre requisitos no funcionales, el diseño y sus relaciones es respaldado por Cysneiros y Leite (2004), quienes establecen que los requisitos no funcionales tienen muchas interdependencias entre sí que pueden requerir “negociaciones” entre las diferentes decisiones de diseño.

En términos del logro de objetivos, las diferentes técnicas de elicitación resultaron ser prometedoras y elicitaron más requisitos no funcionales que otras técnicas utilizadas en los estudios de caso. TE1 genera entre el 20% y el 25% de nuevas clases, y el 46% de las clases existentes se cambiaron para satisfacer los

requisitos no funcionales elicitados. Doerr et ál. (2005) sostienen que con TE2 se descubrieron nuevos e importantes requisitos no funcionales, y que solo cinco de los 54 elicitados no eran medibles. Con TE3 se encontraron 172 cualidades concretas; sin embargo, no se encontró información acerca del porcentaje de nuevos requisitos no funcionales elicitados. Para TE5 y TE6, la elicitación dio lugar a una serie de cambios en las actividades de desarrollo previstas.

2.1.2. Dependencias

La tabla 5 muestra los tres métodos de dependencia (MD) que se han estudiado empíricamente en relación con los requisitos no funcionales. Los tres métodos cuentan con el sustento de herramientas que se utilizan en su evaluación empírica. Además de los métodos identificados, en uno de los trabajos (Olsson et ál., 2007) se investiga acerca de la existencia de interdependencias entre los requisitos no funcionales.

Tabla 5. Métodos de dependencia identificados

MD	Origen	Descripción
MD1	Boehm e In (1996)	QARCC
MD2	In y Boehm (2001) In et ál. (2001)	QARCC y S-COST
MD3	Cleland-Huang et ál. (2005)	Trazabilidad centrada en objetivos

Fuente: presentación propia de los autores.

Zulzalil et ál. (2008) analizaron las interdependencias entre los requisitos no funcionales para tres aplicaciones web diferentes. En dos de esas tres aplicaciones encontraron una fuerte dependencia entre la usabilidad y la eficiencia, la funcionalidad y la eficiencia, y la usabilidad y la fiabilidad. Por otra parte, en una de las aplicaciones encontraron una fuerte dependencia entre la usabilidad y la funcionalidad, y entre la funcionalidad y la fiabilidad.

Al igual que en Zulzalil et ál. (2008), Boehm e In (1996), In y Boehm (2001) e In et ál. (2001) analizan los conflictos entre los requisitos no funcionales. Sin embargo, en estos estudios se evaluaron los métodos MD1 y MD2 para identificar los conflictos. Cleland-Huang et ál. (2005) también evaluaron un método, el MD3. No obstante, su enfoque es el impacto del cambio funcional en los requisitos no funcionales. Los tres métodos fueron evaluados y comparados con los enfoques manuales. MD1 y MD2 fueron comparados con el uso manual del modelo WinWin (In y Boehm, 2001), mientras que MD3 fue comparado

con el número de enlaces que se requieren en un análisis manual. Todos los métodos evaluados mostraron resultados prometedores en comparación con los manuales, como se ilustra en la tabla 6.

Tabla 6. Eficiencia de las herramientas de dependencias

MD	Enfoque manual	Herramienta
MD1	2 conflictos encontrados	10 conflictos encontrados
MD2	7 conflictos encontrados	76 conflictos encontrados
MD3	360 posibles comparaciones	22 enlaces para analizar

Fuente: presentación propia de los autores.

Con MD1 y MD2 se identificaron más conflictos de calidad en comparación con el uso manual de WinWin. En un estudio de In (2001) se identificaron siete conflictos de calidad manualmente y 76 con MD2 (24 insignificantes). Además, con MD2 se identificaron cinco de los siete conflictos encontrados con el enfoque manual; con MD1 y MD2 se identificaron los mismos conflictos hallados con el uso manual de WinWin. Además, con la herramienta identificaron ocho conflictos más, de los cuales cinco fueron significativos. Con MD3 evaluaron la capacidad de los métodos para reducir el número de enlaces para gestionar el impacto del cambio funcional sobre los requisitos no funcionales. Mientras que un análisis manual requiere 360 enlaces para ser analizados (Cleland-Huang et ál., 2005), MD3 reduce ese número a 22.

