

USABILIDAD EN APLICACIONES MÓVILES

Lic. Enriquez Juan Gabriel¹
Dra. Casas Sandra Isabel²

Resumen: La usabilidad es considerada uno de los factores más importantes dentro de la calidad de un producto de software. Debido a esto es de interés poder contar con metodologías para medir la usabilidad de las aplicaciones. Los métodos de análisis de usabilidad que actualmente se utilizan, métodos clásicos, fueron desarrollados para aplicaciones de escritorio.

Con la aparición, el uso masivo y el crecimiento de los dispositivos móviles, especialmente los Smartphones, la medición de usabilidad en aplicaciones móviles se tornó un tema de investigación. Los métodos y métricas actualmente utilizados para medir usabilidad pueden no ser directamente aplicables a este tipo de productos.

Los estudios de usabilidad se han efectuado sobre diferentes contextos, recientemente se empezó a considerar la usabilidad en un contexto móvil. Este trabajo examina los métodos y las métricas utilizados para medir usabilidad y pretende analizar que desafíos existen al momento de realizar pruebas de usabilidad en aplicaciones móviles, donde el contexto que cambia continuamente pasa a tener un rol preponderante.

Palabras Claves: Usabilidad, métricas, aplicaciones móviles, contexto móvil.

¹ Autor: jenriquez@unpa.edu.ar

² Director: scasas@unpa.edu.ar

ICT desarrollado en el marco del proyecto UNPA 29/A273-1.

INTRODUCCIÓN

Las aplicaciones o productos de software cuando son lanzados al mercado se espera que tengan cierto grado de aceptación entre los usuarios, ese grado va a depender de las características particulares que cada usuario considere importantes. Desde el punto de vista de la Ingeniería de Software (SE: Software Engineering), una de las principales características que tiene que tener una aplicación para ser exitosa entre los usuarios es que sea de calidad. Resulta relevante para los desarrolladores de software poder medir esa calidad o realizar pruebas de calidad a las aplicaciones construidas, pero para poder medir se necesita saber qué es lo que hay que medir y cómo.

La calidad del software es el grado que el software posee de una combinación deseada de atributos, esta combinación de atributos deberá ser claramente especificada. Definir calidad de software para un sistema es equivalente a definir una lista de atributos de calidad del software requeridos para ese sistema [1]. Utilizando esta definición se puede afirmar que lo que hay que medir más puntualmente son ciertos atributos del software relacionados a la calidad. Dentro de estos atributos uno de los considerados más importantes es la usabilidad, que indica la facilidad con la que un usuario puede usar una aplicación de software. Por lo tanto resulta de interés poder obtener una medida del grado de usabilidad que tiene una aplicación.

Debido a que en los últimos años el uso de dispositivos móviles (teléfonos móviles, reproductores de audio portátil, asistentes personales digitales, navegadores gps, tablets, cámaras digitales, etc.) se ha incrementado de manera considerable, es importante disponer de metodologías y herramientas que permitan realizar estudios de usabilidad específicos para aplicaciones desarrolladas para estos tipos de dispositivos (aplicaciones móviles).

Las aplicaciones móviles son aquellas que fueron desarrolladas para ejecutarse en dispositivos móviles. El término móvil se refiere a poder acceder a los datos, las aplicaciones y los dispositivos desde cualquier lugar. Para desarrollar software de este tipo se tiene que tener en cuenta ciertas restricciones que tiene el hardware de estos dispositivos, como por ejemplo que son de dimensiones reducidas, tienen bajo poder de cómputo, escasa capacidad de almacenamiento, ancho de banda limitado, etc. Algunos ejemplos de aplicaciones móviles son: mapas y navegación, búsqueda, juegos, mensajería, aplicaciones empresariales.

Los métodos y métricas actualmente utilizados para medir usabilidad fueron creados para aplicaciones de escritorio, sin embargo estos pueden no ser directamente adecuados o apropiados a entornos móviles [2]. Uno de los desafíos consiste en identificar las variables adicionales relacionadas al ambiente de uso (contexto móvil) que pueden impactar en la usabilidad de una aplicación móvil.

Este trabajo se enfoca en considerar los distintos aspectos relacionados con la medición de la usabilidad en general, y particularmente en aplicaciones móviles. Se analizan los métodos y métricas utilizados actualmente en las pruebas de usabilidad de una aplicación e intenta identificar los retos que existen al realizar pruebas de usabilidad en aplicaciones móviles. En las cuales el contexto pasa a tener más relevancia y debe ser considerado en mayor medida en este tipo de análisis.

El informe está organizado de la siguiente manera: en la Sección 1 se describen los conceptos teóricos relacionados a la usabilidad en aplicaciones de software; se brinda una definición y se detallan los distintos aspectos (atributos, métricas y modelo). En la Sección 2 se explican los tipos de pruebas de usabilidad. En la Sección 3 se especifica cuales son las características de las aplicaciones móviles y consideraciones a tener en cuenta en el desarrollo de las mismas. En la Sección 4 se describe el contexto de uso de este tipo de aplicaciones (contexto móvil) y su influencia en la usabilidad. La Sección 5 desarrolla como se realizan las pruebas en las aplicaciones móviles y por último en la Sección 6 se muestran las conclusiones sobre las pruebas y el análisis de usabilidad en sistemas móviles.

USABILIDAD

La usabilidad en general tiene que ver con la forma en que se usa algún elemento (herramienta, dispositivo electrónico, etc.), es la facilidad con que se usa y si permite hacer lo que se

necesita. Particularmente la usabilidad de una aplicación de software se refiere a la facilidad con que los usuarios pueden utilizar la misma para alcanzar un objetivo concreto. Este nivel de usabilidad no puede medirse o ser evaluado directamente, debido a que depende de diferentes factores.

Formalmente, la definición más utilizada o reconocida de usabilidad es la que se expone en la norma ISO 9241-11³, en la cual usabilidad se describe como el grado con el que un producto puede ser usado por usuarios específicos para alcanzar objetivos específicos con efectividad, eficiencia y satisfacción, en un contexto de uso específico [3]. La norma define como especificar y medir la usabilidad de productos y aquellos factores que tienen un efecto en la misma; también destaca que la usabilidad en terminales con pantalla de visualización es dependiente del contexto de uso y que el nivel de usabilidad alcanzado dependerá de las circunstancias específicas en las que se utiliza el producto. El contexto de uso lo forman los usuarios, las tareas a realizar, el equipamiento (hardware, software y materiales), así como también los entornos físicos y sociales que pueden influir en la facilidad de uso de un producto.

De la definición anterior se puede observar que la usabilidad está relacionada con los atributos de una aplicación o sistema, así como también de su contexto; se entiende por atributo la característica o propiedad de una aplicación de software. En la norma mencionada anteriormente los atributos considerados son los siguientes [4]:

Efectividad: Está relacionada con la precisión y completitud con la que los usuarios utilizan la aplicación para alcanzar objetivos específicos. La calidad de la solución y la tasa de errores son indicadores de efectividad.

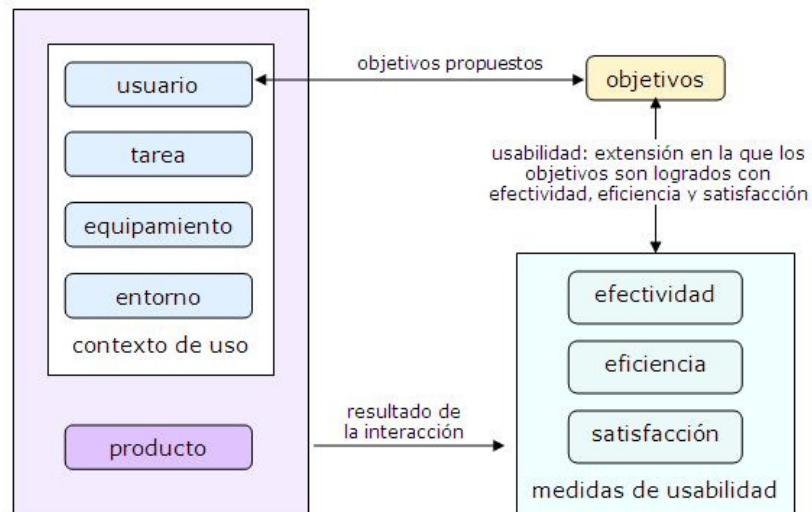
Eficiencia: Es la relación entre efectividad y el esfuerzo o los recursos empleados para lograr esta. Indicadores de eficiencia incluyen el tiempo de finalización de tareas y tiempo de aprendizaje. A menor cantidad de esfuerzo o recursos, mayor eficiencia.

