

Métodos para Evaluación. Aplicación para la medición de Usabilidad en interfaces gráficas

Autores

Msc. Esnolia Noy Monteagudo. esnoy@isdi.co.cu

Ing. Lemay Cruz Pujol. lemayc@isdi.co.cu

Instituto Superior de Diseño (ISDi), Cuba

RESUMEN

La usabilidad ha llegado a ser un campo de actividad establecido en el desarrollo del software y está aumentando de importancia en el campo del diseño de productos de consumo. Es un concepto que se refiere básicamente a la facilidad de uso de una aplicación o producto interactivo.

El poder saber cuan buena es una interfaz de cara al usuario es una tarea que hoy se queda en el plano de las subjetividades.

Para ello es muy importante lograr saber cuál es la medida de la calidad de dicha interfaz. Precisamente la calidad de una interfaz se mide a partir de cuan eficiente y eficaz es el uso de dicha interfaz, se considera entonces que cuan más “usable” es la interfaz, más calidad tiene.

Con este trabajo se persigue realizar un análisis de los diversos métodos de evaluación y como los mismos pueden tributar a la medición de la usabilidad.

Después de analizados 15 métodos de evaluación se concluye que la selección de un método o conjunto de métodos requiere tomar ciertas consideraciones. Los elementos a analizar dentro de cada método son: la estrategia, la localización, el prejuicio, la medida de la Usabilidad, la información, la inmediatez de la respuesta, la intromisión y el costo. Y que la selección de un método o de varios viene dado por un profundo análisis del contexto y del elemento a evaluar.

INTRODUCCIÓN

El éxito de un programa o software depende directamente de la calidad de su interfaz. Al usuario final poco o nada le importa el tipo de tecnología que se utiliza detrás de la interfaz, él no entiende de protocolos, paquetes o enrutamiento, el solo juzga lo que ve y cómo se comporta frente a sus necesidades. Pero lo que sorprende de este asunto, es la poca motivación que presentan las empresas en invertir dinero y tiempo para la evaluación y mejoramiento de sus interfaces, aun sabiendo las altas tasas de retorno de dicha inversión. A esta situación se le atribuyen diferentes causas, entre ellas:

el simple desconocimiento de técnicas o herramientas para ejecutar esta labor, pero en otros, consecuencia de una cultura de desarrollo anticuado centrado en la tecnología y no en el usuario y sus tareas.

Para ello es muy importante lograr saber cuál es la medida de la calidad de dicha interfaz. Precisamente la calidad de una interfaz se mide a partir de cuan eficiente y eficaz es el uso de dicha interfaz, se considera entonces que cuan más “usable” es la interfaz, más calidad tiene.

Retomando la definición de Usabilidad ofrecida por la ISO, (“grado de eficacia, eficiencia y satisfacción con la que usuarios específicos pueden lograr objetivos específicos, en contextos de uso específicos”, se puede identificar dos tipos de atributos:

- Atributos cuantificables de forma objetiva: como la eficacia, número de errores cometidos por el usuario durante la realización de una tarea, por ejemplo; y eficiencia, tiempo empleado por el usuario para la consecución de una tarea.
- Atributos cuantificables de forma subjetiva: como es la satisfacción de uso, cuya medida se aproxima a través interrogatorios al usuario (por ejemplo) y tienen una estrecha relación con el concepto de Usabilidad Percibida (medida de Usabilidad de un sistema teniendo como única base la opinión y percepción del usuario).

Diversos autores han abordado el problema de medir la usabilidad, algunos orientados más a un tipo de atributo que a otro.

Una primera aproximación, quizá la más sencilla de todas, consiste en mantener una guía de estilos para un tipo de aplicación y/o entorno. Cuando se mantiene una imagen uniforme dentro de toda una línea de productos, la familiaridad, tanto en estructura como comandos, se hereda de una interfaz a otra, lo cual garantiza muchos de los aspectos de usabilidad para un sistema (recordación, entrenamientos cortos, entre otros).

Cuando se pretende evaluar la usabilidad de un sistema que se rige bajo este principio, el evaluador determina el grado de concordancia del aplicativo respecto a la guía que implanta o sigue. En términos prácticos, estos sistemas obtienen buenos resultados, pero afrontan los mismos problemas de cualquier otra interfaz cuando es un usuario inexperto el que utiliza la aplicación, lo que no sucede con un usuario avanzado en aplicativos que hayan utilizado la misma guía.