2.1.3. Métricas

De acuerdo con Jacobs (1999), es necesario establecer requisitos medibles. Uno de los objetivos de técnica de elicitación TE2 es lograr un conjunto de requisitos no funcionales medibles, y casi el 90% de los elicitados con la técnica fueron cuantificados. Por otro lado, cuando Olsson et ál. (2007) investigaron una especificación de requisitos de un estudio de caso, en que el 40% de los 2113 requisitos eran no funcionales, fue posible cuantificar cerca de la mitad de ellos. También analizaron la forma en se cuantifican los requisitos no funcionales, y encontraron que se utiliza una mezcla de diferentes escalas:

- Absoluta: sin intervalo dado (58%).
- Min-Max: se especifica un límite inferior y superior, creando un intervalo min-max (7%).
- Límite superior: intervalo unilateral (24%).
- Límite inferior: intervalo unilateral (12%).

Por otra parte, manifiestan que la naturaleza de los intervalos y escalas especificados para las diferentes áreas, incluso dentro del mismo producto de *software*, son importantes para la negociación y la priorización. Es decir, intervalos y escalas deben estar alineados con el mercado y con el valor del costo.

Se identificaron, además, dos enfoques para cuantificar requisitos no funcionales: QUPER (Berntsson et ál., 2008; Regnell et ál., 2008) y el estilo Gilb (Jacobs, 1999). La importancia de los enfoques específicos para cuantificar estos requisitos es apoyado por Jacobs (1999), quien afirma que es más difícil y esencial cuantificar requisitos no funcionales que funcionales. El modelo QUPER considera la alineación entre los intervalos y los valores de mercado sugerido por Olsson et ál. (2007).

Los dos enfoques tienen características similares, ambos tienen conceptos para identificar mediciones para los productos actuales de los proveedores y lo que el mercado espera. Además, ambos enfoques cuantifican requisitos no funcionales para un intervalo, similar al intervalo Min-Max en Olsson et ál. (2007). Sin embargo, el intervalo Min-Max es la escala menos utilizada en Olsson et ál.; pero lo interesante es que tanto QUPER como el estilo Gilb sugieren el uso de intervalos.

Ambos enfoques fueron evaluados suficientemente. En QUPER, la relación entre los intervalos y el valor de mercado es visto como una característica importante, y ajustando el intervalo de requisitos no funcionales con base en el segmento de mercado y la calidad de los competidores, se proporciona una mejor base para las métricas reales. Al introducir el estilo Gilb, se logra una comprensión común de los requisitos no funcionales, lo que, de acuerdo con Jacobs (1999), es crucial. Por otra parte, mediante la especificación de los conceptos del estilo Gilb para cada requisito no funcional, Jacobs encontró que los casos de prueba se pueden definir desde la fase de ingeniería de requisitos.

2.1.4. Estimación de costos

Solo un artículo (Regnell et ál., 2007) relaciona la estimación del costo de los requisitos no funcionales. Sin embargo, no se refiere a cómo se lleva a cabo esta estimación; en su lugar identifica cómo prever las barreras de los costos y evalúan si es posible prever cuándo es necesaria una inversión importante —en términos de costos—, para mejorar la calidad. De este modo se encontró que el costo de alcanzar una determinada calidad está relacionado con la optimización del *software*, la inversión en *hardware*, el esfuerzo de desarrollo, las inversiones en nueva arquitectura y los derechos de las licencias. Además, que la estimación

de costos es más incierta en las primeras etapas de una nueva tecnología en comparación con una tecnología que ya está disponible y alcanzó la madurez en un mercado determinado.

2.1.5. Establecimiento de prioridades

Se identificaron tres estudios (Johansson et ál., 2001; Leung, 2001; Sibisi y van Waveren, 2007), que relacionan la priorización de los requisitos no funcionales más importantes. El análisis hecho a los datos extraídos evidenció que ninguna técnica de priorización que se centre únicamente en los requisitos no funcionales se ha evaluado empíricamente. Sin embargo, en general, las técnicas de priorización de requisitos se pueden utilizar a la hora de darles preferencia a requisitos no funcionales; por lo tanto, están fuera del alcance de esta revisión. Dos de los tres estudios (Johansson et ál., 2001; Leung, 2001) abordan cómo las diferentes partes interesadas priorizan la importancia de los requisitos no funcionales; mientras que (Sibisi y van Waveren, 2007) desarrollan y evalúan un marco de proceso para destacar esa importancia.