Satisfacción: Es el grado con que el usuario se siente satisfecho, con actitudes positivas, al utilizar la aplicación para alcanzar objetivos específicos. La satisfacción es un atributo subjetivo, puede ser medido utilizando escalas de calificación de actitud.

Para poder especificar o medir la usabilidad, es necesario descomponer los atributos y el contexto de uso en componentes medibles y verificables. Las relaciones que existen entre el usuario, el producto, los atributos, el contexto de uso y los objetivos que se quieren lograr se pueden observar en el framework de usabilidad propuesto en la norma citada (Fig. 1).

3 ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) – part 11: Guidance on usability.

Figura 1. Framework de usabilidad (ISO 9241-11).



Diferentes percepciones de usabilidad resultan en una variedad de estructuras de medidas, estas estructuras definen los atributos y sus métricas asociadas. Por ejemplo, Jakob Nielsen, define usabilidad en términos de cinco atributos: facilidad de aprendizaje, eficiencia, memorabilidad, errores y satisfacción [5]. Nielsen puntualiza que la importancia de estos atributos variará en función del contexto y de los usuarios finales. Por ejemplo en un cajero automático, la facilidad de aprendizaje podría ser considerado el atributo principal a evaluar, mientras que en un sistema aeroportuario, la tolerancia y prevención de errores pasa a ser primordial.

El término atributo de usabilidad es bastante ambiguo, esto genera ciertas diferencias, algunos autores usan diferentes nombres para el mismo atributo, tales como memorabilidad y facilidad de aprendizaje. Existen diferentes opiniones sobre que atributos pueden ser considerados y la manera de combinarlos para componer la usabilidad [6].

Los siguientes son algunos de los atributos utilizados para medir el grado de usabilidad de una aplicación de software:

Facilidad de Aprendizaje: La facilidad con la que los usuarios alcanzan objetivos específicos la primera vez que utilizan la aplicación. La primera experiencia que tiene los usuarios con un nuevo sistema es la de aprender a usarlo.

Memorabilidad: La facilidad para memorizar la forma de utilizar la aplicación y alcanzar objetivos específicos, y la facilidad con que vuelven a utilizar la aplicación después de un tiempo. La curva de aprendizaje debe ser significativamente menor para un usuario que ya utilizó el sistema, que para uno que es la primera vez que lo va a utilizar.

Errores: Los errores que comete el usuario al utilizar la aplicación y la gravedad de los mismos. La aplicación debe producir la menor cantidad de errores posibles. Si se producen, es importante que se den a conocer al usuario de forma rápida y clara, además de ofrecer algún mecanismo para recuperarse de ese error.

Contenido: Aspectos relacionados a la distribución del contenido y de los formatos utilizados para mostrar información al usuario.

Accesibilidad: Consideraciones tenidas en cuenta por posibles limitaciones físicas, visuales, auditivas o de otra índole de los usuarios.

Seguridad: Capacidad para alcanzar niveles aceptables de riesgo. Disponibilidad de mecanismos que controlan y protegen la aplicación y los datos almacenados.

Portabilidad: Capacidad de la aplicación de ser transferida de un entorno a otro (diferentes plataformas).

Contexto: Relacionado a los factores o variables del entorno de uso de la aplicación.

A la vez, los atributos de usabilidad, pueden ser clasificados en objetivos y subjetivos [7]. Los atributos objetivos pueden ser medidos a través de la interacción del usuario con la aplicación, no dependen de la percepción del usuario; en cambio los subjetivos están relacionados con el factor humano, se refiere a la actitud del usuario hacia el uso de la aplicación, está vinculado a las emociones y por lo tanto son más difíciles de medir y cuantificar.

Métricas de Usabilidad

Debido a que los atributos de una aplicación son conceptos abstractos, estos no pueden ser directamente medidos. Para medirlos se les asocian distintas métricas, por ejemplo, el atributo eficiencia puede ser evaluado mediante la métrica que calcula el tiempo empleado por un usuario en terminar una tarea específica.

Una métrica (medida) es un valor numérico o nominal asignado a características o atributos de un objeto computado a partir de un conjunto de datos observables y consistentes con la intuición [8]. Una métrica debe cumplir con ciertas características:

- Debe tener características matemáticas deseables.
- Cuando una métrica representa una característica que aumenta cuando se presentan rasgos positivos o que disminuye al encontrar rasgos indeseables, el valor de la métrica debe aumentar o disminuir en el mismo sentido.
- Cada métrica debe validarse empíricamente en una amplia variedad de contextos antes de publicarse o aplicarse en la toma de decisiones.

Una de las clasificaciones para las métricas, es dividir las en estáticas y dinámicas [9]. Las métricas estáticas son utilizadas para medir las características estáticas de una aplicación, como el tamaño del código o la complejidad del mismo. Las dinámicas permiten medir el comportamiento de la aplicación, se calculan con la aplicación en ejecución. Cabe aclarar que las métricas no representan un fin por sí mismas, estas revelan datos e información sobre la experiencia personal del usuario cuando hace uso de una aplicación. La información obtenida de las métricas nos ayuda a realizar un mejor análisis y tomar decisiones más acertadas con respecto a la usabilidad de una aplicación.

La Tabla 1 muestra atributos de usabilidad y las métricas comúnmente asociadas a los mismos para poder cuantificarlos.

Tabla 1. Atributos y métricas asociadas.

Atributos	Métricas
Efectividad	<ul style="list-style-type: none">• Tareas resueltas en un tiempo limitado.• Porcentaje de tareas completadas con éxito al primer intento.• Número de funciones aprendidas.
Eficiencia	<ul style="list-style-type: none">• Tiempo empleado en completar una tarea.• Número de teclas presionadas por tarea.• Tiempo transcurrido en cada pantalla.• Eficiencia relativa en comparación con un usuario experto.

	<ul style="list-style-type: none"> • Tiempo productivo.
Satisfacción	<ul style="list-style-type: none"> • Nivel de dificultad. • Agrada o no agrada. • Preferencias.
Facilidad de Aprendizaje	<ul style="list-style-type: none"> • Tiempo usado para terminar una tarea la primera vez. • Cantidad de entrenamiento. • Curva de aprendizaje.
Memorabilidad	<ul style="list-style-type: none"> • Número de pasos, clicks o páginas usadas para terminar una tarea después de no usar la aplicación por un periodo de tiempo.
Errores	<ul style="list-style-type: none"> • Número de errores.
Contenido	<ul style="list-style-type: none"> • Cantidad de palabras por página. • Cantidad total de imágenes. • Número de páginas.
Accesibilidad	<ul style="list-style-type: none"> • Tamaño de letra ajustable. • Cantidad de imágenes con texto alternativo.
Seguridad	<ul style="list-style-type: none"> • Control de usuario. • Número de incidentes detectados. • Cantidad de reglas de seguridad.
Portabilidad	<ul style="list-style-type: none"> • Grado con que se desacopla el software del hardware. • Nivel de configuración.
Contexto	<ul style="list-style-type: none"> • Grado de conectividad. • Ubicación. • Características del dispositivo.

Dependiendo de la naturaleza de la aplicación a estimar, en las pruebas se consideran relevantes diferentes atributos. La usabilidad del sistema no es una simple adición del valor de estos atributos, sino que se define para cada sistema como un nivel a alcanzar por cada uno de los atributos especificados para ese sistema. En la Tabla 2 se pueden observar distintos atributos utilizados en la medición de usabilidad y el entorno dónde tienen más influencia en la misma. Por ejemplo el atributo “contenido” es más apropiado considerarlo en la medición de un sitio web que en una aplicación de escritorio; el atributo “contexto” está relacionado a la usabilidad de las aplicaciones móviles, en otros entornos se puede considerar pero este no resulta primordial para la evaluación.

Tabla 2. Atributos y entorno de uso.

