Procedimiento para realizar pruebas de usabilidad

Procedure for performing usability test

Roexcy Vega Prieto [1*], Zaylí Rodríguez Luis, Yaneisi Ofelia Justo Morell

[1] Facultad 6. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños, km 2½, Torrens, Boyeros, La Habana, Cuba. CP.: 19370

*Autor para la correspondencia: rprieto@uci.cu

Resumen:

El desarrollo que ha alcanzado la industria del software a nivel mundial propicia la competencia entre todo tipo de producto informático. En la búsqueda de una mayor calidad, la usabilidad se ha convertido en una necesidad y obligación para las empresas dedicadas a este fin, el usuario final es el que le da o no importancia al software y se basa mayormente en su facilidad de uso. Con el objetivo de obtener una mejor interacción entre los clientes y desarrolladores surge la necesidad de realizar la investigación que lleva por título Procedimiento para la realización de pruebas de usabilidad de software. Con la propuesta también se persigue desarrollar habilidades que contribuyan a la planificación de las actividades para realizar las mediciones de la usabilidad. En la investigación se tuvieron en cuenta los métodos existentes para la evaluación de la usabilidad, las opiniones dadas por expertos sobre sus buenas prácticas y los requisitos del sistema informático a evaluar. Teniéndose como base los aspectos descritos anteriormente, se define un procedimiento para realizar pruebas de usabilidad.

Palabras clave: calidad, calidad de software, usabilidad de software, prueba

Abstract:

The development achieved by the software industry worldwide, encouraging competition among all types of software product. In search of a higher quality, usability has become a necessity and obligation for companies dedicated to this purpose, at the end user is the one that gives importance to the software or not and is largely based on its ease of use . In order to obtain a better interaction between customers and developers arrises the need for research entitled Procedure for usability testing of software. The proposal also seeks to develop skills that will help in the planning of activities for measurement of usability research took into account the existing methods for usability evaluation, the opinions given by experts on their best practices and computer system requirements to evaluate. Taking as basis the aspects described above, defines a procedure for testing usability.

Keywords: quality, software quality, usability of software, test

---

1. Introducción

Desde el surgimiento de las tecnologías de la información y las comunicaciones, el ser humano no descansa de buscar alternativas para darle solución a los problemas esenciales de la sociedad, lo cual trae consigo el desarrollo

tecnológico alcanzado a escala mundial en la industria del software. La competencia ha provocado un desborde de técnicas y metodologías en aras de producir un software cada vez más completo.

Esta evolución mundial ayuda a comprender de dónde proviene la necesidad de ofrecer una mayor calidad del producto o servicio que se le proporciona al consumidor y, en definitiva, a la sociedad, y cómo poco a poco se han ido involucrando todas las organizaciones en la consecución de este fin, para de esta forma lograr la satisfacción del cliente.

La calidad no se ha convertido únicamente en uno de los requisitos esenciales del producto, sino que es un factor estratégico del que dependen la mayor parte de las instituciones, no sólo para mantener su posición en el mercado sino incluso para asegurar su supervivencia.

En la producción de software existen varios métodos para medir la calidad, unos comienzan desde sus inicios y otro en sus fases finales. La ventaja que puede tener uno sobre otro está dada por la profundidad de los errores que se pueden encontrar. En un software donde se aplican estos métodos desde que comienza, se van corrigiendo dificultades con cada etapa, si es verificado al final puede ser muy costoso encontrar la raíz de los problemas.

Calidad de software ha sido definida como la concordancia con los requisitos funcionales y de rendimiento explícitamente establecidos con los estándares de desarrollo documentados y con las características implícitas que se esperan de todo software desarrollado profesionalmente (Pressman, 2005). Carrasco la definió como un conjunto de cualidades que lo caracterizan y que determinan su utilidad y existencia (Scalone, 2006). La calidad de software es sinónimo de eficiencia, flexibilidad, corrección, confiabilidad, mantenibilidad, portabilidad, usabilidad, seguridad e integridad. De acuerdo con el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) es el grado con el que un sistema, componente o proceso cumple los requerimientos especificados y las necesidades o expectativas del cliente o usuario (Montañez, 2007).

En el desarrollo de software, la calidad de diseño acompaña a la calidad de los requisitos y especificaciones. La concordancia es un aspecto centrado principalmente en la implementación, la cual si sigue al diseño y el sistema resultante cumple con los objetivos de requisitos y rendimiento, la calidad de concordancia es alta. Entonces a grandes rasgos la calidad de software es el grado con el cual el usuario percibe que el software cumple con sus expectativas.

Durante el ciclo de desarrollo del software se hace necesario darle seguimiento a una serie de parámetros para verificar que el mismo se encuentre libre de errores. En todos los equipos de desarrollo deben existir personas encargadas de ello, las cuales se guían por un plan de pruebas que permite velar por la calidad del software.

Las pruebas es una actividad en la cual un sistema o componente es ejecutado bajo unas condiciones o requerimientos específicos, los resultados son observados y registrados, y una evaluación es hecha de algún aspecto del sistema o componente (Prado, 2007).

La prueba de un sistema se define como el proceso de ejercitar o evaluar el sistema, por medios manuales o automáticos, para verificar que satisface los requerimientos o, para identificar diferencias entre los resultados esperados y los que producen el sistema (Rojas, 2007). Después de analizadas las fuentes bibliográficas se puede definir que.las pruebas de software son la verificación constante del comportamiento del producto a partir de un conjunto de casos de prueba, las cuales se ejecutan para comprobar la satisfacción de los requerimientos solicitados por el cliente.