Se encontró que las diferentes partes interesadas tienen diferentes puntos de vista respecto a la importancia de los requisitos no funcionales, tanto entre las de dentro de la misma empresa, donde los objetivos son los mismos (Johansson et ál., 2001) como entre los diferentes dominios (Johansson et ál., 2001; Sibisi y van Waveren, 2007). Johansson et ál. (2001) encontraron que la fiabilidad era identificada por una multitud de interesados para enfocarla en los más importante requisitos no funcionales; y que en Leung (2001) también era identificada como el requisito funcional más importante para las aplicaciones de intranet.

Entre tanto, Sibisi y van Waveren (2007) informan que la funcionalidad, para sus dos proyectos, es uno de los requisitos no funcionales más importantes. Los resultados revelan que dentro de la misma organización, las partes interesadas, como mercadeo, diseñadores de sistemas y arquitecturas pueden tener diferentes puntos de vista acerca de a cuáles requisitos no funcionales se deba considerar importantes (Johansson et ál., 2001). Una explicación puede ser, por ejemplo, que los arquitectos y los diseñadores de sistemas trabajan en diferentes niveles de la arquitectura (Johansson et ál., 2001).

Para ayudar a las organizaciones a personalizar un modelo de calidad del *software* para sus necesidades específicas, Sibisi y van Waveren (2007) desarrollaron un marco de procesos para vincular desde el estándar ISO/IEC 9126-1 a las necesidades del usuario, las características de calidad y sus subcaracterísticas. El resultado de validación revela que el marco es válido en la esfera de las

características; sin embargo, no se encontró ningún estudio o evaluación acerca del tiempo necesario para aprender y calibrar este marco para las necesidades específicas de cada organización.

2.2. Métodos de investigación empírica (PI2)

En cuanto a qué métodos empíricos se han utilizado, en la tabla 7 se observa que las dos terceras partes de la muestra (doce artículos) incluyeron el uso estudio de casos como método de investigación para evaluar empíricamente los requisitos no funcionales, lo que lo convierte en el método de investigación más común. Siete de estos artículos de estudio de caso utiliza un solo caso, mientras que los otros cinco utilizan múltiples casos.

Tabla 7. Distribución de los métodos de investigación

Método de investigación	Número de publicaciones
Estudio de caso	12 (7 uno solo, 5 múltiples)
Experimento	3
Encuesta	2
Mixto (encuesta y experimento)	1

Fuente: presentación propia de los autores.

Se podría argumentar que esto indica que existe una necesidad de evaluar los requisitos no funcionales y sus técnicas, utilizando otros métodos de investigación diferentes al estudio de caso. Esta idea solo la apoya Seaman (1999), quien argumenta que las cuestiones de la ingeniería del *software* son las más investigadas, pero que para ello se utiliza una combinación de métodos cualitativos y cuantitativos. Sin embargo, mirando qué métodos de investigación se han utilizado para cada una de las categorías identificadas en relación con los requisitos no funcionales, los resultados muestran que todos los estudios, en relación con la elicitación y las métricas, han utilizado estudios de caso. Por otra parte, los tres experimentos se llevan a cabo en relación con las dependencias de los requisitos no funcionales; mientras que todas las encuestas investigan la priorización de estos mismos requisitos.

3. Discusión

El objetivo de la revisión sistemática fue recopilar y comparar la evidencia empírica existente acerca de los requisitos no funcionales y, además, proporcionar el panorama actual en esta temática para servir de base a futuras investigaciones empíricas.

3.1. Beneficios y limitaciones

Los dieciocho estudios de la muestra se clasifican en cinco categorías en relación con las actividades de ingeniería de requisitos. La revisión mostró que se llevan a cabo más estudios empíricos acerca de los requisitos no funcionales respecto a la elicitación y las categorías de dependencias. Los estudios que se centran en la elicitación de requisitos no funcionales no proporcionan un punto de vista unificado de la práctica actual; en cambio, ofrecen un amplio panorama de las experiencias y las técnicas probadas. En esta revisión sistemática se identificaron cinco diferentes técnicas de elicitación de estos requisitos; cada técnica se evaluó en uno, dos o tres sistemas diferentes, y todas resultaron ser prometedoras para reunir requisitos no funcionales.

