

Vea discusiones, estadísticas y perfiles de autor para esta publicación en: <https://www.researchgate.net/publication/297764945>

Una revisión sistemática de mapeo de métodos de evaluación de usabilidad para el proceso de desarrollo de software

Artículo · Enero de 2016

DOI: 10.14257/ijseia.2016.10.1.16

Citas

50

LEA

1,320

2 autores, incluso:



Paz

Pontificia Universidad Católica del Perú

52 PUBLICACIONES 217 Citas

VER EL PERFIL

Algunos de los autores de esta publicación también están trabajando en estos proyectos relacionados:



Nota de usabilidad en herramientas educativas: Una revisión sistemática [Ver Proyecto](#)



Chatbot en la ciudad [Ver proyecto Freddy](#)

Una revisión sistemática de mapeo de métodos de evaluación de usabilidad para el proceso de desarrollo de software

Freddy Paz ¹ y José Antonio Pow-Sang ¹
¹ Pontificia Universidad Católica del Perú, Av.

Universitaria 1801, San Miguel, Lima 32, Perú
fpaz@pucp.pe , japowsang@pucp.edu.pe

Resumen

Dado que la usabilidad es uno de los aspectos más importantes de la calidad del software, se han desarrollado varios métodos para establecer técnicas capaces de evaluar este atributo desde las primeras fases del proceso de desarrollo del software. Sin embargo, la elección del método más apropiado para un escenario particular sigue siendo una decisión difícil, debido a la existencia de una gran cantidad de enfoques que se describen en la literatura para este propósito. Por lo tanto, se realizó una revisión sistemática de mapeo para identificar las técnicas de evaluación de usabilidad más utilizadas en los desarrollos de software. Se identificaron un total de 1169 estudios, de los cuales solo 215 estudios fueron seleccionados para esta revisión. Según el análisis, la mayoría de los estudios de casos establecen el uso de cuestionarios de usabilidad como herramienta de evaluación. Adicionalmente, La informática sanitaria y las aplicaciones web son el dominio del software y el tipo de aplicación que se informa con frecuencia en estas evaluaciones. Este trabajo ha permitido alcanzar resultados prometedores en esta área. Su objetivo es ser una guía para especialistas para apoyar la elección del método más adecuado para un escenario particular.

Palabras claves: *Interacción humano-computadora, revisión sistemática del mapeo, métodos de evaluación de usabilidad, proceso de desarrollo de software, diseño centrado en el usuario*

1) Introducción

Hoy en día, la usabilidad se considera uno de los aspectos más importantes para el éxito de cualquier producto tecnológico. En el contexto del software, si un producto es difícil de usar o proporciona mecanismos que son difíciles de entender, entonces se espera que la aplicación falle [15]. Dado el dominio actual, en el que hay varias alternativas disponibles para cualquier producto de software, los usuarios solo tendrán preferencia por un software capaz de ayudarlos a alcanzar sus objetivos con satisfacción. Por esta razón, la usabilidad se ha vuelto altamente relevante, especialmente, durante todas las fases del proceso de desarrollo de software [8]. Los desarrolladores son conscientes de que solo un pequeño porcentaje de usuarios pasa su tiempo leyendo un manual. Por lo tanto, el diseño de una interfaz gráfica debe ser lo suficientemente intuitivo como para satisfacer las expectativas de los usuarios con respecto a la usabilidad.

La importancia de este atributo de calidad ha llevado al desarrollo de varios métodos de evaluación de usabilidad, cuyo propósito es determinar sistemáticamente el grado en que un producto de software es fácil de usar. Sin embargo, debido a una amplia gama de estas técnicas, la elección del método más adecuado para un escenario particular se ha convertido en una decisión difícil. No hay acuerdo entre especialistas sobre cuál es el mejor método. La llegada de numerosas técnicas ha resultado en una extensa discusión para determinar el método más ampliamente aceptado por la comunidad científica. Además, actualmente se informan nuevas variantes basadas en las propuestas tradicionales en la literatura, sin una evidencia concreta de que estos enfoques todavía se emplean como procedimientos para evaluar la usabilidad de los productos de software [11].

El objetivo de este estudio fue examinar todos los estudios de casos que se describieron recientemente en la literatura para determinar los principales métodos de evaluación de usabilidad. En un trabajo anterior [10], realizamos una revisión preliminar de las técnicas existentes que estaban disponibles para este propósito. En este documento, presentamos una revisión sistemática extendida, mejorada y actualizada para determinar las tendencias actuales en el uso de métodos de evaluación de usabilidad para contextos de desarrollo de software. Un análisis de los resultados ha permitido identificar las técnicas más utilizadas para cada categoría de software. Estos hallazgos están destinados a servir como guía para especialistas para apoyar la elección correcta de un método apropiado en un escenario particular.

Este documento está estructurado de la siguiente manera. En la Sección 2, describimos los conceptos principales que se utilizan en la interacción hombre-computadora para estudios de esta naturaleza. En la Sección 3, presentamos la metodología que se utilizó para realizar este estudio. En la Sección 4, discutimos los resultados de nuestra investigación. Finalmente, las conclusiones y trabajos futuros se establecen en la Sección 5.

2) Antecedentes

2.1. Usabilidad

En Ingeniería de Software, el término "usabilidad" está relacionado con la facilidad de uso de un producto de software. Sin embargo, este concepto se puede aplicar a cualquier interfaz tecnológica que permita la interacción entre humanos y máquinas. De acuerdo con la norma ISO 9241-11 [12], la usabilidad se puede definir como "la medida en que cualquier producto puede ser utilizado por usuarios específicos para lograr objetivos específicos con efectividad, eficiencia y satisfacción en un contexto específico de uso". Aunque esta definición no está especialmente relacionada con los productos de software, se destaca la relevancia de la experiencia del usuario.

En el dominio de la informática, el estándar ISO 9126-1 [13] proporciona una definición más especializada donde la usabilidad se considera como un atributo de la calidad del software. El concepto de usabilidad se define como "la capacidad de un producto de software para ser entendido, gustado, usado y atractivo cuando se usa bajo condiciones específicas". Esta definición enfatiza no solo la relevancia de una interfaz gráfica estética e intuitiva, sino también el potencial de un software específico para satisfacer las expectativas del usuario.

2.1. Métodos de evaluación de usabilidad

Dada la relevancia de la usabilidad en el contexto del proceso de desarrollo de software, han surgido varios métodos de evaluación. El propósito de estas técnicas es establecer, de manera sistemática, el nivel de usabilidad de las interfaces gráficas de software [4].