Atributos	Aplicaciones de Escritorio	Sitios Web	Aplicaciones Móviles
Efectividad	X	X	X
Eficiencia	X	X	X
Satisfacción	X	X	X
Facilidad de Aprendizaje	X	X	X
Memorabilidad	X	X	
Errores	X	X	X
Contenido		X	X
Accesibilidad		X	

Seguridad	X	X
Portabilidad		X
Contexto		X

Modelo de Usabilidad

Diferentes estándares o modelos para cuantificar y evaluar la usabilidad se han propuesto en la Interacción Humano-Computador (HCI: Human-Computer Interaction) y las comunidades de Ingeniería de Software. Sin embargo, la usabilidad no se ha definido de manera consistente a través de estos estándares y modelos. No definen una guía precisa sobre los atributos, métricas y reglas a utilizar en la medición de usabilidad de una aplicación, tampoco de cómo seleccionar y medir aspectos de la misma. Debido a que los factores relacionados a la usabilidad son diversos, estos estándares y modelos no cubren todos los aspectos de la misma. Tampoco están adecuadamente integrados en las actuales prácticas de la Ingeniería de Software, y por lo general carecen de herramientas informáticas que lo soporten. Una consecuencia de estas deficiencias es que la mayoría de los desarrolladores de software no aplican correctamente un modelo en particular para la evaluación de la usabilidad. Las actuales prácticas suelen ser personalizadas de cada desarrollador, esto lleva a que el esfuerzo en medir usabilidad pueda ser desperdiciado por no disponer de un consistente y consolidado modelo [10].

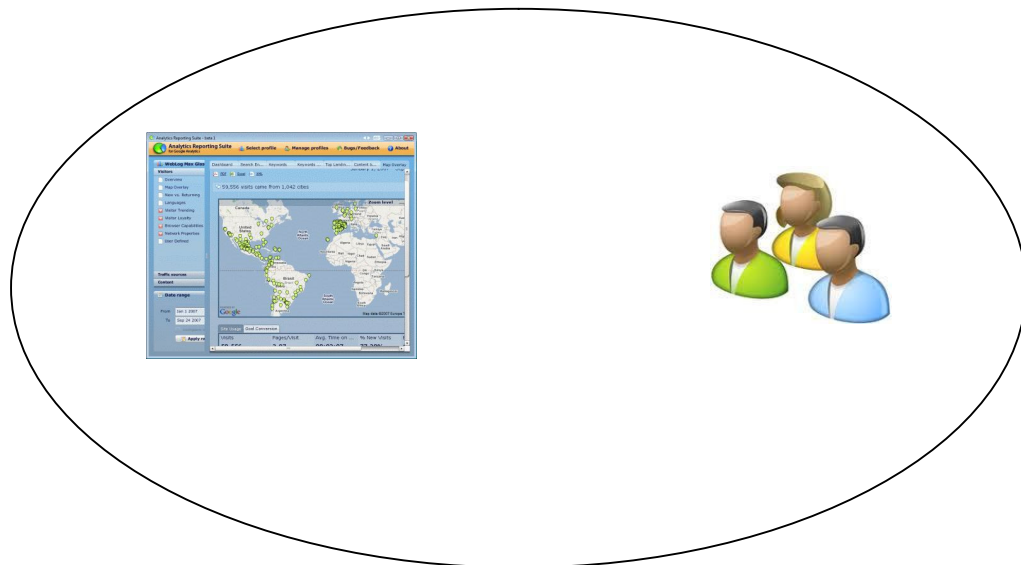
Entre estos estándares o modelos, están los siguientes:

- ISO 9241-11 [3] mencionada anteriormente, explica los beneficios de medir usabilidad en términos de rendimiento y satisfacción del usuario. Estos son medidos por el grado en que los objetivos previstos son alcanzados, los recursos que se han invertido para conseguir esos objetivos, y el grado en que el usuario encuentra aceptable el uso del producto. Esta norma define usabilidad y explica cómo identificar la información que es necesaria tener en cuenta a la hora de especificar o evaluar la usabilidad de un terminal de pantalla visual. Ofrece una guía sobre cómo describir el contexto de uso del producto (hardware, software o servicios) y las medidas de usabilidad relevantes de una manera explícita. La guía se da en forma de principios y técnicas generales, en lugar de en forma de requerimientos específicos.
- MUSiC [11] (Metrics for Usability Standards in Computing), este modelo se refiere específicamente a la definición de medidas de usabilidad del software. Ejemplos de métricas de usabilidad en el marco de MUSiC incluyen medidas de rendimiento, tales como la efectividad en las tareas de los usuarios, eficiencia temporal, y proporción del período productivo. Sin embargo, este modelo no puede reflejar otros aspectos de usabilidad, tales como la satisfacción del usuario o facilidad de aprendizaje. Por tal motivo, como parte del proyecto MUSiC, un cuestionario de satisfacción de los usuarios de 50 ítems llamado Software Usability Measurement Inventory (SUMI) fue desarrollado para proporcionar medidas de satisfacción global.
- SANE [12] (The Skill Acquisition Network), es un modelo de análisis de la calidad de uso para dispositivos interactivos. Este enfoque asume un modelo de interacción del usuario que define las tareas del usuario, la dinámica del dispositivo y los procedimientos para la ejecución de tareas de usuario. Específicamente, un modelo de tarea y un modelo de dispositivo se desarrollan simultáneamente y posteriormente son relacionados. A continuación, se simulan los procedimientos de usuario en el modelo tarea-dispositivo generado. Un total de 60 medidas diferentes se describen en este marco, de los cuales 24 se refieren a medidas de calidad. Las puntuaciones de estos últimos se combinan entonces para formar un total de cinco medidas de calidad compuestas, incluyendo: eficiencia, aprendizaje, adaptabilidad, carga cognitiva de trabajo y esfuerzo para la corrección de errores.

- QUIM [13] (Quality in Use Integrated Measurement), es un modelo consolidado para evaluar usabilidad. Describe una jerarquía en la que se descompone usabilidad en factores, éstos en criterios, y finalmente en métricas específicas. Detalla cómo establecer los requisitos de calidad, así como también identificar, implementar, analizar y validar los procesos y las métricas de calidad del producto. El principal propósito de QUIM es proporcionar un marco de trabajo consistente y un repositorio de factores de usabilidad, criterios y métricas para educación y fines de investigación.

Los modelos de usabilidad esencialmente están formados por la aplicación que se va a evaluar y la interacción que el usuario tiene con esta para alcanzar sus objetivos. La aplicación tiene definidos atributos, y métricas asociadas para medir dichos atributos. Estos componentes se encuentran definidos dentro de un cierto contexto de uso (Fig. 2).

Figura 2. Modelo de usabilidad.



PRUEBAS DE USABILIDAD

La evaluación de la usabilidad de una aplicación de software, consiste en realizar pruebas para obtener medidas e información y observar debilidades relacionadas al uso de la misma. Las cuatro formas básicas de evaluación son: automática (se calculan las métricas mediante la ejecución de la aplicación), empírica (la usabilidad es evaluada testeando la aplicación con usuarios reales), formal (usando modelos formales y fórmulas para el cálculo de medidas de usabilidad), e informal (basados en reglas generales y la habilidad y experiencia de los evaluadores). Este proceso de evaluación implica varias actividades en función del método empleado, las actividades comunes que incluyen son [14]:

Captura: Recolección de datos de usabilidad, tales como el tiempo de terminación de una tarea, errores, valores subjetivos.

Análisis: Interpretar los datos recolectados para identificar problemas de usabilidad.

Crítica: Sugerir soluciones o mejoras para mitigar los problemas encontrados.

Las pruebas de usabilidad se pueden realizar en diferentes etapas del desarrollo del software.

Algunas técnicas de evaluación, sólo se pueden aplicar cuando existe un prototipo de la aplicación, mientras que otras se pueden aplicar a principios de la etapa de diseño. Cada técnica tiene sus propios requerimientos y generalmente diferentes técnicas descubren diferentes problemas.

Los beneficios que se obtienen al realizar pruebas de usabilidad a las aplicaciones son los siguientes:

- Mejorar el diseño del producto de software mediante la información obtenida de la retroalimentación.
- Reducir los costos de desarrollo, al disminuir el número de cambios posteriores requeridos.
- Reducir los costos de mantenimiento y soporte: Los sistemas que son fáciles de usar requieren menos entrenamiento, menos soporte para el usuario y menos mantenimiento.
- Mejorar el uso: los sistemas que mejor se ajustan a las necesidades del usuario reducen el esfuerzo y mejoran la productividad y la calidad de las acciones.
- Mejorar la calidad del producto: productos de mayor calidad de uso, son más competitivos en un mercado que demanda productos de fácil uso.