Davis et ál. (2006) encontraron que la entrevista es la técnica de elicitación más efectiva para recopilar estos requisitos. En esta revisión sistemática, tres de las cinco técnicas utilizan entrevistas para elicitar los requisitos no funcionales. Por otra parte, los hallazgos de Davis et ál. solo se confirman parcialmente en este estudio. Respecto a las limitaciones, Doerr et ál. (2005) sostienen que la elicitación de requisitos no funcionales y funcionales debe estar entrelazada con la arquitectura. Por otra parte, Cysneiros y Leite (2004) sostienen que ambos requisitos no deben tratarse dentro del mismo ámbito de aplicación. La razón es que los requisitos no funcionales necesitan un razonamiento detallado. Sin embargo, la idea de Doerr et ál. (2005) apoya el estudio de van Lamsweerde (2000), quien identificó la necesidad de cerrar la brecha entre requisitos y arquitectura.

Respecto a la manipulación de las dependencias de los requisitos no funcionales, se identificaron tres métodos diferentes. En la práctica actual no se proporciona una visión unificada en este aspecto. En dos estudios (Boehm e In, 1996; In y Boehm, 2001) se hace una comparación entre el enfoque manual del modelo WinWin con los métodos QARCC y QARCC en combinación con el método S-COST. Ambos estudios mostraron que el modelo WinWin fue menos eficaz en la identificación de los conflictos entre los requisitos no funcionales. Sin embargo, no se presenta una comparación entre el método QARCC y la combinación de los métodos QARCC y S-COST.

Van Lamsweerde (2000) sugiere la necesidad de una herramienta de soporte en la ingeniería de requisitos. No se encontró otra comparación de las técnicas de dependencia. En cuanto a las limitaciones, el modelo WinWin es el único método evaluado en más de un estudio; por lo tanto, no fue posible determinar cuál método identifica mejor las interdependencias entre los requisitos no funcionales.

Respecto a la priorización de los requisitos no funcionales, solo se identificaron tres estudios. Sin embargo, cabe señalar que ninguno de los de la muestra describe

qué técnica o método puede ser más eficiente para priorizar los requisitos no funcionales. Ello indica la importancia de comprender cómo se priorizan estos requisitos y, además, que puede ser un vacío en la literatura. En esta SLR se identificó solo un estudio relacionado con los requisitos no funcionales y la estimación de costos. Una posible explicación es que la cadena de búsqueda no identificó los estudios existentes, lo que puede explicarse porque en la literatura se utilizan ciertos sinónimos para los términos buscados, algo que discuten Jørgensen y Shepperd (2007). Estos investigadores llegan a la conclusión de que para que una búsqueda más amplia tenga sentido, se deben incluir más sinónimos, lo que se traducirá en un conjunto muy grande de estudios encontrados.

Otra explicación puede ser que la estimación de costos de los requisitos funcionales y no funcionales se realiza de manera similar. En esta revisión no se encontró ningún estudio en el que se trate la forma de estimar el costo de los requisitos no funcionales; por lo que, de acuerdo con el conocimiento adquirido en esta revisión y las anteriores (Kitchenham et ál., 2009), es posible concluir que falta de evidencia acerca de cómo se lleva a cabo la estimación del costo de los requisitos no funcionales.

Se argumenta que para la estimación de costos, sin importar si es para requisitos funcionales o no funcionales, es posible utilizar el mismo proceso. Sin embargo, los estudios (Brooks, 1987; Cysneiros y Leite, 1999) han demostrado que los no funcionales son caros y difíciles de manejar, y de acuerdo con Chung et ál. (2000), los no funcionales son a menudo poco conocidos. Esto indica la importancia de comprender cómo se lleva a cabo el procedimiento de estimación de costos de estos requisitos, lo que puede ser un vacío en la literatura.