Según Fernández *et. al.*, [2], los métodos de evaluación de usabilidad pueden definirse como "procedimientos compuestos por una serie de actividades bien definidas para recopilar datos relacionados con la interacción entre el usuario final y un producto de software, con el fin de determinar cómo las propiedades específicas de un software en particular contribuir al logro de objetivos específicos ". Por lo general, estos métodos se emplean durante todas las fases del proceso de desarrollo de software para garantizar el diseño de un producto utilizable que pueda cumplir con los estándares de alta calidad.

3) Realización de la revisión sistemática de mapeo

Una revisión sistemática de la literatura es un método para analizar, evaluar e interpretar todos los estudios relevantes a una pregunta de investigación en particular, o área específica, o fenómeno de interés. Una revisión sistemática de mapeo es una variante de esta técnica en la que la evidencia se representa con un alto nivel de granularidad. Aunque la SLR se usa ampliamente en medicina, existen propuestas para utilizar esta metodología en el campo de la ingeniería de software. Kitchenham y Charters [5] establecen un conjunto de pautas para realizar estudios sistemáticos confiables, rigurosos y auditables para temas de ingeniería de software. Este protocolo define procedimientos bien definidos para identificar y resumir toda la información existente sobre un tema

importa de una manera minuciosa e imparcial. Este trabajo se realizó de acuerdo con los parámetros definidos por Kitchenham y Charters [5]. Los pasos de esta metodología se presentan en las secciones siguientes.

3.1. Preguntas de investigación

El propósito de este trabajo fue determinar las tendencias actuales en el uso de métodos de evaluación de usabilidad para procesos de desarrollo de software. Además, identificamos las categorías de software que se informan con frecuencia como parte de estas inspecciones de usabilidad, y las técnicas que se emplean comúnmente para cada tipo. De esta manera, formulamos las siguientes preguntas de investigación:

RQ1: ¿Cuáles son las técnicas más utilizadas para evaluar la usabilidad de los productos de software en el contexto de un proceso de desarrollo?

RQ2: ¿Qué métodos de evaluación de usabilidad se usan comúnmente para cada categoría de aplicación de software involucrada en un proceso de desarrollo?

Para realizar esta revisión, definimos los conceptos generales basados en PICOC. Dado que nuestra investigación no pretende comparar intervenciones, no se consideró el criterio de "comparación". Estos conceptos se detallan en la Tabla 1.

Tabla 1. Definición de los conceptos generales usando PICOC

Criterio	Descripción
Población	Productos de software.
Intervención	Métodos de evaluación de usabilidad. Resultados
Contexto	Estudios de casos en los que una evaluación de usabilidad se considera parte de un proceso de desarrollo de software. Contexto académico, industria del software y todo tipo de estudios empíricos.

3.2. Estrategia de búsqueda

Definimos nuestra estrategia de búsqueda basada en los conceptos generales. Algunos fueron seleccionados para lograr una búsqueda más completa. Solo consideramos estudios relevantes, cuya fecha de publicación fue desde 2012, para analizar las tendencias y el estado actual del arte en este campo. La cadena resultante fue:

("desarrollo de software" O "construcción de software" O "proyecto de software" O "proyectos de software" O "proceso de software" O "procesos de software" O "Ingeniería de software" O "pruebas de software" O "diseño de software" O "verificación de software" O "validación de software") Y ("método" O "técnica" O "proceso" O "procedimiento" O "Acercarse") Y ("prueba" O "evaluación" O "inspección" O "evaluación" O "medición" O "estudiar" O "estudios") Y ("usabilidad" O "usable") Y (año de publicación > 2011)

3.3. Proceso de búsqueda

El proceso de búsqueda se realizó mediante el uso de tres bases de datos reconocidas para buscar estudios primarios: SCOPUS, Biblioteca Digital ACM e ISI (considerando Web of Science & Web of Knowledge). No se consideró ningún estudio adicional. Una vez que se recuperaron los documentos, utilizamos Google Scholar para determinar la relevancia de cada artículo. La literatura gris fue excluida ya que no es revisada por pares.

3.4. Selección de estudios primarios

Todos los autores examinaron cada estudio que se recuperó de la búsqueda automatizada para determinar su inclusión en este estudio de mapeo sistemático. El proceso de evaluación incluyó una revisión de todo el documento: título, resumen, introducción, antecedentes, estado del arte, metodología, caso de estudio, resultados y conclusiones. Uno de

Estas secciones deben cumplir los siguientes criterios de inclusión: *El estudio debe informar al menos una evaluación de usabilidad aplicada a un producto de software en el contexto de un proceso de desarrollo.*

Del mismo modo, establecimos los criterios de exclusión. Los estudios que cumplieron al menos una de las siguientes condiciones fueron excluidos de esta investigación: (1) *el estudio de usabilidad no se aplica a un producto de software, y 2) El estudio de usabilidad no se realiza en un contexto de desarrollo de software.*

3.5. Extracción de datos

Desarrollamos una plantilla para registrar toda la información relevante sobre cada artículo resultante. El proceso de extracción de datos incluyó la siguiente información: (a) ID del artículo, (b) Título del artículo, (c) Autor (es), (d) Tipo de publicación, (e) Nombre de la conferencia o revista en la que se realizó el estudio presentado, (f) Año de publicación, (g) Fecha de extracción, (h) Base de datos en la que se encontró el estudio.

La búsqueda automatizada de nuestra revisión sistemática de mapas se realizó el 20 de junio, 2015. Obtuvimos 1169 estudios de las tres bases de datos consultadas. Después de la aplicación de los criterios de inclusión y exclusión, se seleccionaron 215 de estos documentos para el proceso de revisión. La Tabla 2 muestra los detalles con respecto a la cantidad de estudios que se encontraron durante el proceso de búsqueda.

Tabla 2. Resumen de resultados de búsqueda

Nombre de la base de datos	Resultados de la búsqueda	Papeles duplicados	Papeles relevantes
SCOPUS	488	-	136
ACM	101	75	4 4
ISI (Revistas y Actas)	580	247	75
TOTAL	1169	322	215

4) Análisis de datos y resultados

Para determinar los métodos de evaluación de usabilidad más utilizados, identificamos el número de veces que se informó cada técnica en los documentos relevantes. Todos los resultados se resumen en la Tabla 3.