Las pruebas se encuentran definidas dentro de niveles específicos, donde se figuran los métodos de cada uno de ellos y sus objetivos como tal. A continuación se mencionan los niveles de pruebas que existen: unitarias, de integración, de sistema, de liberación, de aceptación y de regresión.

Las pruebas de sistema tienen en cuenta la usabilidad con la que se desarrollan los sistemas en las cuales se hace énfasis en que el producto tenga una buena facilidad de aprendizaje, brinde un alto nivel de productividad, el número de errores (por parte del usuario) sea prácticamente nulo; todo enfocado en la necesaria satisfacción y aceptación del cliente.

2. Pruebas de usabilidad

Después de definir el concepto de prueba, se hace necesario recalcar que uno de los objetivos de este trabajo es la búsqueda de formas adecuadas para mejorar la satisfacción del cliente. Las pruebas más relacionadas con este tema son las de usabilidad, las cuales son técnicas formales que tienen como objetivo estudiar la usabilidad de una aplicación en un entorno real con usuarios reales (ISO 9241, 2006). El material usado puede ser el plan de evaluación, la lista de tareas, consentimiento de grabación, entre otros. Son una forma de medir el número de errores que puede una persona presentar al usar un objeto hecho por el hombre, como puede ser una página web, una interfaz de usuario, un documento o un dispositivo.

Usabilidad se define como el grado en el que un producto puede ser utilizado por usuarios para conseguir objetivos específicos con efectividad, eficiencia y satisfacción en un determinado contexto de uso. Como se puede comprobar, en esta definición se liga la usabilidad de un sistema a usuarios, necesidades y condiciones específicas. Por tanto, no es un atributo inherente al software, no puede especificarse independientemente del entorno de uso y de los usuarios concretos que vayan a utilizar el sistema. La ingeniería de usabilidad se puede definir como una aproximación al desarrollo de sistemas en la que se especifican niveles cuantitativos de usabilidad, y el sistema se construye para alcanzar dichos niveles, que se conocen como métricas. Proporciona un modo práctico de asegurar que el software desarrollado alcanza un cierto nivel de usabilidad y está basada en la evaluación mediante pruebas de usabilidad con usuarios.

La usabilidad no puede definirse como un atributo simple de un sistema informático, pues implicará aspectos distintos dependiendo del tipo de producto a construir.

Descripción de procesos de evaluación de la usabilidad

En la industria del software, la usabilidad juega un papel importante a la hora de brindar un servicio dado y se asumía que era una propiedad exclusiva de la presentación de la información. Con el paso del tiempo y la realización de estudios sobre el tema se demostró que se estaba cometiendo un error ya que realmente toda la atención va dirigida a los objetivos que los usuarios quieren alcanzar con sus tareas, siendo entonces los usuarios quienes juegan el papel fundamental. Existen varias técnicas para medir la usabilidad de los software, entre las que se encuentran:

1- Entrevistas individuales (Usability, 2010):

Las entrevistas individuales se refieren normalmente a hablar con un usuario a la vez cara a cara, por teléfono, o con la mensajería instantánea u otros medios asistidos por ordenador. Estas entrevistas no tienen que ver con mirar una obra de usuario. Por lo tanto, esto es diferente de entrevistar a los usuarios en una sesión de pruebas de usabilidad o la realización de entrevistas contextuales, ya que pueden dar una profunda comprensión de la gente que viene a su sitio. Se pueden probar sus actitudes, creencias, deseos y experiencias. ¿Cuándo y cómo realizar una entrevista? Una técnica, es el uso de entrevistas individuales para complementar las encuestas en línea. Se pueden hacer las primeras para refinar las preguntas de la encuesta. O se pueden hacer después de un estudio para investigar los detalles y las razones detrás de las respuestas que dan los usuarios en una encuesta. Los primeros pasos para la realización de este tipo de pruebas es la selección de los participantes y de los temas sobre los que se necesita información. Luego se

define un protocolo de entrevista para guiar al entrevistador, donde se incluyen las preguntas para el seguimiento. Las sesiones se pueden grabar y tener uno o más tomadores de notas.

2- Grupo focal (Usability, 2010):

Un grupo focal es una reunión entre 8 y 12 usuarios o posibles usuarios de un sitio, que tienen un debate moderado sobre el mismo. Normalmente dura unas dos horas y cubre una amplia gama de temas programados de antemano. Son una técnica de investigación de mercados tradicionales, por lo que los departamentos de marketing suelen estar más familiarizados con los grupos de enfoque que con las pruebas de usabilidad o entrevistas contextuales. En un grupo de enfoque típico, los participantes hablan acerca de su trabajo. En este tipo de técnica se trabaja en la búsqueda de información acerca de las actitudes del usuario, las creencias, los deseos y sus reacciones al mostrarles diferentes ideas o prototipos. En la realización de grupos focales se seleccionan los participantes de tal forma que estos sean representantes de usuarios con interés de usar la aplicación. También es necesario tener un plan donde esté plasmado lo que se necesita saber, es decir lo que se quiere lograr con la realización de la prueba, para que sirva de guía al moderador. Este último debe ser una persona capaz de motivar a la participación de los seleccionados en el debate y guiar el mismo a través de una serie de preguntas sin un orden lógico, es decir provoca una discusión fluida y sin problemas. Se debe de tener también una o más personas tomando notas, en dependencia de la cantidad de involucrados en el debate.