3.2. Fuerza de la evidencia

La discusión de la solidez de las evidencias se basa en dos categorías: el diseño y la calidad del estudio (tabla 3). En Dybå y Dingsyr (2008), el diseño del estudio se refiere a los métodos de evaluación utilizados, donde los experimentos se consideran la evidencia más fuerte, mientras que los estudios observacionales son evidencias débiles. Además, Seaman (1999) argumenta que, para investigar los problemas de ingeniería del *software*, es preferible una combinación de métodos cualitativos y cuantitativos. Doce de los dieciocho estudios incluidos en la muestra de la actual revisión son estudios de caso, y ello puede implicar que la evidencia total de los estudios combinados es baja en relación con el diseño del estudio. Respecto a la calidad de los métodos de los estudios incluidos, a veces, no están bien descritos. Cuestiones como parcialidad, validez y procedimientos de análisis de datos adecuados no siempre se tratan (véase tabla 3). Solo

en cuatro de los dieciocho estudios se abordan cuestiones de validez de forma sistemática. Con base en estos resultados, se concluye que existen limitaciones para la calidad de los estudios.

4. Conclusiones

En esta revisión sistemática de la literatura se clasificaron los estudios en cinco áreas en relación con los requisitos no funcionales: elicitación, dependencias, métricas, estimación de costos y priorización. En cuanto a la pregunta de investigación PI1: ¿qué se conoce actualmente de las investigaciones empíricas acerca de requisitos no funcionales en relación con la elicitación, la priorización, la estimación de costos, las dependencias, y las métricas?, los estudios se clasifican en cinco categorías.

Los estudios que tratan la elicitación de los requisitos no funcionales no proporcionan un punto de vista unificado. Por ejemplo, Cysneiros y Leite (2004) argumentan que estos requisitos no deben tratarse en el mismo ámbito de los funcionales; mientras que Doerr et ál. (2005) y Hassenzahl et ál. (2001) sostienen que los requisitos no funcionales y otros aspectos como el diseño y la arquitectura deben tratarse de forma conjunta. Se evaluaron cinco técnicas de elicitación diferentes, de las cuales tres se basan en entrevistas. Este resultado está en línea con Davis et ál. (2006), quienes encontraron que las entrevistas constituyen la técnica de elicitación más comúnmente utilizada.

Respecto al manejo de las dependencias de los requisitos no funcionales, se identificaron tres métodos con el soporte de una herramienta. Estos métodos fueron más eficientes que los enfoques manuales utilizados en la evaluación. Los resultados empíricos de esta revisión indican que las herramientas ayudan a los profesionales a hacer frente a las dependencias entre los requisitos no funcionales de forma más eficiente que con los métodos manuales. Este resultado está en línea con van Lamsweerde (2000), quien apoya la necesidad de una herramienta de soporte en los procesos de la ingeniería de requisitos. Sin embargo, no se encontraron comparaciones entre los tres métodos; por lo tanto, no es posible identificar cuál es el método más adecuado para identificar las dependencias entre los requisitos no funcionales.

En cuanto a las métricas para los requisitos no funcionales, se encontraron dos enfoques con características similares para la cuantificar estos requisitos. Ambos enfoques sugieren la cuantificación de los requisitos no funcionales mediante intervalos; sin embargo, Olsson et ál. (2007) encontraron que un intervalo Min-Max es la escala menos utilizada en la industria, mientras que un valor absoluto es la cuantificación más común.

Se identificó solo un estudio relacionado con la estimación de costos de los requisitos no funcionales; sin embargo, ese estudio no presenta un método o técnica para dicha estimación. Una razón puede ser que los métodos de estimación de costos son empíricamente evaluados para requisitos en general, y sin importar si se trata de funcionales o no funcionales, se puede utilizar el mismo método. Sin embargo, los estudios de Brooks (1987) y Cysneiros y Leite (1999) muestran que los no funcionales son más difíciles de manejar.

Ninguno de los estudios relacionados con la priorización tuvo en cuenta técnicas o métodos para priorizar los requisitos no funcionales. En cambio, la evidencia empírica muestra que las diferentes partes interesadas tienen diferentes puntos de vista respecto a la importancia de estos requisitos, incluso dentro de la misma empresa, donde los objetivos son los mismos. Además, en dos de los tres estudios, la fiabilidad fue identificada como el requisito no funcional más importante, mientras que la calidad como aspecto de la funcionalidad es vista como el más importante en el tercer estudio.