Tabla 3. Frecuencia de uso de cada método de evaluación de usabilidad

Método de evaluación de usabilidad	Número de veces que se usó el método	Porcentaje (%)
Encuesta / Cuestionario	104	26,26%
Pruebas de usuario	56	14,14%
Evaluación heurística	50	12,63%
Entrevista	41	10,35%
Pruebas de usuario: pensar en voz alta / pensar en voz alta	38	9,60%
Métricas de software / Métricas de usabilidad	19	4,80%
Evaluación automatizada mediante herramienta de software	dieciséis	4,04%
Tutorial cognitivo	11	2,78%
Evaluación de prototipo	11	2,78%
Grupo de enfoque	6 6	1,52%
Verificación de lista de verificación	5 5	1,26%
Lápiz y papel	5 5	1,26%
Otros	34	8,59%
TOTAL	396	100.00%

Tabla 4. Estudios que informan el uso de cada método de evaluación de usabilidad

Método de evaluación de usabilidad	Estudios que informan el uso del método (ver Apéndice A)
Encuesta / Cuestionario	001, 002, 005, 006, 009, 015, 017, 019, 020, 023, 034, 037, 040, 043, 046-1, 046-2, 048, 051, 052, 055, 056, 057, 058, 059, 062, 063, 064, 066, 068, 069, 070, 074, 076, 077, 078, 080, 081, 083, 085, 087, 094, 096, 097, 099, 102, 103, 105, 106, 108, 111, 113, 114, 117, 118, 119, 121, 123, 124-2, 126, 127, 131, 134, 136, 137, 140, 141, 142, 144, 145, 147, 149, 151, 154, 158, 160, 161, 167, 169, 170, 174, 175, 176, 177, 178, 179, 181, 182, 185, 188, 190, 192, 194, 198, 199, 201, 202, 206-1, 206-2, 207, 208, 209, 211, 213, 214
Pruebas de usuario	007, 027, 032, 034, 043, 047, 048, 050, 055, 056, 057, 067, 070, 072, 073, 076, 077, 081, 085, 088, 089, 094, 101, 102, 114, 123, 124-1, 124-2, 125, 127, 129, 132, 133, 138, 139, 140, 146, 150, 153, 158, 160, 161, 165, 176, 179, 181, 183-1, 183-2, 192, 193, 204, 205, 207, 208, 209, 214
Evaluación heurística	002, 008, 011-1, 011-2, 012-1, 012-2, 024, 031-1, 031-2, 037, 041, 042, 043, 044, 045, 049, 050, 054, 056, 060, 061, 065, 068, 070, 075, 077, 079, 082, 086, 090, 114, 117, 128, 132, 133, 140, 143, 156, 157, 158, 161, 171, 172, 180, 183-1, 195, 205, 206-1, 206-2, 215
Entrevista	003, 007, 014, 015, 023, 028, 030, 036, 038, 039, 046-1, 046-2, 051, 052, 055, 058, 063, 071, 081, 083, 095, 105, 110, 111, 119, 123, 127, 138, 140, 144, 146, 156, 161, 163, 173, 174, 182, 183-2, 186, 204, 207
Pruebas de usuario: pensar en voz alta / pensar en voz alta	002, 003, 004, 005, 014, 019, 028, 029, 036, 037, 039, 040, 046-1, 046-2, 052, 063, 068, 074, 082, 083, 093-1, 093-2, 095, 099, 104, 109, 117, 118, 134, 156, 159, 162, 173, 174, 182, 184, 188, 197
Métricas de software / Métricas de usabilidad	010, 011-1, 016-1, 016-2, 016-3, 022, 026, 070, 077, 099, 100, 108, 113, 114, 124-2, 155, 178, 200, 202
Evaluación automatizada mediante herramienta de software	013, 016-1, 016-2, 016-3, 025, 053, 091, 108, 112, 135, 145, 189, 191, 196-1, 196-2, 196-3
Tutorial cognitivo	008, 031-1, 031-2, 033, 035, 061, 117, 118, 122, 140, 156
Evaluación de prototipo	007, 038, 050, 110, 120, 121, 156, 183-1, 183-2, 190, 193
Grupo de enfoque	015, 050, 074, 107, 111, 183-1
Verificación de lista de verificación	084, 098, 121, 168, 212
Lápiz y papel	037, 039, 144, 165, 192
Inspección de usabilidad basada en perspectiva	130-1, 130-2, 206-1, 206-2
Observación de campo / estudio de campo	095, 105, 111
Registro visual	052, 164, 166
Haga clic en Mapa / Mapa de desplazamiento / Mapa de calor	014, 160
Opinión Minería	021, 148
Proceso de evaluación de usabilidad web	012-1, 012-2
Pensamiento retrospectivo en voz alta	115, 198
Análisis de tareas cognitivas	051, 165
Pautas de usabilidad	187
Clasificación de tarjetas	183-2
Clasificación de tarjetas de Iona	165
Creación de sentido retrospectivo	162
Personas	165
Flujo de trabajo del usuario	165
Jogthrough cognitivo	152
Inspección específica de dominio	172
Evaluación heurística participativa	116
Método de inspección semiótica	210
Método de evaluación de usabilidad y comunicabilidad	092
Tutorial pluralista simplificado	203
Tutorial cognitivo simplificado y simplificado	203
Método de medición de rendimiento musical	018

En la Tabla 4, especificamos los documentos que informaron el uso de cada método de evaluación de usabilidad específico. Sin embargo, algunos estudios describieron el uso de más de un método único. Por esta razón, existe una diferencia entre el número de trabajos primarios que se identificaron y el número de evaluaciones de usabilidad que se encontraron. Para distinguir estos estudios de las referencias, se asignó un número de tres dígitos a cada artículo. Esta lista de estudios se puede encontrar en el Apéndice A. En el caso de los documentos 011, 012, 016, 031, 046, 093,

124, 183, 196 y 206, se consideró más de un producto de software para una evaluación de usabilidad. El número junto a la ID especifica un software diferente que se evaluó en el mismo caso de estudio.

4.1. Métodos de evaluación de usabilidad

En esta sección, detallamos los métodos de evaluación de usabilidad que se encontraron a través de la ejecución de **la presente revisión sistemática de mapeo**. Según Nielsen [14], Zhang *et. al.*, [21], Paz *et. al.*, [9] y Otaiza *et. al.*, [18] estos métodos se pueden definir como:

Encuesta / Cuestionario: Es una lista de ítems del cuestionario que los usuarios representativos deben responder de acuerdo con una escala Likert. Cada declaración de la encuesta está destinada a medir un aspecto de usabilidad particular del sistema de software o una dimensión específica de la satisfacción del usuario.

Pruebas de usuario: Una cantidad representativa de usuarios finales interactúa con el software siguiendo una lista de tareas predefinidas. Las observaciones exhaustivas de estas interacciones humano-sistema permiten la identificación de problemas de usabilidad relacionados con el sistema. Este método de evaluación se aplica comúnmente en un laboratorio de usabilidad cuyo equipo permite el registro de los gestos del usuario y la pantalla de la computadora del usuario para su posterior análisis.