3- Evaluación heurística (Usability, 2010):

Una evaluación heurística es un método de inspección de usabilidad de software que ayuda a identificar problemas en la interfaz de usuario. En concreto, supone que los evaluadores examinan y juzgan su conformidad con los principios reconocidos los cuales son llamados también heurísticas de usabilidad. Las evaluaciones heurísticas generalmente se llevan a cabo por un pequeño grupo (de uno a tres) de probadores, los cuales de forma independiente examinan una interfaz de usuario y juzgan el cumplimiento de un conjunto de principios de usabilidad. El resultado de este análisis es una lista de problemas. A pesar de que se requiere un cierto nivel de conocimientos y experiencia para aplicar la heurística eficaz, este método puede proporcionar una retroalimentación rápida y relativamente barata para los diseñadores, donde los comentarios pueden ser obtenidos al inicio del proceso de diseño. La asignación de la heurística puede ayudar a sugerir las mejores medidas correctivas para los diseñadores.

4- Cuestionarios (Usability, 2010):

Existen varios tipos de cuestionarios, actualmente los más usados con fines de usabilidad son los cuestionarios en línea para los cuales se estructuran las entrevistas con los usuarios, donde se muestran una lista de preguntas y respuestas.

Al realizar un cuestionario en línea, se tiene la oportunidad de conocer muchos tipos de información acerca de quiénes son los usuarios de un sitio, cómo lo utilizan, y sus opiniones sobre el mismo. La realización de encuestas se puede llevar a cabo en cualquier etapa del proceso de desarrollo. Una encuesta en línea también puede ser un tema persistente que permite medir continuamente la forma en que los usuarios lo ven e identifican maneras de hacerlo mejor. En la realización de una encuesta se incluyen los siguientes componentes:

- Su objetivo y dónde se puede encontrar respuestas.
- Debe de tener menos de 10 artículos que puedan ser respondidos en un tiempo estimado entre 5 y 10 minutos.

- Se debe considerar la posibilidad de una mezcla de composición abierta (darle la opción a los usuarios de completar la respuesta) y preguntas cerradas (por ejemplo, de opción múltiple o selección de la respuesta).

- También se debe tener una breve encuesta y preguntar si el entrevistado está dispuesto a responder a más preguntas en profundidad en una encuesta de seguimiento. Se debe dar la opción de facilitar una dirección de correo electrónico y enviar la encuesta de seguimiento.

5- Análisis de tareas (Usability, 2010):

El análisis de tareas consiste en aprender acerca de las metas de los usuarios, lo que quiere hacer en un sitio web, cómo funcionan a través de las tareas específicas que hacen para alcanzar sus objetivos y qué medidas toman para llevarlas a cabo. El análisis de tareas complementa la comprensión de usuario, ya que se pueden ver directamente las tareas generales que están tratando de lograr o cómo las realizan actualmente. El análisis puede ayudar a redefinir la navegación o la búsqueda de una mejor asistencia a los usuarios y nuevas metas para crear aplicaciones que coincidan con los usuarios, objetivos, tareas y pasos.

La aplicación de métodos que evalúen el nivel de usabilidad que tiene un sistema ayuda a la búsqueda de la aceptación y éxito de un software, lo cual se debe a que un producto en el cual se han validado satisfactoriamente tanto las necesidades como las facilidades que requiere un usuario para interactuar con él, es más usado y brinda mayor satisfacción a los clientes.

2.4. Atributos de usabilidad

La usabilidad es una cualidad demasiado abstracta como para ser medida directamente. Para poder estudiarla se descompone habitualmente en los siguientes atributos (Nielsen, 2001):

- **Facilidad de aprendizaje:** está dado por la facilidad con que los usuarios puedan realizar las tareas en el sistema.

- **Facilidad de Uso:** facilidad con la que el usuario hace uso de la herramienta, con menos pasos o más naturales a su formación específica. Tiene que ver con la eficacia y eficiencia de la herramienta.

- **Satisfacción:** medida en que los usuarios están satisfechos con los objetivos logrados.

- **Satisfacción subjetiva:** se basa en qué tan placentera es la utilización del sistema para los usuarios.

- **Número de errores por parte de los usuarios:** la aplicación debe ayudar a que el usuario cometa el menor número de errores posibles mientras esté interactuando.

- **Tiempo requerido para realizar una tarea:** los usuarios interactúan con la aplicación en busca de resultados rápidos, por lo que el tiempo para la respuesta del sistema debe ser de un tiempo relativamente corto.

- **Eficacia:** en su interacción con el sistema, el usuario debe tener un alto nivel de productividad.

- **Retención sobre el tiempo:** se basa en qué tanto puede recordar un usuario la realización de una tarea pasado un tiempo de haberla efectuado.

- **Flexibilidad:** variedad de posibilidades con las que el usuario y el sistema pueden intercambiar información. También abarca la posibilidad de diálogo, la multiplicidad de vías para realizar la tarea y la similitud con tareas anteriores.

- **Robustez:** es el nivel de apoyo al usuario que facilita el cumplimiento de sus objetivos.

- **Privacidad:** el usuario debe confiar en que sus datos personales y actividades sólo van a ser visibles para quien ellos elijan.

Algunos de estos atributos no contribuyen a la usabilidad del sistema en la misma dirección, pudiendo ocurrir que el aumento de uno de ellos tenga como efecto la disminución de otro.