Los resultados para la pregunta de investigación PI2: ¿qué métodos de investigación empírica se han utilizado para evaluar los requisitos no funcionales? muestran que el estudio de caso es el método de investigación más utilizado. No obstante, mirando qué método de investigación se ha utilizado en cada una de las cinco áreas, los resultados son diferentes. Solo los estudios de caso se utilizan en las áreas de elicitación y las métricas, mientras que las encuestas se utilizan en relación con la priorización de los requisitos no funcionales.

Respecto a las limitaciones, no se percibe un punto de vista unificado de las prácticas actuales. Pocos estudios son reproducidos, por lo que la posibilidad de sacar conclusiones con base en variaciones es limitada. Para que los profesionales utilicen estos resultados a efecto de seleccionar qué método o técnica pueden aplicar, deben considerarlos y compararlos con el contexto real en el que el método o la técnica se va a aplicar.

Una conclusión clara de esta revisión es la necesidad de aumentar el número y la calidad de los estudios acerca de los requisitos funcionales; por lo tanto, se proponen algunos trabajos futuros:

Evaluar empíricamente los requisitos no funcionales en diversas actividades de ingeniería de requisitos para llenar las lagunas identificadas en la literatura.

La priorización de los requisitos no funcionales requiere más atención. Ni un solo estudio de la muestra examina las técnicas de priorización de estos requisitos.

Es necesario repetir los estudios en contextos diferentes y ampliarlos a entornos más complejos.

Referencias

- BERNTSSON, S. R.; OLSSON, T. y REGNELL, B. Introducing support for release planning of quality requirements: An industrial evaluation of the QUPER model. *The Second International Workshop on Software Product Management (IWSPM 08)*. Barcelona, España, 2008, pp. 18-26.
- BOEHM, B. e IN, H. Identifying quality-requirements conflict. *IEEE Software*. 1996, vol. 12, núm. 6, pp. 25-35.
- BREITMAN, K. K.; LEITE, J. C. S. P y FINKELSTEIN, A. The world's stage: a survey on requirements engineering using a real-life case study. *Journal of the Brazilian Computer Society*. 1999, vol. 6, núm. 1, pp. 13-38.
- BROOKS, F. P. No silver bullet: essences and accidents of software engineering. *IEEE Computer*. 1987, vol. 4, Apr. pp. 10-19.
- CHUNG, L.; NIXON, B. A.; YU, E. y MYLOPOULOS, J. *Non-functional requirements in software engineering*. Norwell, USA: Kluwer Academic Publishers, 2000.
- CLELAND-HUANG, J. et ál. Goal-centric traceability for managing non-functional requirements. *27th International Conference on Software Engineering (ICSE 05)*. St. Louis, USA, 15-21, 2005, pp. 362-371.
- CYSNEIROS, L. M. y LEITE, J. C. S. P Integrating non-functional requirements into data model. *Fourth IEEE International Symposium on Requirements Engineering (ISRE 99)*. Limerick, Ireland, 7-11, 1999, pp. 162-171.
- CYSNEIROS, L. M. y LEITE, J. C. S. P Nonfunctional requirements: from elicitation to conceptual models. *IEEE Transactions on Software Engineering*. 2004, vol. 30, núm. 5, pp. 328-349.
- DAVIS, A. et ál. Effectiveness of requirements elicitation techniques: empirical results derived from a systematic review. *14th IEEE International Requirements Engineering Conference (RE'06)*. Minneapolis, USA, 11-15, 2006, pp. 176-185.
- DOERR, J. et ál. Nonfunctional requirements in industry - Three case studies adopting an experience-based NFR method. *13th IEEE International Conference on Requirements Engineering (RE 05)*, Paris, France, 2005, pp. 373-382.
- DYBÅ, T. y DINGSYR, T. Empirical studies of agile software development: A systematic review. *Information and Software Technology*. 2008, vol. 50, núm. 9-10, pp. 833-859.
- FINKELSTEIN, A. y DOWELL, J. A comedy of errors: the London ambulance service case study. *Eight International Workshop on Software Specification and Design*. Schloss Velen, Germany, 22-23, 1996, pp. 2-4.
- HASSENZAHN, M.; WESSLER, R. y HAMBORG, K. C. Exploring and understanding product qualities that users desire. *5th Annual Conference of the Human-Computer Interaction Group of the British Computer Society (IHM-HCI 01)*, Lille, France, 2001, pp. 95-96.
- IN, H. y BOEHM, B. W. Using WinWin quality requirements management tools: a case study. *Annals of Software Engineering*. 2001, vol. 11, núm. 1, pp. 141-174.