A continuación se describen las características principales de los métodos de evaluación de usabilidad más reconocidos.

Métodos Empíricos

Este tipo de pruebas se desarrollan en laboratorios que intentan simular las condiciones reales bajo las cuales se usa la aplicación. Se le solicita al usuario que realice tareas previamente definidas, relacionadas con la funcionalidad de la aplicación. La interacción del usuario con la aplicación es evaluada para detectar los problemas que experimenta el usuario con el uso de la misma.

Para realizar la evaluación se necesita analizar la información de usabilidad que se desprende de la interacción. Para capturar esa información, por un lado están las métricas que se van obteniendo en tiempo real al interactuar el usuario con la aplicación. Por otro lado, se puede registrar la forma de realizar las tareas por parte del usuario mediante una grabación de video y de audio; también se pueden efectuar entrevistas o cuestionarios.

La ventaja de este tipo de pruebas es que directamente se monitorea el uso de la aplicación por parte de usuarios reales. La desventaja es que se deben realizar sobre aplicaciones implementadas casi en su totalidad, lo que implica que cualquier modificación que se tenga que realizar demandará mayor tiempo [15]. Si la aplicación es un prototipo, los resultados de las pruebas son utilizados para mejorar la versión final de la misma.

Métodos de Inspección

Inspección de usabilidad es el nombre genérico para un conjunto de formas rentables de evaluación de las interfaces de usuario para encontrar problemas de usabilidad, son muy informales y fáciles de usar.

Este método consiste en formar un grupo de expertos en usabilidad que analizan o inspeccionan la aplicación considerada. Estos realizan un informe comentando sobre distintos aspectos de usabilidad de la aplicación, basándose en su experiencia en el área. Este informe es utilizado para realizar los cambios o ajustes necesarios en la aplicación, para resolver los problemas indicados. Dos de los métodos más utilizados de este tipo son: evaluación heurística y recorrido cognitivo [15].

Evaluación heurística: Un grupo de evaluadores inspecciona el diseño de la interfaz basándose en un conjunto de heurísticas de usabilidad (principios de usabilidad establecidos). La evaluación heurística es fácil de realizar, económica y capaz de encontrar varios problemas de usabilidad. Sin embargo, puede no encontrar problemas específicos del dominio. Es por eso que el uso de apropiadas heurísticas es muy significativo.

Recorrido cognitivo: Se utiliza una descripción de la interfaz de usuario, un conjunto de escenarios de trabajo y las acciones específicas que el usuario debe realizar para cumplir las tareas. Los inspectores examinan y analizan cada paso en la secuencia de acciones definidas, documentando los problemas encontrados.

El método MUSIC

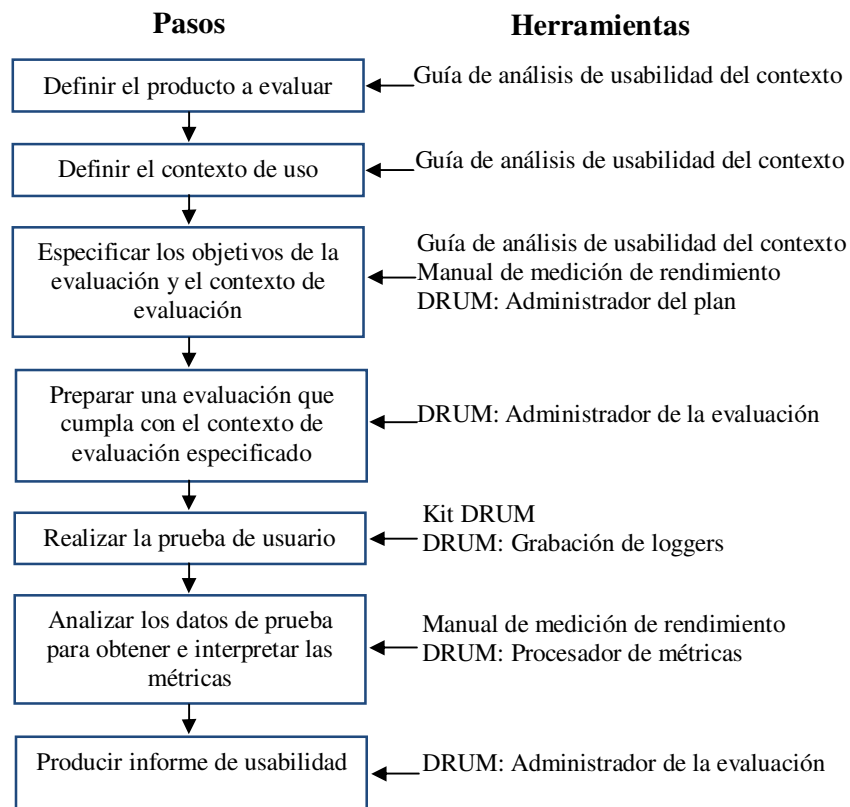
El proyecto MUSiC (Metrics for Usability Standards in Computing) [11], es una metodología rigurosa que proporciona un medio válido y fiable para la especificación y medición de la usabilidad. Fue desarrollado y perfeccionado para satisfacer la demanda de aplicaciones comerciales. Este es soportado por una serie de herramientas que se pueden elegir y usar de acuerdo a las necesidades del desarrollo, el presupuesto y los plazos de tiempo.

La salida básica del método son medidas relacionadas a la efectividad y la eficiencia. También se pueden generar las siguientes salidas opcionales: eficiencia relativa del usuario (indicador de facilidad de aprendizaje), periodo productivo (proporción de tiempo sin tener problemas en el uso) y tiempo empleado en resolver problemas, en búsquedas de ayuda o en navegación no productivas o a través del sistema.

El uso principal del método es formar parte integral de un proceso de desarrollo basado en prototipos y mejoramiento iterativo. Esto significa que las pruebas de usabilidad deben ejecutarse desde las etapas tempranas del desarrollo hasta la entrega final.

La Fig. 3 resume la secuencia de pasos requeridos por el método para evaluar usabilidad y obtener las correspondientes medidas. La columna de la derecha muestra las herramientas que provee el método, las guías y soporte para cada paso. Los pasos 4 a 7 se pueden repetir dentro del proceso.

Figura 3. Pasos y Herramientas usadas en el método MUSiC.



Para que las evaluaciones se puedan ejecutar sin problemas y eficientemente, es conveniente realizarlas en laboratorios de usabilidad. Como alternativa, los datos pueden ser capturados en el lugar de trabajo, si por ejemplo, los factores claves en un sistema de información o su entorno ambiental no pueden ser adecuadamente replicados en un laboratorio.

Una de las herramientas que da soporte al método se llama DRUM (Diagnostic Recorder for Usability Measurement) [16], consiste de un software para realizar estudios observacionales de video asistido. Las sesiones de evaluación se graban en video y posteriormente se pueden analizar con la ayuda de DRUM. Un primer análisis se puede realizar en tiempo real durante la grabación. Este software permite automatizar actividades, crear registros con marca de tiempo de cada sesión de evaluación y calcular las medidas y métricas. También asiste a la generación y entrega de información de diagnóstico relativa a defectos de usabilidad.

APLICACIONES MÓVILES

Las tecnologías móviles y su continuo avance están propiciando una nueva generación de aplicaciones, estas son las denominadas “aplicaciones móviles”. Se considera aplicación móvil, a aquel software desarrollado para dispositivos móviles. Móvil se refiere a poder acceder desde cualquier lugar y momento a los datos, las aplicaciones y los dispositivos. Este tipo de aplicaciones se desarrollan teniendo en cuenta las limitaciones de los propios dispositivos, como por ejemplo el bajo poder de cómputo, la escasa capacidad de almacenamiento, ancho de banda limitado, etc. Los dispositivos móviles son suficientemente livianos como para ser transportados por personas y disponen de la capacidad de batería adecuada para funcionar de forma autónoma. Algunos de estos dispositivos se pueden observar en la Fig. 4.