La usabilidad del sistema informático no es una simple adición del valor de estos atributos, sino que se define para cada uno como un nivel a alcanzar para algunos de ellos. Estos atributos pueden descomponerse a su vez para conseguir una mayor precisión en los aspectos de usabilidad en los que se quiere poner mayor énfasis. Por ejemplo, rendimiento en uso normal y uso de opciones avanzadas son ambos subatributos de eficiencia, mientras que primera impresión es un subatributo de satisfacción.

2.5 Etapas de las pruebas de usabilidad

En estudios realizados por el Consultor Independiente Eduardo Mercovich se describen 3 etapas para guiar el proceso de pruebas de usabilidad, las cuales están basadas en las actividades a realizar. Las etapas son las siguientes (Mercovich, 2009):

1. **Planificación**:

- **Desarrollo del Plan de pruebas**: documenta los pasos a seguir, permite coordinar los esfuerzos de todos los participantes y definir claramente el rol de cada uno. Además, detalla otros recursos necesarios y sirve como herramienta de comunicación dentro de la organización para con otras áreas o personas que no están directamente relacionadas.
- **Selección de participantes**: se analiza la audiencia del sitio, se definen los criterios y cuestionario de selección. Se suele comenzar con una prueba piloto y 4 ó 5 usuarios. Esta cantidad permite encontrar aproximadamente entre el 65 y el 75% de los problemas importantes de usabilidad.
- **Preparación de los materiales**: se preparan todos los elementos que se utilizarán durante la prueba. Entre otros está el cuestionario pre y post prueba, la introducción para el participante, las definiciones y escenarios para las tareas, otros elementos necesarios vinculados con las tareas y los temas del interrogatorio post prueba.

2. **Prueba:**

- **Prueba piloto:** se evalúa el diseño de pruebas con un participante para verificar que todo funciona como corresponde y todos (Observadores, Facilitador y Participante) conocen sus roles y tareas y las cumplen a la perfección. Al finalizar se hace un pequeño análisis para determinar si es necesario hacer modificaciones al protocolo de pruebas. Para poder hacerlo con tranquilidad, es que la prueba piloto suele llevarse a cabo un día completo antes de comenzar con la prueba definitiva.
- **Prueba definitiva:** se lleva a cabo la prueba con los participantes elegidos. Individualmente se le da la bienvenida, se explica lo que se va a hacer y se procede a la prueba luego de la cual se aplica el cuestionario.

3. **Conclusión:**

- **Análisis de los datos:** se reúnen los datos de los observadores, se analizan estadística y semánticamente, y se discuten sus posibles implicanciones. En la discusión participan los observadores, el facilitador y opcionalmente otros integrantes del entorno de desarrollo.

- **Elaboración del informe:** se elaborará un informe que contendrá los resultados y las recomendaciones para mejorar el sitio de acuerdo a lo observado. El mismo no sólo cumple con la formalidad de terminar la prueba, sino que es el respaldo para explicar a otras personas que no estuvieron involucradas los motivos de las decisiones tomadas.

3.1 Propuesta de procedimiento para pruebas de usabilidad

Desde el momento en que se planifica realizar pruebas de usabilidad a un producto, se debe tener en cuenta varios aspectos como la categoría a utilizar, los atributos y métricas. En busca de un objetivo general que es satisfacer las necesidades de los usuarios, se realiza un conjunto de acciones, que serán descritas a continuación. Es importante que los participantes sientan que forman parte de un equipo que busca obtener un software con una buena calidad.

Figura 1. Etapas de las pruebas de usabilidad.

En la figura número 1 se mencionan las etapas propuestas para el procedimiento de pruebas de usabilidad. Primeramente se realiza la etapa de planificación, luego la de selección de los participantes. Con el plan de pruebas realizado y los participantes seleccionados se pasa a la realización de la prueba, la misma emite los resultados necesarios para brindar el criterio que se necesita en las conclusiones. Cada actividad a realizar es ubicada en una etapa específica, como queda descrito en la figura que a continuación se expone, la cual se lee de izquierda a derecha.

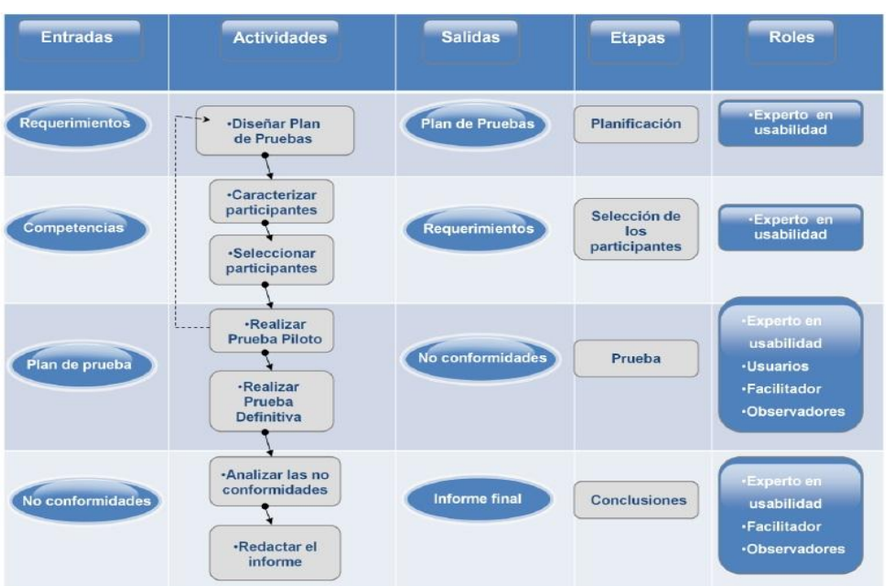

Figura 2. Flujo de actividades de las pruebas de usabilidad.