Para ello es importante conocer si es posible utilizar los métodos de evaluación para evaluar la usabilidad de las interfaces gráficas. Con el presente trabajo se persigue hacer un análisis de los mismos y ver si logran y en qué medida resolver esta problemática.

La clasificación de los métodos de evaluación de usabilidad obedece a diferentes criterios: el grado de implicación del usuario, escenarios de tarea, el empleo de reglas o por el objetivo de la evaluación, se dispone de una clasificación de métodos de evaluación de la usabilidad propuestos por diferentes autores.

La selección de un método o conjunto de métodos requiere tomar ciertas consideraciones. A partir de los análisis de diversos autores se obtienen los siguientes:

Estrategia: Estado del ciclo de vida del proceso de desarrollo en el cual se ejecuta la evaluación.

Localización: La localización en la cual se realizará la evaluación: en un laboratorio o en un entorno real.

Prejuicio: Nivel de subjetividad u objetividad inherente en el método. Los resultados de la realimentación pueden ser altamente influenciados por el prejuicio con el cual se desarrolle la evaluación.

Medida de la Usabilidad: Tipo de medida provisto por el método. Puede ser cuantitativo o cualitativo.

Información: Información o realimentación provista por el método. La información describe elementos de bajo nivel: combinación apropiada de colores, fuente de letra de mejores características, íconos de mayor comprensión, etc. El concepto de realimentación es de alto nivel más que una primera impresión. Detalla problemas de mayor complejidad que no siempre tienen solución con un cambio de formato o color.

Inmediatez de la respuesta: Concierne a cuán rápido el método retorna realimentación. Una observación de video es un proceso que demanda mayor tiempo en su análisis, pero puede

llegar a encontrar problemas complejos de usabilidad; por otro lado, los resultados de un prototipo en papel estarán disponibles tan pronto como finalice la observación.

Intromisión: concierne al hecho que el comportamiento de algunos usuarios se vea alterado por la presencia de un observador o un sistema de registro.

Costo: Recursos involucrados para la ejecución y análisis del método.

Para poder generar elemento de juicio sobre la selección de un posible método a utilizar durante la evaluación de la usabilidad, se realiza a continuación un resumen de algunos de los métodos más citados tanto por autores como empresas.

Evaluación Heurística

Descripción: Analizar la conformidad de la interfaz con unos principios reconocidos de usabilidad (las heurísticas) mediante la inspección de varios evaluadores expertos.

Estrategia: se puede aplicar durante todo el proceso de desarrollo, cada etapa puede considerar la utilización de este método manejando diferentes enfoques y artefactos, por ejemplo, durante la captura de requisitos, la evaluación podría ser aplicada a sistemas actuales o alternativas de la competencia.

Localización: laboratorio

Prejuicio: alta subjetividad del experto. Los resultados y recomendaciones son criterio del evaluador.

Medida de la Usabilidad: típicamente cualitativa

Información: abundante información (problemas a bajo nivel: colores, fuente, etc.); baja realimentación (problemas de alto nivel)

Inmediatez de la respuesta: rápido retorno de información.

Intromisión: ninguna, la evaluación heurística no involucra la participación de usuarios reales.

Costo: bajo costo de recursos y tiempo

Recorrido Pluralista

Descripción: se reúne tanto a desarrolladores y expertos como con usuarios, se trabaja sobre prototipos, pueden ser en papel, se pide que ejecuten ciertas tareas sobre el prototipo mientras anotan cada paso de la interacción. Luego se debaten los problemas y decisiones tomadas durante el proceso. Combina tres visiones diferentes del sistema: lo que se debe (experto), se quiere (usuarios) y se puede (desarrollador).

Estrategia: se encuentra una mayor fortaleza para este método en las primeras fases del desarrollo, el impacto ante los cambios propuestos y la generación de un lineamiento de trabajo para futuros modelo de mayor nivel, es evidente.

Localización: laboratorio

Prejuicio: alta subjetividad de los participantes.

Medida de la Usabilidad: cualitativa

Información: alta información (problemas a bajo nivel: colores, fuente, etc.); baja retroalimentación (problemas de alto nivel), la atención no está centrada en la observación del usuario, sino en el debate entre las partes.

Inmediatez de la respuesta: rápido retorno de información.

Intromisión: baja, los usuarios realizan su observación y sacan sus propias conclusiones, luego viene el proceso de debate en el cual la estar frente a un equipo de desarrollo y un grupo de expertos muchas veces intimida al usuario.

Costo: mediana inversión de recursos, sobre todo de tiempo y organización.