Estos dispositivos están dominados por diferentes plataformas tecnológicas, incluyendo diferentes sistemas operativos. Cada uno tiene sus particularidades en cuanto al manejo por parte del usuario, como así también al momento de desarrollar una aplicación. Los sistemas operativos para móviles son mucho más simples que los de una computadora y están más orientados a la conectividad inalámbrica.

Figura 4. Dispositivos Móviles.



Existen dos categorías en las que se pueden clasificar las aplicaciones móviles [17]: aplicaciones nativas y aplicaciones Web.

Aplicaciones nativas: Las aplicaciones nativas son desarrolladas específicamente para un tipo de

dispositivo y su sistema operativo, se basan en la instalación de código ejecutable en el dispositivo del usuario. Estas tienen la ventaja de acceder a las funciones del dispositivo, como por ejemplo: almacenamiento, GPS (sistema de posicionamiento global), SMS (servicio de mensajes cortos), mails, etc. Existen repositorios de los cuales se pueden descargar e instalar este tipo de aplicaciones, según el sistema operativo. El principal inconveniente de estas aplicaciones es que se deben desarrollar para cada plataforma y por lo tanto incrementa el tiempo de desarrollo, costo y esfuerzo.

Aplicaciones Web: Las aplicaciones móviles de este tipo se encuentran ejecutándose en servidores, estas incluyen páginas web optimizadas para ser visualizadas en dispositivos móviles y se pueden desarrollar en HTML, Java Script, CSS, etc. Por definición, estas aplicaciones serán accedidas utilizando algún navegador web. La ventaja que tiene desarrollar aplicaciones móviles Web es que son fáciles de implementar y de integrar con aplicaciones existentes, además de necesitar menos requerimientos del hardware de los dispositivos móviles. El problema que tienen es que no pueden acceder a las funcionalidades propias del dispositivo. Por ejemplo, una aplicación web no puede emplear la cámara de un Smartphone, en el caso que la tuviera, para capturar imágenes o realizar una filmación.

Desarrollo de aplicaciones móviles

Debido a la heterogeneidad de los dispositivos móviles se requiere que el desarrollador lleve a cabo una cuidadosa planificación y diseño antes de la implementación de la aplicación. Ejecutar una aplicación en un dispositivo móvil introduce una serie de consideraciones que un desarrollador/programador debe conocer [18]:

- Distintas velocidades y características de la red. Al ser dispositivos móviles atravesarán distintas redes.
- Errores de red. La falta de una red de datos disponible es muy probable, en estos casos la aplicación debe tener un modo fuera de línea.
- Variación del rendimiento de la plataforma de hardware. Lo ideal es que la aplicación esté disponible para tantos dispositivos como sea posible. Esto significa soportar diferentes dispositivos y diferentes plataformas.
- Distintos tamaños y resoluciones de pantallas. Los diferentes dispositivos cuentan con diferentes pantallas con distintas funcionalidades.
- Difícil de probar las aplicaciones por completo. Dada la variedad de dispositivos, se torna difícil probar todos los dispositivos actuales y los nuevos que ingresan al mercado.

Las aplicaciones de este tipo tienen que manejar ciertos requerimientos [19]:

- Potencial interacción con otras aplicaciones.
- Manejo de sensores como por ejemplo el acelerómetro para responder a movimientos del dispositivo o pantalla táctil para responder a gestos.
- Seguridad en la aplicación. Numerosos dispositivos utilizan software embebido, por lo tanto son cerrados y no resulta sencillo invadir el software incorporado. En cambio las plataformas móviles son abiertas, lo que permite instalación de software malicioso que puede afectar el funcionamiento general del dispositivo.
- Bajo consumo de energía. Muchos aspectos de una aplicación afectan el consumo de energía, con esto se ve afectada la capacidad de la batería, es decir la vida útil del dispositivo.

CONTEXTO MÓVIL

Los dispositivos junto con las aplicaciones móviles son utilizados en un cierto contexto, donde las características del mismo cambian continuamente. Al utilizarlos, los usuarios van a tener ciertas

particularidades, tendrán diferentes objetivos, realizarán diferentes tareas; también se manipularán en diferentes entornos físicos y sociales. Todos estos factores y otros, relacionados al contexto móvil, influyen en la forma de uso de una aplicación.

Como ejemplo, en [20] se evidencia la influencia que tiene la información del contexto al realizar una evaluación de usabilidad. Una de las observaciones que muestra este trabajo es la relación entre el movimiento y su impacto en la entrada de datos; estar en movimiento provoca distracciones para el usuario y hace difícil ingresar datos en un dispositivo móvil. Además, restricciones físicas en el contexto a veces generan problemas en la interacción con la aplicación, por ejemplo, la falta de buena iluminación.

El proyecto MUSiC, en un intento de garantizar que se preste adecuada atención a los factores del contexto, puntualizó lo siguiente:

- La usabilidad de un producto depende de su contexto de uso.
- Los productos deben ser diseñados para contextos específicos.
- La medición de la usabilidad debe ser siempre realizada en un contexto apropiado.
- Las evaluaciones de usabilidad siempre deben ir acompañada de una descripción detallada del contexto de las mediciones.

Para medir la usabilidad en aplicaciones móviles y obtener resultados reales, es necesario considerar el contexto como parte integral de la aplicación. Si no se tiene en cuenta el entorno se está realizando un análisis sesgado de usabilidad en comparación con lo que sucede cuando un usuario usa la aplicación en el mundo real. En ese entorno real la conectividad (ancho de banda) puede ir cambiando según el lugar donde se encuentre el usuario, afectando el uso de la aplicación.

Para el dominio de las aplicaciones móviles cuyo contexto de uso cambia continuamente, como se plantea en [2], existen algunos desafíos al momento de examinar la usabilidad:

Entorno móvil: Se puede definir como cualquier información que caracteriza una situación relacionada con la interacción entre usuarios, aplicaciones y sus alrededores. Esto incluye la ubicación, las identidades de las personas cercanas, objetos, elementos del entorno que pueden distraer la atención del usuario, etc. Este puede estar cambiando continuamente.

Conectividad: El ancho de banda limitado y la baja confiabilidad de las conexiones, son características comunes en la conectividad de estos tipos de dispositivos. Este problema afecta principalmente a los tiempos de descarga de datos y la calidad de la transmisión de audio y video. La intensidad de la señal y la velocidad de transferencia de la red pueden variar en diferentes momentos y lugares, agravadas también por la movilidad del usuario.

Capacidad de procesamiento: El poder computacional y la capacidad de memoria de los dispositivos móviles son reducidos con respecto a dispositivos considerados de escritorio.

Pantallas pequeñas: Las limitaciones físicas propias de los dispositivos móviles, como el tamaño pequeño de las pantallas afectan la usabilidad de una aplicación.

Pantallas con diferentes resoluciones: La baja resolución de la pantalla puede degradar la calidad de la información multimedia mostrada. Como resultado, diferentes resoluciones de pantalla puede causar diferentes grados de usabilidad.

Método de entrada de datos: El ingreso de datos en estos dispositivos no es una tarea sencilla. Diminutos botones y etiquetas limitan la eficacia y eficiencia de los usuarios en la introducción de datos, esto reduce la velocidad de entrada y aumenta los errores. En general no es eficiente y nada amigable.

USABILIDAD EN APLICACIONES MÓVILES

En las tecnologías móviles a diferencia de otras (Web, escritorio, TV digital), la usabilidad es un problema más significativo, esto es debido a que una gran mayoría de las aplicaciones móviles son difíciles de usar, son poco flexibles y no son robustas.

Debido a la movilidad que permiten estos dispositivos las pruebas de usabilidad en un ambiente real de uso son difíciles de realizar. En primer lugar, puede ser complicado establecer estudios realistas que reflejan la riqueza del contexto de uso descrito anteriormente. En segundo lugar, está lejos de ser trivial aplicar técnicas de evaluación clásicas, cuando la prueba se realiza en condiciones reales de uso. Por último, las pruebas en un entorno real complican la recolección de datos y limitan el control sobre éstos y sobre el usuario, ya que el mismo se está moviendo físicamente en un contexto con un número de variables desconocidas que afectan potencialmente la evaluación [21].