En la figura número 2, la descripción comienza desde que el equipo de desarrollo de un software detecta la necesidad de aplicar pruebas de usabilidad, brindando al grupo de calidad los requisitos documentados en el modelo de sistema. A partir de los requerimientos se elabora un plan de pruebas, donde, una de las actividades es definir las competencias

de los roles para dar paso a la selección los participantes. Luego de redactado el plan y seleccionado los que desarrollaran cada rol, se aplica la prueba piloto donde, si se detectan deficiencias, se vuelve a la etapa de planificación para corregirlas. En caso de no detectarse ningún problema, se realiza directamente la prueba definitiva y basado en sus resultados se emite un criterio en las conclusiones.

3.1.1. Planificación de las pruebas

La planificación debe tener una atención esmerada debido a que es la etapa donde se definen las bases para la aplicación de la prueba de usabilidad.

Figura 3. Etapa de planificación.

En la figura número 3 se describe la etapa de planificación, en la que se desarrolla un documento denominado plan de pruebas, donde quedan detallados los atributos, métricas, categorías, roles, requisitos a medir y técnicas a utilizar. En cada prueba de usabilidad no es recomendable medir todos los atributos a la vez, ya que se obtienen mejores resultados si se selecciona una muestra de ellos y toda la atención se centra en detectar los problemas enfocados en dichas propiedades.

Métricas

Las métricas son medidas del grado en que un sistema posee un atributo dado, es decir son descritas para medir condiciones de un componente, proceso o producto en general. Basado en los descritos anteriormente se seleccionan las siguientes métricas:

- **Recuerdo:** métrica mediante la cual se mide la facilidad de aprendizaje y la retención en el tiempo. Está basada en verificar qué tanto recuerda un usuario sobre la realización de alguna tarea en la aplicación, después de haber pasado un tiempo de realizada.

- **Exactitud:** se basa en el número de errores que cometen los usuarios en su interacción con el software y el tiempo que necesitan para la realización de cada una de las tareas.

- **Respuesta emocional:** en esta métrica se tiene en cuenta la eficacia, facilidad de uso, satisfacción y la satisfacción subjetiva. Para su medición se deben tener bien claras las diferencias entre los dos últimos atributos, la cual está dada porque uno está enfocado en la tarea y otro en el usuario. La satisfacción normalmente se ve cuando un usuario tiene la capacidad de lograr el efecto que desea. Pero en su búsqueda no siempre ganan ambos tipos de atributos, ya que puede quedar satisfecho con el objetivo logrado pero no con los pasos realizados para lograrlo.

- **Intercambio:** se basa en la medida de la flexibilidad, y robustez. Se realiza haciendo énfasis en la facilidad que brinda el software para realizar tareas de varias formas. También se tiene en cuenta cómo guía al usuario para intercambiar información.

- **Confianza:** se mide observando la cantidad de datos reales que el usuario es capaz de brindar al software mientras interactúa con él.

Medidas para evaluar las métricas

La evaluación de las métricas se expresa en Buena, Regular o Mala de acuerdo con los resultados que se obtienen de realizar la prueba. Para emitir este criterio se tuvo en cuenta el método descrito en Memorias II del Congreso Latinoamericano de Ingeniería Biomédica (Las pruebas de usabilidad en los equipos médicos, 2001). En el mismo se realiza una descripción detallada y basado en dicho estudio se selecciona la clasificación a utilizar que se expone a continuación. En este tipo de evaluación se tienen en cuenta los casos de uso vinculados con la interacción usuario-producto y el número de pasos que se realizan para lograr completar una tarea. Ello está dado principalmente porque las pruebas de usabilidad están enfocadas a satisfacer las necesidades de los usuarios y a que cada no conformidad vinculada una métrica la afecta de forma directa. Todos los software no deben ser medidos de la misma forma, en los productos que tienen mayor cantidad de funcionalidades vinculadas a los usuarios, los clientes tienden a dejar pasar si encuentran una deficiencia pequeña, más si el sistema sólo permite la realización de unas pocas acciones el cliente no soportaría el error por pequeño que fuera.

Basado en el párrafo anterior se llega a la conclusión de que las evaluaciones se deben emitir de acuerdo a la cantidad de no conformidades que afecten una métrica determinada. En todos los sistemas informáticos sólo se evalúa de bien la que no es afectada por ninguna deficiencia. Para las otras evaluaciones se debe tomar el número de casos de uso que tengan como actor a los usuarios, dividirlo entre dos y el resultado es el límite de no conformidades que se pueden tener afectando una métrica para emitir un criterio de regular, en caso que sea mayor la calificación es de mala. Esta forma de evaluación sólo varia en dos casos específicos, donde las funcionalidades tengan mayor dificultad y se necesite personal calificado para su uso y si por el contrario, solo se tiene un pequeño número de actividades. En el primer caso la calificación sería más flexible y en el segundo la mitad del número de no conformidades afectando una métrica ya es considerada para emitir la calificación de mal. La usabilidad del producto a probar en general también es calificada, de acuerdo a la cantidad de métricas con una evaluación dada. Se realiza de la siguiente forma:

- Solo se evalúa de bien si todas tienen esta calificación.
- Si se tiene sólo una evaluada de mal, y el resto es de bien, ya el resultado final es regular.
- Al menos una regular y las demás bien, se considera regular.
- En caso de que haya una mal y el resto regular; es regular
- En caso de que sean 2 o más con mal, el resultado es mal.