Caminata Cognitiva

Descripción: este método tiene por eje central evaluar el rápido aprendizaje del sistema. Está basado en los recorridos estructurales tradicionales que se usan en la comunidad de la ingeniería de software. Partiendo de una interfaz detallada, un escenario y una tarea, se pide a usuarios reales interactuar con el sistema. Cada paso que el usuario realice será anotado y contrastado con la secuencia de acciones bajo las cuales el analista planteo la interfaz. Las sesiones son grabadas en video y analizadas para descubrir el origen de fallos, errores o frustraciones de los usuarios. Se trata de reconstruir paso a paso la interacción del usuario con el sistema.

Estrategia: se recomienda su utilización durante las fases de diseño e implementación

Localización: laboratorio

Prejuicio: alta subjetividad de los usuarios.

Medida de la Usabilidad: cuantitativa o cualitativa, se puede manejar como un grado de coincidencia entre la ruta esperada y los caminos de solución tomados por el usuario.

Información: media información; alta realimentación, se evidencian problemas críticos de la interfaz que vayan más allá del modelo estético.

Inmediatez de la respuesta: media, es necesario un análisis detallado de las conductas del usuario durante la interacción.

Intromisión: media, la sección se registra sobre video, lo cual puede condicionar ciertas decisiones del usuario.

Costo: mediana inversión de recursos.

Inspección de estándares

Descripción: Este método se realiza por medio de un experto en un estándar que puede ser de facto de la interfaz. El experto realiza una inspección minuciosa a la interfaz para comprobar que cumple en todo momento y globalmente todos los puntos definidos en el estándar.

Estrategia: fases avanzadas del desarrollo, pruebas, e implantación.

Localización: laboratorio

Prejuicio: objetivo

Medida de la Usabilidad: cuantitativa, porcentaje de cumplimiento con el estándar.

Información: alta información; baja o nula realimentación.

Inmediatez de la respuesta: alta, resultados disponibles en cuanto termine el proceso. El tiempo en la ejecución de las verificaciones dependerá de las herramientas con que cuente el experto y la complejidad del estándar.

Intromisión: nula, no hay presencia de usuario.

Costo: baja inversión de recursos.

Estudio de Campo / Método Etnográfico

Descripción: centra su atención en los aspectos tangibles del entorno de ejecución de la aplicación. Selecciona perfiles de usuarios reales, diseña entrevistas y organiza sesiones de observación directa del entorno de trabajo. Registro en fotos, archivos de trabajo, etc. El resultado final de este método, es el reflejo del modelo mental del usuario con respecto a la tarea y artefactos (objetos tangibles en la interacción). Información muy útil dentro de etapas de definición del sistema

Estrategia: fases de definición, captura de requisitos e implantación, vista como los requerimientos de una segunda versión del producto.

Localización: entorno real

Prejuicio: subjetiva, las apreciaciones están a juicio del observador.

Medida de la Usabilidad: cualitativa, el propósito es conocer, más que juzgar.

Información: baja información; alta realimentación, modelo mental de las tareas.

Inmediatez de la respuesta: media, los resultados requieren de un análisis de los datos observados para la generación del modelo mental.

Intromisión: media, la observación y entrevistas interfieren con el usuario.

Costo: mediana inversión de recursos.

Grupo de discusión dirigido (focus group)

Descripción: reúne de 6 a 9 usuarios para discutir aspectos relacionados con el sistema. Un ingeniero de factores humanos hace las veces de moderador, que tiene que preparar la lista de aspectos a discutir y recoger la información que necesita de la discusión. Esto puede permitir capturar reacciones espontáneas del usuario e ideas que evolucionan en el proceso dinámico del grupo.

Estrategia: pruebas e implantación.

Localización: laboratorio

Prejuicio: subjetiva, las apreciaciones se gestan en debate.

Medida de la Usabilidad: cualitativa. Apreciaciones del sistema.

Información: media información; media realimentación.

Inmediatez de la respuesta: media, los datos recogidos tienden a tener una validez baja y son muy difíciles de analizar debido a su naturaleza no estructurada y de flujo libre.

Intromisión: alta, el proceso de debate termina por alterar la opinión de los usuarios.

Costo: mediana inversión de recursos.

Entrevistas

Descripción: consta de una serie de preguntas y tópicos el cual el usuario debe responder a un revisor; aspectos tales como su experiencia con el sistema, actitud e impresiones. El revisor debe adaptar la entrevista tal que pueda obtener el máximo beneficio en cada sesión.