El trabajo realizado en [22] propone un modelo, llamado mGQM (mobile Goal Question Metric), el cual define un conjunto de métricas para evaluar la usabilidad de las aplicaciones móviles. El modelo propone un conjunto de métricas objetivas y subjetivas, las que a continuación se detallan:

Métricas Objetivas

Tiempo requerido para introducir los datos: Esta métrica mide el tiempo tomado por el usuario para ingresar datos de entrada.

Número de errores al teclear los datos: Esta métrica mide el número de errores cometidos, mientras el usuario introduce datos al utilizar la aplicación.

Tiempo empleado para instalar: Esta métrica mide el tiempo empleado por cada aplicación que se instala en el dispositivo móvil.

El número de interacciones durante la instalación de la aplicación: Esta métrica mide la cantidad de interacciones entre el usuario y el dispositivo móvil cuando se está instalando una aplicación.

Tiempo necesario para aprender: Este indicador mide el tiempo empleado por los usuarios para aprender a utilizar la aplicación.

Número de errores mientras aprende: Este indicador mide el número de errores cometidos por los usuarios cuando están aprendiendo a utilizar la aplicación.

Número de errores: Este indicador mide si la aplicación es precisa o imprecisa.

Tiempo necesario para completar la tarea: Este indicador medirá el tiempo empleado por el usuario para completar las tareas dadas.

Número de tareas exitosas en el primer intento: Esta métrica indica el número de tareas terminadas con éxito en el primer intento.

Número de tareas exitosas en el tiempo dado: Este indicador proporciona información sobre si los usuarios están familiarizados con la aplicación mediante el recuento del número de tareas completadas con éxito en un tiempo dado.

Tiempo necesario para iniciar la aplicación: Este indicador mide el tiempo empleado por los dispositivos móviles para iniciar una aplicación.

Tiempo necesario para responder: Esta métrica mide el tiempo empleado por una aplicación para

responder a la entrada del usuario.

Tiempo necesario para conectarse a la red (vía satélite o Wi-Fi): El tiempo empleado por la aplicación para conectarse a la red es una medida esencial porque refleja la satisfacción de los usuarios.

Número de recursos del sistema mostrados: Esta métrica cuenta el número de recursos del sistema (nivel de la señal o de la batería) que se muestra durante el estudio de usabilidad.

Número de solicitudes de actualización de la aplicación: Esta métrica cuenta el número de solicitudes de las aplicaciones para actualizar la versión.

Porcentaje de la batería usada durante la instalación: Este indicador mide el porcentaje de batería que se utiliza cuando se instalan una aplicación en un dispositivo móvil.

Porcentaje de la batería usada: Este indicador mide el porcentaje de la batería usada por una aplicación para una tarea determinada.

Métricas Subjetivas

Satisfacción con el teclado virtual: Este indicador mide el grado de satisfacción del usuario con el teclado proporcionado por la aplicación móvil o por el dispositivo móvil.

Satisfacción con la salida: Este indicador mide el nivel de satisfacción con la salida proporcionada por las aplicaciones.

Satisfacción con el proceso de instalación: Este indicador mide el nivel de satisfacción de los usuarios cuando instalan aplicaciones en los dispositivos móviles.

Satisfacción con la optimización del tamaño de pantalla: Este indicador mide el nivel de satisfacción de los usuarios con la optimización del tamaño de la pantalla.

Satisfacción con la ayuda: Este indicador mide el grado de satisfacción de los usuarios con la ayuda proporcionada por las aplicaciones.

Satisfacción con los contenidos: Satisfacción con el contenido es una medida esencial para asegurar que los contenidos suministrados cumplirán las necesidades de los usuarios.

Placer: Esta métrica mide el disfrute de los usuarios cuando utilizan las aplicaciones móviles.

Satisfacción con la interfaz: Satisfacción con la interfaz es también una medida importante porque una buena interfaz atraerá a más usuarios a utilizar la aplicación.

Seguridad durante la conducción: Este indicador mide el nivel de seguridad para las personas que utilizan las aplicaciones móviles mientras conducen.

Facilidad para encontrar ayuda: Esta métrica mide cuán fácil es para los usuarios encontrar ayuda en las aplicaciones.

Estrés: Esta métrica mide el nivel de estrés que sienten los usuarios cuando usan las aplicaciones móviles.

Satisfacción con el indicador de señal: Esta medida indica la satisfacción del usuario con el indicador

de señal proporcionada por las aplicaciones móviles.

Satisfacción con el joystick virtual: Esta métrica indica el nivel de satisfacción de los usuarios con el joystick virtual.

Satisfacción mientras aprenden: Este indicador mide el nivel de satisfacción de los usuarios cuando aprenden a utilizar las aplicaciones.

Satisfacción con el texto: Este indicador mide el nivel de satisfacción de los usuarios con el texto que muestran las aplicaciones.

Satisfacción con el sistema de navegación: Este indicador mide el grado de satisfacción con el sistema de navegación (como los usuarios pasan de un menú a otro menú o submenú).

Satisfacción con la pantalla táctil: Esta métrica mide el nivel de satisfacción de los usuarios con la pantalla táctil.

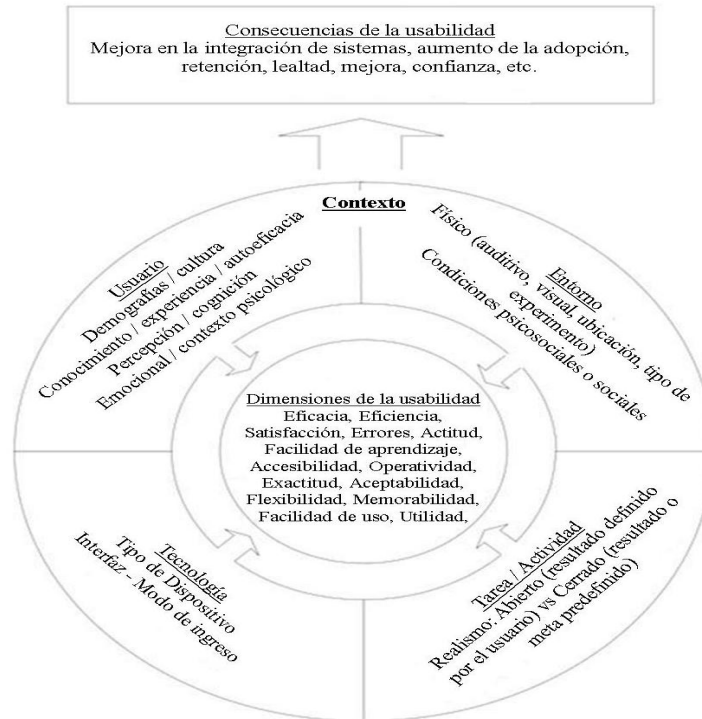
Satisfacción con el botón de menú: Esta medida indica el nivel de satisfacción de los usuarios con el botón de menú.

Como se mencionó anteriormente, otro factor influyente al momento de medir el grado de usabilidad de una aplicación móvil es el ambiente o contexto real donde el usuario utiliza la aplicación. Este contexto móvil da lugar a diferentes factores que afectan la usabilidad, los usuarios pueden estar de pie, caminando, o sentados en un ambiente oscuro o luminoso mientras usan una aplicación. Estos factores toman valor en este tipo de aplicaciones, producen que los enfoques tradicionales para evaluar usabilidad resulten inapropiados.

Uno de los principales problemas de las pruebas clásicas, para entornos de escritorio, al pretender utilizarlas en aplicaciones móviles, es que ignoran la importancia del contexto. Diferentes evaluaciones de usabilidad han sido propuestas en diferentes contextos, tales como, sitios web, librerías digitales, productos electrónicos audiovisuales y muchos otros [23]. En el trabajo citado se propone un framework para usabilidad en un contexto móvil, que se basa en una revisión cualitativa de estudios empíricos sobre usabilidad móvil. Este representa una forma fácil y directa para identificar y abordar las diferentes dimensiones contextuales de usabilidad móvil.