Basado en ello el número de no conformidades que afecta a una métrica específica no debe ser mayor que 3 sino la evaluación de la misma sería de mala, si está entre 1 y 3 se considera regular, y sólo es buena si no se encuentran deficiencias relacionadas a la misma.

Categorías

Para aplicar pruebas de usabilidad, se deben seleccionar las categorías necesarias, las cuales están definidas por el tipo de usuarios que se necesita para realizar las tareas. Basada en la selección también es escogida la técnica a utilizar. Existen tres tipos de categorías: pruebas automatizadas, con usuarios expertos, con usuarios reales.

Las pruebas automatizadas se realizan dependiendo de la herramienta a utilizar, las actividades estarían descritas en dependencia de cómo se mediría la usabilidad del software a probar a través de una herramienta.

En las pruebas con usuarios expertos sólo participan expertos en el tema o personal capacitado y facilitadores, donde uno de los participantes es encargado de la planificación. Es válido aclarar que estos usuarios no son solamente conocedores de temáticas relacionadas con la usabilidad, sino que deben tener conocimientos sobre el trabajo en general con aplicaciones web. Este tipo de pruebas es guiado por las etapas y centrado en buscar la satisfacción de los

usuarios finales de determinado software, basado en los conocimientos de los participantes sobre cómo mejorar la interfaz de un producto.

Las pruebas con usuarios reales se realizan con la participación de un experto que la planifica, facilitadores para guiar al usuario dentro de la aplicación, observadores que toman notas sobre todo lo que presencian basado en lo atributos y los usuarios que son los que realizan las tareas para evaluar el software.

Competencias de los roles

Los participantes deben ocupar roles específicos, por lo que se deben documentar las competencias de cada uno.

Figura 4. Competencias que deben cumplir los participantes para desarrollar cada rol.

En la figura número 4 se describen las competencias de los roles, luego se seleccionan los materiales y el lugar para la realización de la prueba. Para ello se tiene en cuenta la forma de acceso a la interfaz del software, la disponibilidad de computadoras, laboratorios y el tiempo de los participantes. Es importante destacar que no es necesario un laboratorio de calidad para realizar las pruebas, es decir que con un local condicionado con los requisitos antes descritos es posible realizar la evaluación.

3.1.2. Selección de los participantes

La selección de los participantes queda descrita como etapa de pruebas de usabilidad y no solamente como una tarea, los roles que intervienen en la realización de las pruebas de usabilidad, no son los que se acostumbran a tener en los equipos, grupos o proyectos de calidad. Por ello se hace necesario un estudio de las competencias que deben tener para ser seleccionados.

Figura 5. Etapa de selección de participantes.

En la figura número 5 se describen las actividades que se realizan en la segunda etapa de las pruebas de usabilidad. Los participantes deben ser seleccionados primeramente por la disponibilidad y luego basados en las competencias. El experto en usabilidad es seleccionado antes de comenzar la prueba por el jefe del equipo de calidad. En caso de

necesitar aplicar la evaluación y no tener experto se hace necesario preparar a alguien para desempeñar este rol, enfatizando en que conozca acerca de la forma de medición de cada métrica.

Selección del facilitador

El facilitador debe ser propuesto por el jefe del equipo de desarrollo del software a medir. Lo cual se debe a que el interesado en participar desarrollando este rol, debe tener conocimientos amplios sobre la realización de las tareas que se pueden hacer en la interacción con el producto. Generalmente es escogido alguien vinculado al proyecto a evaluar. El número de participantes con este rol debe estar dado por la complejidad de las tareas a realizar, es decir, si son tareas sencillas basta sólo con la presencia de uno, si son tareas para las cuales hay que tener ciertos conocimientos se seleccionan entre 2 o 3.

Selección de los observadores En el grupo de calidad debe existir personal capacitado para este rol, de ser así el experto se encarga de informarle su participación en la prueba, los atributos y métricas a evaluar. De no haber ningún integrante con los conocimientos necesarios, el experto selecciona personas del equipo de calidad y los capacita para que cumplan esta función. Las instrucciones que se le dan a los implicados son fundamentalmente sobre las métricas, su forma de medición y relación con las actividades que van a realizar los usuarios. El número de observadores debe estar en relación con el de los usuarios, debido a que tienen que estar al tanto de todas las acciones que se realizan. Normalmente se selecciona un observador por cada una o dos personas a interactuar con el sistema informático.

Selección de los usuarios Para el rol de usuario, se deben seleccionar personas con intereses afines al contenido del producto. Son designados por el experto a través de conversaciones con los interesados, donde se les explica que la prueba es para medir el software y en caso de no realizar alguna tarea correctamente no es culpa suya sino del producto a evaluar. Con estos elementos aclarados y un vocabulario donde no se utilice ninguna palabra técnica, se les proporciona la confianza y satisfacción a los participantes con este rol para desarrollar la prueba de forma relajada. Para la selección del número de usuarios existen varios criterios. Algunos expertos son partidarios de que mientras más participantes para el desempeño de este rol tenga la prueba es mejor, pero en un estudio realizado por *Jakob Nielsen*, considerado el padre de la usabilidad, quedó demostrado que esto sólo traería resultados repetitivos. Un grupo pequeño de personas interactuando con el software a medir brinda un menor costo que uno grande y los mismos o mejores resultados, debido a que al utilizar menor cantidad de usuarios todos los participantes pueden tener un mejor control de la prueba. *Nielsen* basado en su investigación expone que para probar sitios grandes, con muchas funcionalidades, la prueba tiene éxitos con un número de usuarios entre 5 y 8, mientras que para un producto pequeño, sólo son necesarios entre 3 y 5.