Estrategia: pruebas e implantación.

Localización: laboratorio, entorno real

Prejuicio: subjetiva, apreciaciones del usuario

Medida de la Usabilidad: cualitativa, apreciaciones del sistema.

Información: baja información; alta realimentación.

Inmediatez de la respuesta: alta.

Intromisión: media.

Costo: baja inversión de recursos.

Cuestionarios

Descripción: es un procedimiento menos flexible que una entrevista, pero permite llegar a más usuarios y generar informes estadísticos con mayor facilidad.

Estrategia: pruebas e implantación.

Localización: laboratorio, entorno real

Prejuicio: subjetiva y objetiva. Aun cuando los cuestionarios se llenen subjetivamente, un número considerable de muestras conlleva a una visión objetiva del sistema.

Medida de la Usabilidad: cualitativa, apreciaciones del sistema.

Información: baja información; alta realimentación.

Inmediatez de la respuesta: baja.

Intromisión: baja o nula.

Costo: baja inversión de recursos, sobre todo aprovechando los recursos tecnológicos.

Grabación de Uso (logging)

Descripción: implica disponer en el computador de una ampliación del sistema que recoja automáticamente estadísticas sobre el uso detallado del sistema. Es útil porque muestra cómo los usuarios realizan su trabajo real y porque es fácil recoger automáticamente datos de una gran cantidad de usuarios que trabajan bajo diversas circunstancias.

Estrategia: pruebas e implantación.

Localización: laboratorio, entorno real

Prejuicio: objetiva

Medida de la Usabilidad: cualitativa

Información: baja información; alta realimentación.

Inmediatez de la respuesta: baja, depende de la estrategia de análisis de los datos.

Intromisión: baja, la captura puede hacerse de forma transparente al usuario.

Costo: alto, consecuencia del diseño y construcción de los elementos de captura de la información, más el tiempo de análisis de los datos.

Pensando en voz alta (thinking aloud)

Descripción: se pide a los usuarios que expresen en voz alta sus pensamientos, sentimientos y opiniones mientras que interactúan con el sistema. Es muy útil en la captura de un amplio rango de actividades cognitivas.

Estrategia: diseño, pruebas e implantación.

Localización: laboratorio, entorno real

Prejuicio: subjetivo

Medida de la Usabilidad: cualitativa

Información: media información; alta realimentación. Ayuda a comprender el modelo mental de la tarea y proporciona información como terminología empleada, funcionalidades deseables y necesidades en la documentación.

Inmediatez de la respuesta: alta

Intromisión: alta, no constituye un proceso natural para el usuario.

Costo: alto

Interacción Constructiva (codescubrimiento)

Descripción: es una derivación del método pensando en voz alta e implica el tener en vez de uno, dos usuarios que hagan el test al sistema conjuntamente. La principal ventaja es ser mucho más natural que el pensar en voz alta, ya que las personas normalmente verbalizan cuando tratan de resolver un problema conjuntamente y además hacen muchos más comentarios; la desventaja es que los usuarios pueden tener diferentes estrategias de aprendizaje.

Estrategia: diseño, pruebas e implantación.

Localización: laboratorio

Prejuicio: subjetivo

Medida de la Usabilidad: cualitativa.

Información: media información; alta realimentación. Al tenerse estrategias de aprendizaje distintas se produce una interferencia entre las opiniones de los participantes, esto reduce la capacidad de abstraer el modelo mental de la tarea.

Inmediatez de la respuesta: alta

Intromisión: baja, convirtiendo el proceso de pensar en voz alta, en algo normal.

Costo: alto

Examen Retrospectivo

Descripción: se parte del hecho que, si se ha realizado una grabación en vídeo de la sesión de evaluación, es posible recoger más información haciendo que el usuario revise la grabación. De este modo, el usuario puede dar opiniones más amplias acerca de su interacción y el observador puede preguntar sin interferir el proceso. Demanda un mayor tiempo por sesión.

Estrategia: diseño, pruebas e implantación.

Localización: laboratorio

Prejuicio: subjetivo

Medida de la Usabilidad: cualitativa

Información: media información; alta realimentación.

Inmediatez de la respuesta: alta

Intromisión: baja, las preguntas se realizan examinando el video de la sesión.