Evaluación heurística: Un grupo de especialistas en usabilidad juzga si cada elemento de diálogo del sistema de software sigue los principios de usabilidad establecidos, llamados "heurísticas".

Entrevista: Tanto el usuario final como el especialista en usabilidad participan en una sesión de discusión sobre la usabilidad de una aplicación de software.

Pruebas de usuario - Pensar en voz alta / Pensar en voz alta: Esta versión de prueba de usuario implica la ejecución del "protocolo de pensamiento en voz alta". Los usuarios tienen que verbalizar sus pensamientos mientras interactúan con el sistema de software. Los supervisores deben alentar a los usuarios finales a expresar sus opiniones durante la actividad. En algunos casos, esta indicación solo se solicita al comienzo de la prueba.

Métrica de usabilidad / Métrica de software: El propósito de este método es establecer mediciones cuantitativas. Las métricas de usabilidad cuantifican la usabilidad de un sistema en cuanto a efectividad, eficiencia y satisfacción. Por lo general, algunas ecuaciones se utilizan para determinar valores numéricos sobre la usabilidad de un sistema. Se requiere la participación de un número representativo de usuarios para generalizar los resultados obtenidos.

Evaluación automatizada mediante herramienta de software / software: Se utiliza una herramienta de software para realizar todas las actividades que se requieren en una evaluación de usabilidad. Dependiendo del tipo de software, esta herramienta puede simular acciones humanas. Otras aplicaciones solo realizan un seguimiento de las actividades del usuario y realizan mediciones basadas en métricas. Además, estos sistemas pueden generar un archivo de registro que puede analizarse después de la prueba.

Tutorial cognitivo: Un especialista en usabilidad simula las acciones de un usuario novato del sistema. Durante esta interacción, el inspector debe identificar posibles problemas de usabilidad.

Evaluación del prototipo: Tanto los usuarios finales como los especialistas en usabilidad participan de una reunión en la que se les pide a los usuarios que expliquen sus expectativas sobre un prototipo en papel o una maqueta.

Grupo de enfoque: Se solicita a un grupo representativo de usuarios finales que participen en una discusión abierta para analizar la interfaz gráfica de un producto de software. En este método, los participantes son libres de escuchar y hablar con otros miembros del grupo. De esta forma, pueden desarrollar ideas propias basadas en comentarios anteriores.

Verificación de la lista de verificación: Un especialista en usabilidad verifica si una interfaz gráfica de usuario cumple con una serie de especificaciones de diseño bien definidas. Una lista de verificación de verificación ayuda a los inspectores a administrar todos los detalles de usabilidad que deben considerarse en un producto de software en particular. Esta lista

Lápiz y papel: Los usuarios evalúan aspectos de un prototipo en papel. Son libres de modificar el diseño de la interfaz con un lápiz. Además, pueden escribir sus comentarios y hacer anotaciones para especificar sus observaciones en detalle.

Inspección de usabilidad basada en perspectiva: En cada sesión de inspección, el especialista se enfoca en un subconjunto específico de problemas de usabilidad cubiertos por una de varias perspectivas de usabilidad. Cada perspectiva proporciona al inspector una lista de preguntas que representan los problemas de usabilidad para verificar y un procedimiento específico para realizar la inspección. Se supone que con una atención enfocada y un procedimiento bien definido, cada sesión de inspección puede detectar un mayor porcentaje de los problemas relacionados con la perspectiva utilizada, y que la combinación de diferentes perspectivas puede descubrir más problemas que la combinación del mismo número de sesiones de inspección utilizando una técnica de inspección general.

Observación de campo / estudio de campo: Este método implica que un especialista en usabilidad observe el comportamiento natural del usuario en su "hábitat natural", el campo donde se realiza la actividad diaria o el lugar de trabajo donde se implementará el producto de software. El facilitador le da al usuario una tarea y observa, toma notas y hace preguntas mientras el usuario emplea el producto de software para completar la tarea definida. La observación puede ser directa, cuando el inspector está presente durante la tarea, o indirecta, utilizando un software especial para capturar las acciones del usuario en la computadora y grabar la sesión.

Registro visual: Este método implica medir el lugar donde mira el usuario (el punto de mirada) o el movimiento de un ojo durante el uso de un producto de software. Existen varios dispositivos para realizar este tipo de evaluación, tales como: monitores especiales, cámaras específicas, sensores e incluso software especializado. Al analizar la ruta visual de los usuarios finales a través de la interfaz, es posible determinar la información relevante, las secciones que se ignoran, el contenido que se pasa por alto cualquier otra pregunta relacionada con la mirada.

Haga clic en Mapa / Mapa de desplazamiento / Mapa de calor: Clickmaps muestra dónde los usuarios hacen clic en una interfaz de software. Esta información permite a los inspectores identificar las secciones más populares y ver qué secciones confunden los usuarios con los enlaces. Este mapa a menudo está representado por colores que indican la cantidad de clics en un área específica. Se puede obtener un mapa de clics mediante el uso de herramientas de software especiales.

Opinión Minería: Este método se refiere al uso de procesamiento de lenguaje natural y análisis de texto para identificar información subjetiva sobre la usabilidad de un producto de software. Para este propósito, un grupo representativo de usuarios tiene que escribir su opinión sobre cierto software en tres factores de usabilidad: efectividad, eficiencia y satisfacción. Luego, estos comentarios se analizan utilizando técnicas especializadas de Ciencias de la Computación para determinar qué tan positivos o negativos son en cada categoría [1].

Proceso de evaluación de usabilidad web: Este método implica la descomposición del concepto de usabilidad en subcaracterísticas y atributos medibles, que luego son

asociado con las métricas para cuantificarlas numéricamente. Esta técnica ha sido desarrollada especialmente para aplicaciones web. El propósito es proporcionar comentarios durante todas las fases del proceso de desarrollo de software. Los autores de este método proporcionan un modelo completo, que incluye todos los atributos de subcaracterísticas y sus métricas asociadas [3].

Pensamiento retrospectivo en voz alta: Este método es otra variante de las pruebas de usuario. Es una práctica similar al protocolo de pensamiento en voz alta, sin embargo, en este método, los usuarios tienen que verbalizar sus pensamientos después de las actividades de sesión de prueba del usuario, en lugar de durante ellos. Se solicita a los usuarios que utilicen el sistema y realicen ciertas tareas en silencio. Los participantes verbalizan sus pensamientos después mientras miran una grabación de su actuación [20].