El framework (Fig. 5) se conforma de tres elementos. En primer lugar, el círculo exterior muestra los cuatro factores contextuales (Usuario, Tecnología, Tarea / Actividad, y Entorno). El círculo interior detalla las dimensiones de usabilidad claves (Eficacia, Eficiencia, Satisfacción, Errores, Actitud, Facilidad de aprendizaje, Accesibilidad, Operatividad, Exactitud, Aceptabilidad, Flexibilidad, Memorabilidad, Facilidad de uso, Utilidad, Alegría). Por último, en el cuadro de la parte superior de los factores contextuales se observa una lista de las consecuencias de la usabilidad (Mejora en la integración de sistemas, aumento de la adopción, retención, lealtad, mejora, confianza, etc.).

Figura 5. Framework de usabilidad móvil.



Relacionando el contexto y su influencia en la usabilidad de una aplicación, en [24] se propone un enfoque que se basa en el monitoreo de las condiciones ambientales durante la prueba de usabilidad, tales como la luz, la aceleración, el sonido, la temperatura, y la humedad, y relacionar estas con las acciones del usuario. Los datos de los sensores pueden utilizarse para analizar las causas de los problemas encontrados en una prueba.

Existen dos formas de realizar las pruebas de usabilidad en las aplicaciones móviles: una son los experimentos de laboratorio y la otra los estudios de campo. El primero es muy útil para comparar diferentes diseños de interfaces o diferentes mecanismos de entrada de datos. Una limitación importante al realizar pruebas en un laboratorio es que se ignora el contexto de uso, donde el usuario realmente utiliza la aplicación. Los estudios de campo tienen como principal ventaja disponer del contexto dinámico móvil, características difíciles de simular en un laboratorio. Otra prueba menos desarrollada son las pruebas remotas.

Pruebas de laboratorio

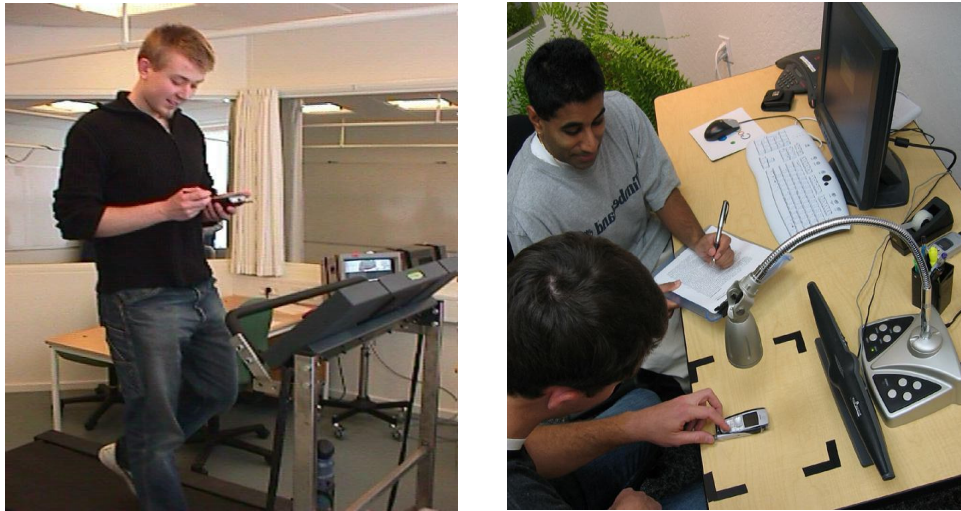
La evaluación de los sistemas móviles en laboratorios intenta simular artificialmente el entorno de uso. Sin embargo algunos factores del contexto como el social puede resultar difícil de representar.

Una forma de simular las condiciones reales de uso, como la movilidad, es hacer caminar al usuario mientras utiliza el sistema móvil que se está evaluando. Esto se lleva a cabo ya sea en una cinta caminadora o en una pista específicamente definida para la prueba. Los datos cualitativos y cuantitativos recolectados, pueden ser el tiempo de realización de las tareas, la tasa de error, la frecuencia cardíaca, la velocidad de marcha promedio, etc. [21].

La Fig. 6 por un lado (izquierda) muestra una prueba de laboratorio incorporando como factor del contexto la movilidad; por otro lado (derecha) se puede ver como se realiza una evaluación de una aplicación móvil utilizando técnicas clásicas de evaluación, en la cual no se tiene en cuenta el contexto de uso.

La recolección de datos en las pruebas de laboratorio es mucho más fácil. En este tipo de pruebas se emplean los métodos de recolección de datos clásicos, tales como los logs, la observación, la entrevista, cuestionarios, y el protocolo verbal. La principal desventaja de las pruebas de laboratorio es que carecen de la información contextual y su influencia.

Figura 6. Prueba de Laboratorio⁴.



Pruebas de campo

Las pruebas de campo permiten obtener una idea de los aspectos relacionados al uso real de la aplicación, estos son cruciales para el éxito en el diseño de tecnologías móviles. Las evaluaciones de campo complican la recolección de datos y limitan el control de usuario, ya que los usuarios se están moviendo físicamente en un entorno con un número de variables desconocidas que pueden afectar la evaluación [25].

La Fig. 7 muestra a un usuario utilizando la aplicación que está siendo analizada, acompañado por dos evaluadores. Un evaluador hace de observador, mientras que el otro registra las acciones mediante una cámara de video.

⁴ Fotos extraídas de “Experimental Evaluation of Techniques for Usability Testing of Mobile Systems in Laboratory Setting” [21] (izquierda) y de “Towards the Perfect Infrastructure for Usability Testing on Mobile Devices” [32] (derecha).

Figura 7. Prueba de Campo⁵.



Con respecto a la recolección de datos en las pruebas de campo, los investigadores han desarrollado nuevas técnicas, tales como diarios de correo de voz [26] y diarios web [27]. En el método que utiliza diario de correo de voz, los participantes llaman a una línea de correo de voz dedicado, para informar de problemas o sugerencias sobre el uso de la aplicación móvil. Mediante la información obtenida a través de los diarios de correo de voz, los evaluadores pueden analizar problemas reportados de usabilidad y las sensaciones de los participantes sobre la aplicación móvil. Para el método del diario web, los participantes tienen que anotar información detallada acerca de la aplicación móvil en sus dispositivos, luego en algún momento, conectarse a un servidor Web predeterminado y subir sus notas en un diario. Los evaluadores pueden obtener información del servidor durante el estudio. Si tienen consultas, pueden contactar a los participantes para obtener más detalles. Además de estos dos métodos, otros enfoques tales como reuniones periódicas, informes de correo electrónico o cuestionarios en línea pueden ser adoptados para la recolección de datos.

Pruebas remotas

En este tipo de pruebas, se separa en espacio y/o tiempo al evaluador del usuario, el evaluador se encuentra separado físicamente del usuario. Las pruebas pueden ser sincrónicas o asincrónicas. En las sincrónicas, la prueba y la transferencia de la información de usabilidad se hacen en tiempo real; el evaluador se encuentra separado en espacio del usuario. Las pruebas asincrónicas, significa que el evaluador y el usuario están separados tanto espacial como temporalmente; la diferencia con la prueba de característica sincrónica es que la transferencia de datos no se realiza en tiempo real, sino que la información se almacena en el dispositivo y en algún otro momento se transfiere al evaluador (Fig. 8).

⁵ Foto extraída de “Experimental Evaluation of Techniques for Usability Testing of Mobile Systems in Laboratory Setting” [21].

Figura 8. Prueba Remota.



En [28] se presenta un método para evaluar remotamente la usabilidad de las aplicaciones en los teléfonos móviles con sistema operativo Android, está diseñado con una estructura Cliente/Servidor. Este método utiliza la potencia de cómputo de los Smartphone para registrar las actividades del usuario, tales como las acciones, los sonidos y las pantallas. Los datos registrados se transfieren a través de una red Wi-Fi al evaluador de manera silenciosa sin intervención del usuario. Se realizaron experimentos y los resultados mostraron que el registro de datos puede trabajar de manera eficiente, dependiendo de la cantidad de datos recolectados y de la situación de la red Wi-Fi.

CONCLUSIONES

Las metodologías existentes para analizar la usabilidad de un producto de software no ofrecen una guía consolidada que sea precisa en definir una representación de usabilidad y de cómo medirla, esto incluye las aplicaciones móviles. Además, las actuales métricas y métodos definidos en distintos modelos de usabilidad no fueron desarrollados para ser aplicados en contextos móviles.