Costo: alto

Método del conductor (Coaching method)

Descripción: contrario a muchos otros métodos, donde la interferencia del observador es evitada en lo posible, este método guía o conduce a los usuarios en la dirección correcta mientras usan el sistema. Este método se centra en el usuario inexperto y su propósito es descubrir las necesidades de información de los usuarios de tal manera que se proporcione un mejor entrenamiento y documentación al mismo tiempo que un posible rediseño de la interfaz para evitar la necesidad de preguntas.

Estrategia: diseño, pruebas e implantación.

Localización: laboratorio

Prejuicio: subjetivo

Medida de la Usabilidad: cualitativa.

Información: alta información; alta realimentación.

Inmediatez de la respuesta: alta

Intromisión: alta, la sesión completamente guiada.

Costo: alto

Método de Modelos

Descripción: utilización de los modelos GOMS (Goal, Operations, Methods and Selection Rules). Un modelo GOMS es una descripción del conocimiento que un usuario debe tener para poder llevar a cabo una tarea sobre un dispositivo o un sistema. Consiste de la descripción de los Métodos necesarios para culminar una Meta específica. Los Métodos están compuestos de Operadores que son una serie de pasos que el usuario interpreta. Un Método puede exigir el cumplimiento de sub-Metas para ser culminado, por tanto, los métodos tienen una estructura jerárquica. Si hay más de un método comprometido en una Meta, entonces Reglas de Selección escogen el Método

apropiado dependiendo del contexto. Describir las Metas, Operadores, Métodos y Reglas de Selección para un conjunto de tareas de una manera formal constituye un análisis GOMS o define el denominado modelo GOMS. El más simple de los modelos GOMS es el Keystroke-Level Model descrito por primera vez por Card en 1980, en el cual el total del tiempo de ejecución de una tarea se predice por el total de tiempo para las acciones elementales a nivel de teclado.

Estrategia: diseño, pruebas e implantación.

Localización: laboratorio

Prejuicio: objetivo

Medida de la Usabilidad: cuantitativa y cualitativa.

Información: baja información; nula realimentación. Si por ejemplo debe seleccionarse entre dos sistemas, el GOMS es un buen criterio de elección.

Inmediatez de la respuesta: alta

Intromisión: baja complejidad en el uso de estos modelos

Costo: bajo

Además de estos y otros métodos, existe una serie de técnicas auxiliares que pueden ayudar a los diseñadores e ingenieros en Usabilidad a tomar mejores decisiones; técnicas que van desde acercamientos con el usuario (Stakeholder meeting), organización y estructuración de la información (Card Sorting, Affinity Diagramming, Brainstorming), hasta guías relacionadas directamente con el desarrollo (Prototipos Rápidos, Guías de Estilo).

Parte del estudio del Diseño Centrado en el Usuario es proveer estas herramientas y conceptos adaptados al contexto particular de uso. Cada organización o equipo de trabajo puede, a partir de su experiencia y conocimiento, generar una variante de los métodos expuestos o crear uno completamente nuevo que se ajuste a sus condiciones particulares de entorno.

CONCLUSIONES

- El concepto usabilidad engloba la eficiencia de uso, la eficacia y la satisfacción, y este está muy relacionado al diseño de interfaz
- La selección de un método o conjunto de métodos requiere tomar ciertas consideraciones. Los elementos a analizar dentro de cada método son: la estrategia, la localización, el prejuicio, la medida de la Usabilidad, la información, la inmediatez de la respuesta, la intromisión y el costo.
- La selección de un método o de varios viene dado por un profundo análisis del contexto y del elemento a evaluar

BIBLIOGRAFÍA

- Apple Computer. "Human interface guidelines". 1987
- Chapanis, A. "Human Factors in Systems Engineering". 1996.
- Holcomb and Tharp. "Design for successful guessing". 1991
- ISO 9241-11. "Dialogue principles". 2001
- McCormick, Ernest J. "Ergonomía. Factores humanos en ingeniería y diseño", 1980.
- Nielsen, Jakob. "Usabilidad. Diseño de sitios Web". 2000.
- Nielsen, Jacob. "Usability heuristics". 1993
- Norman, Donald. "La psicología de los objetos cotidianos 1988.
- Norman, Donald A. "Seven principles that make difficult task easy". 1988
- Norman, Donald. "Emotional Design. Why we love (or hate) everyday things". 2004.
- Polson and Lewis. "Design for successful guessing". 1990
- Ravden and Johnson. "Evaluation check list for software inspection". 1989 Salvendy, G. "Handbook of Human Factors and Ergonomics". 1997.
- Shneiderman. "Eight golden rules of dialogue design". 1986