Análisis de tareas cognitivas: Esta técnica implica el proceso de aprender sobre usuarios comunes observando su interacción con un producto de software específico para comprender en detalle cómo realizan sus tareas y alcanzan los objetivos previstos. El análisis de tareas ayuda a identificar las tareas que una aplicación de software debe soportar y también puede ayudarlo a refinar la navegación o la búsqueda. Este método se enfoca en comprender tareas que requieren toma de decisiones, resolución de problemas, memoria, atención y juicio.

Pautas de usabilidad: Un grupo de especialistas tiene que evaluar la interfaz gráfica de un producto de software de acuerdo con las pautas de usabilidad predefinidas. Aunque esta técnica es similar a la evaluación heurística, el procedimiento es diferente. En esta técnica, cada inspector puede trabajar individualmente. No es necesario calificar la gravedad y la importancia crítica de cada problema de usabilidad. La herramienta de evaluación no es necesariamente un conjunto de heurísticas de usabilidad. Los inspectores pueden incluso usar pautas proporcionadas por la empresa de desarrollo de software.

Clasificación de tarjetas: Este método puede usarse para verificar la organización y la estructura de la información que aparece en una aplicación de software. Para este tipo de evaluación, se requieren algunas tarjetas de papel. Cada tarjeta debe contener una palabra o frase escrita en un lado. Esta expresión tiene que representar un concepto específico que se considera parte de la interfaz gráfica de usuario. Los participantes reciben una pila de tarjetas y se les pide que las agrupen, ya que tiene sentido para ellas. Organizan los temas en categorías y también pueden ayudar a etiquetar estos grupos. Si se hace visible una taxonomía aceptada y estandarizada, sería apropiado aplicar esa taxonomía en la interfaz.

Clasificación de tarjetas de lona: Esta técnica es una variación de la clasificación de cartas clásica. Este método requiere que los usuarios seleccionen los conceptos más valiosos y los organicen en una plantilla predefinida. En esta versión, las categorías principales están establecidas previamente, y los usuarios solo tienen que colocar cada tarjeta en uno de los grupos.

Creación de sentido retrospectivo: Este método se basa en un protocolo retrospectivo, en el que se pide a los usuarios que verbalicen sus pensamientos después de completar un conjunto de tareas. Esta técnica específica establece el uso de preguntas abiertas para alentar a los usuarios a procesar información de la memoria a largo plazo, proporcionando justificaciones y explicaciones de ciertas acciones que realizaron durante su interacción con una interfaz. Las preguntas deben orientarse para analizar el proceso cognitivo a través del cual las personas experimentan problemas y eligen realizar ciertas acciones, entre otras alternativas, para resolver los problemas experimentados en un momento específico [16].

Personas: Este método implica la descripción de diferentes usuarios ficticios de la aplicación de software. Estas representaciones deben incluir un breve perfil de objetivos y características que representen las necesidades de un grupo más grande de usuarios reales. La evaluación implica un análisis de la interfaz gráfica de usuario considerando las metas, posibles comportamientos, actitudes, motivaciones y objetivos comerciales de cada perfil.

Flujo de trabajo del usuario: Este método establece la elaboración de diagramas para representar todas las rutas que están disponibles en un sistema de software para realizar una tarea específica. Este diagrama permite a los especialistas analizar el logro de múltiples objetivos que involucran muchas subtareas. Además, permite examinar las preferencias de los diferentes usuarios y el orden en que se realizan ciertas tareas.

Jogthrough cognitivo: Este método es una versión alternativa del tutorial cognitivo. En esta versión, mientras los inspectores trabajan en una serie de tareas, se hacen una serie de preguntas desde la perspectiva del usuario. Las respuestas a estas preguntas deben clasificarse según el porcentaje de usuarios potenciales que se espera que tengan problemas (de 0 a 3 en una escala Likert) [7].

Inspección específica de dominio: Este método implica el uso de un modelo [17] que se puede adaptar a cualquier dominio de software. Los especialistas deben determinar las áreas y atributos que son más relevantes para el software que van a evaluar. La inspección debe realizarse de acuerdo con las pautas establecidas para cada atributo de usabilidad.

Evaluación heurística participativa: es una extensión de la evaluación heurística tradicional donde se consideran algunos principios para evaluar la interfaz gráfica de usuario. La evaluación heurística participativa utiliza la misma técnica. Sin embargo, implica la participación de los usuarios finales en el proceso de evaluación como "inspectores expertos de dominio". Además, se agregan algunas heurísticas adicionales para incluir algunos aspectos de usabilidad que no son considerados por la propuesta tradicional de Nielsen.

Método de inspección semiótica: El propósito de este método es el análisis de los mensajes transmitidos a través de la metacomunicación de diseñador a usuario. Estos mensajes se expresan con una amplia gama de signos y símbolos en la interfaz, desde uno o más sistemas de significación. El objetivo del método de inspección semiótica es la evaluación de estos elementos, buscando problemas reales o potenciales de comunicación y oportunidades de rediseño para mejorar la comunicación [6].

Método de evaluación de usabilidad y comunicabilidad: En este método, los evaluadores tienen que identificar fallas de comunicación mientras una cantidad representativa de usuarios interactúa con el software del producto. Hay trece expresiones de desglose de comunicación o etiquetas para clasificar los problemas de comunicabilidad y usabilidad. El evaluador debe interpretar estos problemas y reconstruir el mensaje para identificar posibles mejoras.

Tutorial pluralista simplificado: Los usuarios y diseñadores participan juntos en una reunión para evaluar nuevas ideas con respecto a la interfaz gráfica de usuario de un producto de software. El método no requiere un prototipo funcional. Pueden desarrollar un diseño a partir de solo ideas. Los diseñadores del sistema pueden obtener información valiosa sobre las tareas de los usuarios además de los comentarios sobre el diseño.

Tutorial cognitivo simplificado simplificado: Este método establece el mismo procedimiento que el trote cognitivo. La diferencia es que los evaluadores solo requerían hacer dos preguntas en cada paso de la inspección [19]. Además, implicaba elaborar menos documentación.

Método de medición de rendimiento musical: Este método establece que la usabilidad de un producto se mide por el grado en que usuarios específicos logran objetivos específicos en un entorno específico. Se utilizan algunas métricas para determinar datos cualitativos con respecto a la usabilidad. La técnica indica que los experimentos controlados deben realizarse tan cerca de un entorno de trabajo real. Se puede utilizar una herramienta de software llamada DRUM para analizar archivos de registro.

4.2. Resultados por dominio de software

En esta sección, detallamos los principales dominios de las aplicaciones de software que se informaron en los estudios primarios. Los resultados se resumen en la Tabla 5.