3.1.3. Ejecución de las pruebas.

Luego de desarrollado el plan de pruebas y seleccionados los participantes se pasa a la etapa de pruebas, donde se miden los requisitos descritos y se ejecuta todo lo planeado.

Figura 6. Etapa de prueba.

Categoría de pruebas automatizadas

Esta etapa depende de la categoría seleccionada, por ejemplo si es elegida la de pruebas automatizadas, la descripción estaría dada por los desarrolladores de la herramienta a utilizar. Existen herramientas con este fin en el mundo, como *Usabilla y FiveSecondTest*, las cuales fueron hechas para un tipo de software en específico.

Categoría de pruebas con usuarios expertos

La realización de pruebas con usuarios expertos se basa directamente en el tipo de técnica a utilizar, por ejemplo la de evaluaciones heurísticas. Esta prueba se basa en la reunión entre 3 y 5 especialistas, donde se evalúa la interfaz del software a través de heurísticas de diseño descritas por varios autores. Los expertos se reúnen en un laboratorio donde tengan acceso a la aplicación, y emiten un criterio, acerca de la interfaz, basado en varias heurísticas como son:

- **Visibilidad del estado del sistema:** El sistema siempre debe mantener informado al usuario acerca de lo que está pasando, a través de información adecuada en un plazo razonable.

- **Coincidencia entre el sistema y el mundo real:** El sistema debe hablar el lenguaje de los usuarios, con palabras, frases y conceptos familiares al mismo, en lugar de términos orientados al sistema. Seguir las convenciones del mundo real, por lo que la información debe aparecer en un orden natural y lógico.

- **Control de usuarios y libertad:** Los usuarios a menudo eligen funciones del sistema por error y será necesario un marcado claramente como "salida de emergencia" para salir del estado no deseado sin tener que pasar a través de un amplio diálogo.

- **Prevención de errores:** Incluso mejor que buenos mensajes de error es un diseño cuidado que impide que un problema se produzca en el primer lugar. Cualquiera de eliminar las condiciones propensas a errores o detectar los usuarios actuales y con una opción de confirmación antes de comprometerse a la acción.

- **Reconocimiento más que recordar:** Reducir al mínimo la carga de memoria del usuario haciendo que los objetos, acciones y opciones visibles. El usuario no debería tener que recordar la información de una parte del diálogo a otra. Instrucciones de uso del sistema deben ser visibles o fácilmente recuperables cuando sea apropiado.

- **La flexibilidad y la eficiencia en el uso de aceleradores no vistos por el usuario novato:** A menudo se puede acelerar la interacción para el usuario experto de tal manera que el sistema puede servir tanto a usuarios inexpertos y experimentados. Permitir a los usuarios sus acciones frecuentes.

- **Estéticas de diseño:** Los diálogos no deben contener información que es irrelevante o raramente necesaria. Cada unidad adicional de información en un diálogo compite con las dependencias pertinentes de la información y disminuye su visibilidad relativa.

- **Ayuda a los usuarios a reconocer, diagnosticar y recuperarse de los errores:** Los mensajes de error deben ser expresados en un lenguaje sencillo (sin códigos), indicar con precisión el problema y sugerir una solución constructiva.

- **Ayuda y documentación:** A pesar de que es mejor si el sistema puede ser usado sin documentación, puede ser necesaria para proporcionar ayuda y documentación. Toda esa información debe ser fácil de la búsqueda, centrado en la tarea del usuario, lista de medidas concretas que se llevarán a cabo, y no ser demasiado grande.

En esta prueba en vez de utilizar las métricas se usan los principios de diseño, pero al igual se llega a un conjunto de no conformidades para luego redactar las conclusiones.

Categoría de pruebas con usuarios reales

En la categoría de pruebas con usuarios reales primeramente se realiza una prueba piloto, es decir a modo de ensayo se repasa todo lo que se va a realizar. Para esta pequeña evaluación se selecciona sólo un usuario, que puede ser cualquier persona sin medir si cumple o no los requisitos, y se le pide la realización de una tarea, alrededor se ponen en funcionamiento los observadores, facilitador y experto.

Esta evaluación no se lleva hasta el final ya que es para verificar que los pasos a desarrollar estén en orden y que los participantes conocen sus roles y responsabilidades. En caso de detectarse algún problema se vuelve a la etapa de planificación para revisar el plan de pruebas y corregir los errores. Después de verificado que todo está en orden, se pasa a la aplicación de la prueba definitiva, con la presencia de todos los usuarios y el resto de los roles. Inicialmente se brinda una bienvenida y se expone la importancia de lo que se está realizando. Luego a cada uno individualmente se le explica la tarea que deben realizar. Una vez comenzada la prueba el facilitador sólo puede explicar algo si los usuarios lo solicitan. Los observadores toman sus posiciones en un lugar donde tengan visibilidad de los usuarios que antes han escogido para realizar su rol. A medida que los usuarios van terminando se les agradece a través de palabras estimulantes el tiempo dedicado. Luego de que todos hayan terminado se recogen las no conformidades en un documento para dar paso a la etapa de conclusiones.