- IN, H.; BOEHM, B. W. y DEUTSCH, M. Applying WinWin to quality requirements: a case study. *23rd International Conference on Software Engineering (ICSE 01)*, Miami, USA, 2001, pp. 555-564.
- JACOBS, S. Introducing measurable quality requirements: a case study. *Fourth IEEE International Symposium on Requirements Engineering (ISRE 99)*, Limerick, Ireland, 1999, pp. 172-179.
- JOHANSSON, E. et ál. The importance of quality requirements in software platform development-a survey. *34th Annual Hawaii International Conference on System Sciences*, Hawaii, USA, 3-6, 2001.
- JØRGENSEN, M. y SHEPPERD, M. A systematic review of software development cost estimation studies. *IEEE Transactions on Software Engineering*, 2007, vol. 31, núm. 1, pp. 33-53.
- KITCHENHAM, B. A. et ál. Systematic literature reviews in software engineering - A systematic literature review. *Information and Software Technology*. 2009, vol. 51, núm. 1, pp. 7-15.
- KITCHENHAM, B. A. Guidelines for performing Systematic Literature Reviews in Software Engineering. *Technical Report* [documento en línea]. *EBSE-2007-001*, UK. 2010. <<http://www.dur.ac.uk/ebse/>>.
- KUSTERS, R. J.; VAN SOLINGEN, J. y TRIENEKENS, J. J. M. Identifying embedded software quality: two approaches. *Quality and Reliability Engineering International*. 1999, vol. 15, núm. 6, pp. 485-492.
- KUSTERS, R. J.; VAN SOLINGEN, J. y TRIENEKENS, J. J. M. Strategies for the identification and specification of embedded software quality. *Ninth International Workshop Software Technology and Engineering Practice (STEP 99)*, Pittsburgh, USA, 1999-2, pp. 33-39.
- LEUNG, H. K. N. Quality metrics for intranet applications. *Information and Management*. 2001, vol. 38, núm. 3, pp. 137-152.
- OLSSON, T.; BERTSSON S. R. y REGNELL, B. Non-functional requirements metrics in practice: An empirical document analysis. *Workshop on Measuring Requirements for Project and Product Success (MeReP 07)*, Palma de Mallorca, Spain, 2007.
- REGNELL, B.; BERTSSON, S. R. y OLSSON, T. Supporting roadmapping of quality requirements. *IEEE Software*. 2008, vol. 25, núm. 2, pp. 42-47.
- REGNELL, B.; HÖST, M. y BERTSSON S. R. A Quality performance model for cost-benefit analysis of nonfunctional requirements applied to the mobile handset domain. *13th International Working Conference on Requirement Engineering: Foundation for Software Quality (REFSQ 07)*, Trondheim, Norway, 11-12, 2007, pp. 277-191.
- RUNESON, P. y HÖST, M. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*. 2009, vol. 14, núm. 2, pp. 131-164.
- SEAMAN, C. Qualitative methods in empirical studies of software engineering. *IEEE Transactions on Software Engineering*. 1999, vol. 25, núm. 4, pp. 557-572.
- SIBISI, M. y VAN WAVEREN, C. C. A process framework for customising software quality models. *IEEE AFRICON Conference*, Windhoek, Namibia, 26-28, 2007, pp. 547-554.

- van LAMSWEERDE, A. Requirements engineering in the year 00: a research perspective. *2000 International Conference on Software Engineering (ICSE'00)*, Limerick, Ireland, 4-11, 2000, pp. 5-19.
- ZULZALIL, H. et ál. Relationships analysis between quality factors for Web applications. *International Symposium on Information Technology (ITSIM 08)*, Kuala Lumpur, Malaysia, 26-28, 2008, pp. 1-5.