Tabla 5. Dominio principal del software involucrado en los estudios

Dominio de software	Número de aplicaciones de software que fueron evaluado	Porcentaje
Informática de salud	50	20,33%
Educación	36	14,63%
Herramientas de desarrollo de software	23	9,35%
Comercio electrónico	20	8,13%
Juego de azar	19	7,72%
Gerente de información personal	9 9	3,66%
Rutas y Guías del Sistema	6 6	2,44%
Sistemas de uso general	6 6	2,44%
Sistemas expertos	5 5	2,03%
Sistemas de Información Geográfica	5 5	2,03%
Software para la comunicación de los usuarios	5 5	2,03%
Planificación de recursos empresariales	4 4	1,63%
Herramientas y utilidades	4 4	1,63%
Otros dominios	39	15,85%
No hay información sobre el software.	15	6,10%
TOTAL	246	100.00%

La mayoría de los sistemas estaban relacionados con *Informática de salud*. En esta categoría se consideraron: aplicaciones para la salud, sistemas de gestión hospitalaria, sistemas de soporte de decisiones clínicas, sistemas electrónicos de registro de pacientes, software para dispositivos médicos especializados, *etc.*

En *Educación*, Se consideraron todo tipo de productos de software que se utilizan para apoyar actividades de aprendizaje y enseñanza, como plataformas de E-Learning, herramientas de enseñanza y sistemas de gestión para universidades.

En la categoría de *Herramientas de desarrollo de software*, revisamos todas las aplicaciones que se utilizan para respaldar cualquier actividad del proceso de desarrollo de software. Algunos ejemplos incluyen IDE, herramientas CASE, software para realizar casos de prueba, *etc.*

En el comercio electrónico, consideramos todo el software relacionado con la venta de productos y servicios. Esta categoría abarcaba la mayoría de las aplicaciones web transaccionales.

Gerente de información personal está relacionado con el software que adquiere, organiza, mantiene y recupera información personal. Algunos ejemplos incluyen aplicaciones de correo electrónico, software para administrar tarjetas de crédito, calendarios web, *etc.*

La categoría de *Rutas y Guías del Sistema* incluye todo el software que proporciona información sobre las rutas disponibles de autobuses y trenes, incluidos horarios y mapas. En este dominio, también consideramos aplicaciones que guían a los usuarios a ubicaciones específicas de acuerdo con la información de tráfico en tiempo real.

Sistemas de uso general están relacionados con aplicaciones web cuyo propósito es solo mostrar información. Por ejemplo, el sitio web de algunas empresas muestra información relacionada con los productos y servicios que ofrecen, sin embargo, no proporcionan mecanismos para comprar en línea. Los motores de búsqueda también se consideraron en esta categoría.

En *Sistemas expertos* fueron considerados sistemas informáticos que emulan la capacidad de toma de decisiones de un experto humano. La mayoría de estas aplicaciones estaban relacionadas con sistemas médicos expertos que podían diagnosticar una enfermedad y dar una receta.

En *Sistema de información geográfica*, consideramos todas las aplicaciones que ofrecen información sobre diferentes lugares de una ciudad en función de la ubicación actual del usuario. Del mismo modo, la categoría de *Software para la comunicación de los usuarios* incluye aplicaciones móviles y web que permiten la comunicación en tiempo real.

En la categoría de *Planificación de recursos empresariales*, Se consideraron aplicaciones de software que son capaces de recopilar, almacenar, administrar e interpretar datos de cualquier actividad comercial.

Finalmente, el dominio de *Herramientas y utilidades* incluye aplicaciones proporcionadas por el sistema operativo.

La categoría "otros dominios" incluye: *Sistemas de gestión de la relación con el cliente, Sistemas de información de recursos humanos, Sistemas de soporte de decisiones, software para astronomía y ciencia espacial, Software financiero, Herramientas de gestión de proyectos, redes sociales, Sistemas de software bancario, Real- Sistemas de localización de hora, sistemas operativos, software para músicos y procesamiento de audio, Software para medir aspectos psicológicos, software de biología molecular, Sistemas en línea de radio y televisión, Gobierno y soluciones de software del sector público, Software para el sector del transporte, Software para Procesos de Negocio, Software para fábricas industriales, Software para militares, seguridad y defensa, software de gestión de la cadena de suministro, Software para el sector marítimo, Software para aeronaves, Software para física, Software para biología, editores de fotos, Software controlado por Brain*

y *Software matemático y estadístico.*

Tabla 6. Principales métodos de evaluación de usabilidad por dominio de software

Dominio de software	Método utilizado para evaluar la usabilidad.	Número de veces que se usó el método	Porcentaje (%)
Informática de salud	Encuesta / Cuestionario	28	30,43%
	Pruebas de usuario: pensar en voz alta	dieciséis	17,39%
	Pruebas de usuario	13	14,13%
	Evaluación heurística	10	10,87%
	Otros	25	27,17%
Educación	Encuesta / Cuestionario	19	28,79%
	Pruebas de usuario	10	15,15%
	Entrevista	10	15,15%
	Evaluación heurística	7 7	10,61%
	Otros	20	30,30%
Desarrollo de software Herramientas	Encuesta / Cuestionario	10	28,57%
	Pruebas de usuario	7 7	20,00%
	Entrevista	5 5	14,29%
	Pruebas de usuario: pensar en voz alta	4 4	11,43%
	Otros	9 9	25,71%

En la Tabla 6 y la Tabla 7, presentamos los métodos más utilizados para cada dominio. A pesar de las diferencias, la encuesta es una de las más utilizadas en todos los dominios de software.

Tabla 7. Principales métodos de evaluación de usabilidad por dominio de software

Dominio de software	Método utilizado para evaluar la usabilidad.	Número de veces que se usó el método	Porcentaje (%)
Comercio electrónico	Evaluación heurística	9 9	25,00%
	Encuesta / Cuestionario	8	22,22%
	Pruebas de usuario	5 5	13,89%
	Métrica de software	3	8,33%
	Otros	11	30,56%
Juego de azar	Encuesta / Cuestionario	8	27,59%
	Evaluación heurística	6 6	20,69%
	Pruebas de usuario	3	10,34%
	Pruebas de usuario: pensar en voz alta	3	10,34%
	Otros	9 9	31,03%
Información personal Gerente	Encuesta / Cuestionario	3	15,79%
	Evaluación heurística	3	15,79%
	Pruebas de usuario	3	15,79%
	Métrica de software	3	15,79%
	Otros	7 7	36,84%
Rutas y sistema Guías	Encuesta / Cuestionario	3	25,00%
	Entrevista	2	16,67%
	Evaluación de prototipo	2	16,67%
	Verificación de lista de verificación	2	16,67%
	Otros	3	25,00%