Otras técnicas

Existen otras formas de evaluar la usabilidad, pero ya serían un complemento para las pruebas tanto con usuarios reales como expertos. Como ejemplo de estas pruebas se tienen:

- **Grupo focal:** se realiza con un grupo de usuarios habituales o posibles del sitio a medir. Para ello se pide una reunión con los participantes, donde debe participar al menos un experto, para mediante temas preparados con anterioridad, conocer los criterios del público.

- **Entrevistas individuales:** tiene semejanza con el grupo focal, pero se realiza sólo entre un usuario y el experto. Para ello son también planificados temas vinculados con el software a evaluar y la usabilidad. Se realiza cara a cara, por teléfono, en entornos tanto formales como informales. Es una técnica donde se busca encontrar al usuario en su entorno normal, para que ocurra el diálogo sin presiones.

- **Cuestionarios:** es uno de los más utilizados para complementar las pruebas. Se realiza normalmente al terminar el usuario de realizar las tareas de una prueba de usabilidad, para así obtener resultados más detallados sobre la satisfacción o cualquier métrica en específico.

Cuando la prueba es complementada con alguna de estas técnicas, los resultados de ambas se unen y se toman como un único grupo de no conformidades. Luego se pasa a la etapa final.

3.1.4. Conclusiones de las pruebas

Con el documento de no conformidades se inicia la realización de las actividades en esta última etapa. Basado en los resultados de la etapa de pruebas se puede emitir un criterio acerca de las métricas a evaluar.

Figura 7. Etapa de conclusiones.

En la figura número 7 se describen las actividades a realizar en esta última etapa. Los conocedores de las métricas y atributos descritos se reúnen después de la prueba para analizar las no conformidades detectadas. Cada no conformidad está relacionada con una métrica para la cual se emite una evaluación de buena, regular o mala en el documento de las conclusiones. El objetivo que se persigue con el análisis es emitir un criterio sobre cada métrica o atributo respaldado por los errores detectados, quedando justificada cada evaluación en el documento de las conclusiones que debe realizar después del análisis el experto en usabilidad. Este documento es entregado al equipo de desarrollo del software, para que sean arregladas las deficiencias. La usabilidad es un aliado para conquistar clientes asegurando que los productos cuentan con calidad y satisfacen la experiencia de uso en cada una de sus funcionalidades.

4. Conclusiones

En el desarrollo de software se busca obtener productos con una calidad cada vez mayor y así resaltar todas las funcionalidades. El proceso de prueba de usabilidad se enfoca en satisfacer las necesidades de los usuarios finales basados en métricas definidas durante la planeación. El desarrollo de este tipo de evaluación no debe de ser olvidado por simple que parezca, ya que un software que durante su desarrollo haya recibido este tipo de prueba va a ser aceptado más rápidamente por sus usuarios que otro en el cual no se hayan tenido en cuenta.

Los referentes teóricos estudiados y los métodos investigativos aplicados permitieron realizar un proceso de sistematización de la teoría concerniente a la elaboración de un Procedimiento de Pruebas de Usabilidad para lograr mayor satisfacción de los clientes.

El aspecto decisivo para determinar la calidad de los productos de software es la experiencia que se lleva el usuario final al utilizarlos. La aplicación del procedimiento descrito en la investigación ayuda a los clientes de un sistema informático a entender fácilmente las funcionalidades que tiene y con ello a que el software obtenga un mayor éxito.

Referencias

PRESSMAN ROGER, S. Ingeniería de Software: Un enfoque Práctico [Libro].- 2005.

SCALONE FERNANDA. Estudio comparativo de los modelos y estándares de calidad de software. [en línea] 2006 [Consultado el: 5 de diciembre de 2012]. Disponible en: [http://laboratorios.fi.uba.ar/lsi/scalone-tesis-maestria-ingenieria-en-calidad.pdf].

MONTAÑEZ LUIS EDUARDO GRATEROL. Diseño de un modelo para el control de la calidad de los proyectos de desarrollo de software en las organizaciones de los estados venezolanos. [en línea] 2006 [Consultado el: 5 de diciembre de 2012]. Disponible en: [http://biblioteca2.ucab.edu.ve/anexos/biblioteca/marc/texto/AAQ9075.pdf].

PRADO ELENA RAJA. Casi todas las pruebas del software. [en línea] 2007. [Consultado el: 10 de diciembre de 2012.]. Disponible en: [http://www.sistedes.es/sistedes/pdf/2007/pris-07-rajactps.pdf].

ROJAS JOHANNA ROJAS Y EMILIO JOSÉ BARRIOS. Investigación sobre estado del arte en diseño y aplicación de pruebas de software. [en línea] 2007 [Consultado el: 30 de noviembre de 2012]. Disponible en: [http://www.udistrital.edu.co].

ISO 9241. 2006. Términos Generales. [en línea] 2006. [Consultado el: 15 de diciembre de 2012.] [http://ver.megared.net.mx/~jccz/iso8402.html].

USABIITY. 2010. [en línea] 2010. [Consultado el: 15 de diciembre de 2012.]. Disponible en [http://Usability.gov.]

NIELSEN JAKOB. Designing Web Usability. [en línea] 2001 [Consultado el: 18 de diciembre de 2012]. Disponible en: [http://www.slideshare.net/webconomia/jakob-nielsen-usabilidad].

MERCOVICH EDUARDO**.** Workshop: cómo hacer un test de usabilidad de un sitio. [en línea] 2009 [Consultado el: 02 de enero de 2013.]. Disponible en: [http://www.gaiasur.com.ar/infoteca/siggraph99/test-de-usabilidad-de-un-sitio.html#que-es-un-test-de-usabilidad].