Propósito general Sistemas	Encuesta / Cuestionario	4 4	36,36%
	Pruebas de usuario	2	18,18%
	Pruebas de usuario: pensar en voz alta	2	18,18%
	Otros	3	27,27%
Sistemas expertos	Encuesta / Cuestionario	4 4	80,00%
	Pruebas de usuario: pensar en voz alta	1	20,00%
	Otros	0 0	00,00%
Sistemas de Información Geográfica	Pruebas de usuario: pensar en voz alta	2	28,57%
	Otros	5 5	71,43%
Software para la comunicación de los usuarios	Evaluación heurística	2	25,00%
	Pruebas de usuario	2	25,00%
	Pruebas de usuario: pensar en voz alta	2	25,00%
	Otros	2	25,00%
Recursos empresariales Planificación	Encuesta / Cuestionario	2	22,22%
	Evaluación heurística	2	22,22%
	Métrica de software	2	22,22%
	Otros	3	33,33%
Herramientas y utilidades	Encuesta / Cuestionario	2	28,57%
	Otros	5 5	71,43%

4.3. Resultados por tipo de aplicación

Para determinar las tendencias actuales por tipo de aplicación, desarrollamos un mapeo (Figura 01) para representar la cantidad de veces que se usaron los métodos más populares en estos contextos particulares. Es posible determinar que la tendencia, que se estableció en la Sección 4, continúa. El cuestionario es la técnica más utilizada debido a su simplicidad técnica. No existe una diferencia significativa entre las pruebas de usuario y la evaluación heurística. Sin embargo, la "evaluación heurística" se informa más que la "prueba de usuario" para el contexto de las aplicaciones móviles. Finalmente, pensar en voz alta es el protocolo menos empleado para estos contextos, excepto los móviles, donde la entrevista obtuvo menos puntaje.

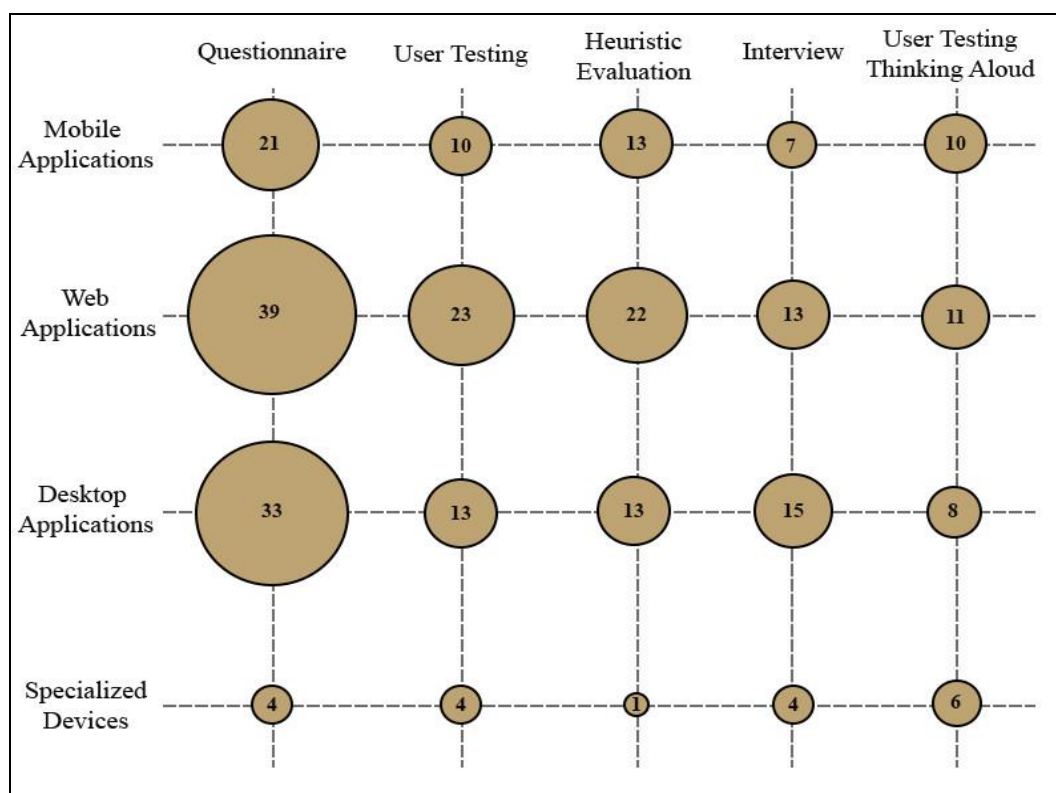


Figura 1. Diagrama categórico de burbujas Acerca de la evaluación de usabilidad principal
Métodos por tipo de aplicación

5) Conclusiones y trabajos futuros

Se han propuesto varios métodos de evaluación para determinar el nivel de usabilidad de las aplicaciones de software. A pesar de esta gran cantidad de métodos, aún queda por determinar la técnica más adecuada para un escenario particular. En este estudio, realizamos una revisión sistemática de mapeo del uso de métodos de evaluaciones de usabilidad para procesos de desarrollo de software. Nuestro estudio está destinado a servir como contribución para apoyar la toma de decisiones en la elección de una técnica.

Siguiendo un protocolo predefinido, identificamos 1169 estudios, de los cuales se seleccionaron 215. Este trabajo permitió determinar que: (1) cuestionario, (2) pruebas de usuario, (3) evaluación heurística, (4) entrevista y (5) protocolo de pensamiento en voz alta son las técnicas más empleadas según la literatura. Además, en este estudio hemos determinado los tipos de aplicaciones que se informan con frecuencia en la literatura como parte de una evaluación de usabilidad en desarrollos de software. La mayor parte de la aplicación pertenece a las categorías de (1) Informática de salud, (2) Educación, (3) Desarrollo de software, (4) Comercio electrónico y (5) Juegos.

Del análisis, observamos además que algunas técnicas están adaptadas para cubrir todos los aspectos de usabilidad en algunos tipos de productos de software. La aparición de categorías híbridas ha obligado a los académicos a proponer herramientas de evaluación particulares, como cuestionarios de usabilidad para dominios específicos, heurística para un tipo particular de software, variantes de un método de usabilidad, *etc.* También es necesario un análisis más profundo en cada categoría, especialmente en la metodología, estableciendo las diferencias y cómo afectan el resultado final.

Apéndice A. Documentos identificados en el estudio de mapeo sistemático

Esta información está disponible en:

http://inform.pucp.edu.pe/~jpowsang/usability/mapping_study_appendix.htm

Referencias

- [1] AM El-Halees, "Evaluación de usabilidad de software utilizando la minería de opinión", *Journal of Software*, vol. 9, no. 2, (2014), pp. 343-349.
- [2] A. Fernández, E. Insfran y S. Abrahão, "Métodos de evaluación de usabilidad para la web: un estudio de mapeo sistemático", *Tecnología de la información y el software*, vol. 53, no. 8, (2011), pp. 789-817. [3] A. Fernández, S. Abrahão y E. Insfran, "Un proceso de evaluación de usabilidad web para el desarrollo web basado en modelos", *Actas de la 23ª Conferencia Internacional sobre Ingeniería de Sistemas de Información Avanzada, Londres, Reino Unido, (2011) 20-24 de junio.* [4] A. Holzinger, "Métodos de ingeniería de usabilidad para desarrolladores de software", *Communications ACM*, vol. 48, no. 1, (2005) pp.71-74. [5] B. Kitchenham y S. Charters, "Directrices para realizar revisiones sistemáticas de literatura en ingeniería de software", *Keele University and Durham University Joint Report, Tech. Rep., (2007).*
- [6] CS de Souza, CF Leitão, RO Prates y EJ da Silva, "El método de inspección semiótica", *Actas del VII simposio brasileño sobre factores humanos en sistemas informáticos, Natal, Brasil, (2006) 19-22 de noviembre.* [7] DE Rowley y DG Rhoades, "The Cognitive Jogthrough", *Actas de la Conferencia SIGCHI sobre factores humanos en los sistemas de computación, California, EE. UU., (1992) 3-7 de mayo.* [8] F. Gündüz y AK Pathan, "Sobre los factores clave de usabilidad en aplicaciones móviles de pantalla táctil de pequeño tamaño", *International Journal of Multimedia and Ubiquitous Engineering*, vol. 8, no.3, (2013), pp. 115-138. [9] F. Paz y JA Pow-Sang, "Tendencias actuales en los métodos de evaluación de usabilidad: una revisión sistemática", *Actas de la 8ª Conferencia Internacional sobre Ingeniería de Software Avanzada y sus aplicaciones, Hainan, China, (2014) 20-23 de diciembre.* [10] F. Paz y JA Pow-Sang, "Métodos de evaluación de usabilidad para el desarrollo de software: un estudio de mapeo sistemático", *Actas de la 8ª Conferencia internacional sobre ingeniería avanzada de software y sus aplicaciones, Isla de Jeju, Corea, (2015) 25-28 de noviembre.* [11] F. Paz, D. Villanueva, C. Rusu, S. Roncagliolo y JA Pow-Sang, "Evaluación experimental de la heurística de usabilidad", *Actas de la Décima Conferencia Internacional sobre Tecnología de la Información: Nuevas Generaciones, Las Vegas, NV, EE. UU., (2013) 15-17 de abril.*

- [12] ISO, "Requisitos ergonómicos para el trabajo de oficina con terminales de pantalla visual (vdts), - Parte 11: Orientación de usabilidad ", Organización Internacional de Normalización, Ginebra, Suiza, ISO 924111: 1998, (1998).
- [13] ISO, "Ingeniería de software - Calidad del producto - Parte 1: Modelo de calidad", Organización Internacional para Normalización, Ginebra, Suiza, ISO 9126-1: 2001, (2001).
- [14] J. Nielsen, "Ingeniería de usabilidad", Academic Press, San Diego, CA, EE. UU., (1993).
- [15] N. Vatankhah, KT Wei y S. Letchmunan, "Medición de usabilidad de sitios web de turismo en línea de Malasia", *Revista Internacional de Ingeniería de Software y sus aplicaciones*, vol. 8, no. 12, (2014), pp. 1-18. [dieciséis] P. Balatsoukas, J. Ainsworth, R. Williams, E. Carruthers, C. Davies, J. McGrath, A. Akbarov, C. Soiland-Reyes, S. Badiyani e I. Buchan, "Protocolos verbales para evaluar la usabilidad de Apoyo a la decisión clínica: el protocolo de toma de sentido retrospectivo ", *Tecnología e informática*, vol. 192, (2013) pp. 283-287. [17] R. AlRoobaea, AH Al-Badi y PJ Mayhew, "Generación de un método de evaluación de inspección específica de dominio a través de un marco adaptativo", *Revista Internacional de Informática Avanzada y Aplicaciones*, vol. 4, no. 6, (2013), pp. 72-91. [18] R. Otaiza, "Evaluación de la usabilidad de los sitios web transaccionales", Tercera Conferencia Internacional sobre Avances en Interacciones Computadora-Humanos, Saint Maarten, (2010) 10-15 de febrero. [19] R. Spencer, "El método de tutorial cognitivo simplificado, evitando las restricciones sociales encontradas en una empresa de desarrollo de software", *Actas de la Conferencia SIGCHI sobre factores humanos en sistemas informáticos*, California, EE. UU., (1992) 3-7 de mayo. [20] S. Eling, L. Lentz y M. de Jong, "Método retrospectivo de pensar en voz alta: usar los movimientos oculares como una señal adicional para las verbalizaciones de los participantes", *Actas de la Conferencia SIGCHI sobre Factores Humanos en Sistemas de Computación*, Vancouver, Canadá, (2011) 7-12 de mayo. [21] Z. Zhang, V. Basili y B. Shneiderman, "Inspección de usabilidad basada en perspectiva: una validación empírica de la eficacia", *Empirical Software Engineering*, vol. 4, (1999) 43-69.

Autores



Freddy Paz es profesor a tiempo parcial en *Pontificia Universidad Católica del Perú* (PUCP). Sus intereses de investigación incluyen la interacción de la computadora humana y los lenguajes de programación. Recibió su Maestría en Informática e Ingeniería Informática de la PUCP - Perú, y *Pontificia Universidad Católica de Valparaíso* (PUCV) - Chile. Actualmente es estudiante de doctorado en Ingeniería. Tiene una licenciatura en Ingeniería de Sistemas de *Universidad Nacional Pedro Ruiz Gallo* (UNPRG), Perú.



José Antonio Pow-Sang es profesor titular y director ejecutivo de la Escuela de Postgrado en *Pontificia Universidad Católica del Perú* (PUCP). Sus intereses de investigación incluir empírica software ingeniería, métricas de software, educación en ingeniería de software e interacción humano-computadora. El es Ph.D. en Ingeniería Informática, y tiene una Maestría en Ingeniería de Software de la Universidad Técnica de Madrid, España. Tiene una licenciatura en Ingeniería Informática de la PUCP. Es miembro senior de IEEE Computer Society y miembro de ACM.