La usabilidad de las aplicaciones móviles está fuertemente influenciada por el contexto de uso, sobre este aspecto existen trabajos [23, 29, 30] que definen los posibles factores que pueden ser parte del contexto y que afectan la usabilidad de una aplicación. Pero definir una metodología general para evaluar cualquier tipo de aplicación móvil es una tarea difícil debido a la complejidad asociada con el contexto de uso.

Con el fin de entender completamente cómo un usuario interactúa con una aplicación, el evaluador debe analizar tanto las interacciones directas del usuario con el dispositivo y la aplicación así como el contexto externo que el usuario percibe durante su interacción [31].

Resulta complicado poder realizar pruebas de usabilidad en el entorno real de uso, debido a los problemas mencionados anteriormente. Un tema de investigación es lograr obtener los datos arrojados por la prueba y la información relacionada al contexto de forma automática y transparente para el usuario. Las métricas y la información del contexto se adquieren en el momento que el usuario interactúa con la aplicación, la información contextual resulta valiosa al analizar los datos recolectados de la prueba.

Para obtener información del contexto y complementar los datos de usabilidad obtenidos de la

pruebas, se ha propuesto la utilización de sensores incorporados a los dispositivos móviles, como por ejemplo un acelerómetro, que permite detectar y registrar información contextual relevante para las pruebas. Por ejemplo una métrica puede medir el tiempo empleado por un usuario al realizar una tarea y la información contextual indicar que el usuario cuando realizo la tarea se encontraba en movimiento, estos datos le permiten al evaluador realizar un análisis más preciso de la interacción entre el usuario y la aplicación.

Un método menos desarrollado consiste en realizar pruebas de usabilidad en situaciones reales de uso sin conocimiento por parte del usuario. Lo que se pretende es realizar las mediciones de las métricas y obtener información del contexto de uso de forma transparente para el usuario. Pero para satisfacer estas condiciones, por un lado se necesita que la tecnología avance hacia dispositivos móviles más potentes para realizar cálculos y que las redes inalámbricas permitan mayor capacidad de transferencia de datos. Por otro lado, se requieren nuevas técnicas y herramientas que permitan capturar información desde la perspectiva del usuario sin influir demasiado en la interacción con la aplicación.

Este tipo de métodos transparente para el usuario puede ser una forma práctica para evaluar la usabilidad de las aplicaciones móviles. Para que las herramientas que dan soporte a las pruebas de usabilidad resulten útiles para los evaluadores, estas deben obtener la información de usabilidad de manera remota, instantánea y automatizada.

REFERENCIA BIBLIOGRÁFICA

1. IEEE Standards Association (1998): IEEE Standard for a Software Quality Metrics Methodology. IEEE Std 1061-1998 (R2009).
2. Zhang, D., Adipat, B. (2005): Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. *International Journal of Human-Computer Interaction*, vol. 18, no. 3, pp. 293–308.
3. ISO International Standard, ISO 9241-11 (1998): Ergonomic requirements for office work with visual display terminals (VDTs.) - Part 11: Guidance on usability.
4. Frøkjær, E., Hertzum, M., Hornbæk, K.: Measuring Usability (2000): Are Effectiveness, Efficiency and Satisfaction Really Correlated?. *Proceedings of the CHI 2000, conference on Human factors in computing systems*, vol. 1, pp. 345–352. The Hague, The Netherlands.
5. Nielsen, J. (1992): The Usability Engineering Life Cycle. *IEEE Computer*, vol. 25, no. 3, pp. 12–22.
6. Folmer, E., Bosch, J. (2004): Architecting for usability: a survey. [*Journal of Systems and Software*](#), vol. 70, no. 1/2, pp. 61–78.
7. Hornbæk, K. (2006): Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, vol. 64, pp. 79–102.
8. DeMarco, T. (1986): *Controlling Software Projects: Management, Measurement and Estimation*. USA.
9. Tahir, A., Ahmad, R. (2010): An AOP-Based Approach for Collecting Software Maintainability Dynamic Metrics. *IEEE Second International Conference on Computer Research and Development*, pp. 168–172.
10. Ahmed, S., Mohammad, D., Rex, B. K., Harkirat, K. P. (2006): Usability measurement and metrics: A consolidated model. *Journal Software Quality Control*, vol. 14, pp. 159–178.

11. Macleod, M., Bowden, R., Bevan, N., Curson, I. (1997): The MUSiC Performance Measurement Method. *Behaviour and Information Technology*, vol. 16, no. 4/5, pp. 279–293.
12. Macleod, M.: Usability (1994): Practical Methods for Testing and Improvement. *Proceedings of the Norwegian Computer Society Software Conference*, pp. 1–4. Sandvika, Noruega.
13. Seffah, A., Kececi, N., Donyaee, M. (2001): QUIM: a framework for quantifying usability metrics in software quality models. *Conference on [Quality Software. Proceedings. Second Asia-Pacific](#), pp. 311–318.*
14. Ivory, M. Y., Hearst, M. A. (2001): The state of the art in automating usability evaluation of user interfaces. *Journal ACM Computing Surveys (CSUR)*, vol. 33, no. 4, pp. 470–516.
15. Chang, E., Dillon, T. S. (2006): A Usability-Evaluation Metric Base on a Soft-Computing Approach. *IEEE Transactions on Systems, Man, and Cybernetics, part A: Systems and Humans*, vol. 36, no. 2, pp. 356–372.
16. Macleod, M., Rengger, R. (1993): The Development of DRUM: A Software Tool for Video-assisted Usability Evaluation. *People and Computers VIII, Proceedings of The BCS HCI'93*. Loughborough, UK.
17. Sambasivan, D., John, N., Udayakumar, S., Gupta, R. (2011): Generic Framework for Mobile Application Development. *The Second Asian Himalayas International Conference on Internet (AH-ICI)*, pp. 1–5.
18. Tracy, K. W. (2012): Mobile Application Development Experiences on Apple's iOS and Android OS. *Potentials, IEEE*, vol. 31, no. 4, pp. 30–34.
19. Wasserman, A. I. (2010): Software Engineering Issues for Mobile Application Development. *FoSER 10, Proceedings of the FSE/SDP workshop on Future of software engineering research* pp. 397–400. USA.
20. Po, S., Howard, S., Vetere, F., Skov, M. B. (2004): Heuristic Evaluation and Mobile Usability: Bridging the Realism Gap. In *Mobile HCI*, pp. 49–60.
21. Beck, E., Christiansen, M., Kjeldskov, J., Kolbe, N., Stage, J. (2003): Experimental Evaluation of Techniques for Usability Testing of Mobile Systems in a Laboratory Setting. *Proceedings of OzCHI 2003*. Brisbane, Australia.
22. Hussain, A. (2012): Metric based evaluation of mobile devices: mobile goal question metric (mGQM), PhD thesis, Salford, University of Salford.
23. Coursaris, C. K., Kim, D. J. (2011): A Meta-Analytical Review of Empirical Mobile Usability Studies. *Journal of Usability Studies*, vol. 6, no. 3, pp. 117–171.
24. Hummel, K. A., Hess, A., Grill, T. (2008): Environmental Context Sensing for Usability Evaluation in Mobile HCI by Means of Small Wireless Sensor Networks. *MoMM '08, Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*, pp. 302–306.

25. Kjeldskov, J., Skov, M. B. (2003): Creating Realistic Laboratory Settings: Comparative Studies of Three Think-Aloud Usability Evaluations of a Mobile System. Proceedings of the 9th IFIP TC13 International Conference on Human-Computer Interaction, pp. 663–670.
26. Palen, L., Salzman, M. (2002): Voicemail diary studies for naturalistic data capture under mobile conditions. Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, pp. 87–95. New Orleans, Louisiana, USA.
27. Kim, H., Kim, J., Lee, Y., Chae, M., Choi, Y. (2002): An Empirical Study of the Use Contexts and Usability Problems in Mobile Internet. Proceedings of the 35th Hawaii International Conference on System Sciences (HICSS-35'02), pp. 1767–1776. Big Island, Hawaii.
28. Liang, H., Song, H., Fu, Y., Cai, X., Zhang, Z. (2011): A Remote Usability Testing Platform for Mobile Phones. IEEE International Conference on Computer Science and Automation Engineering (CSAE), vol. 2, pp. 312–316. Shanghai.
29. Maguire, M. (2001): Context of Use within usability activities. International Journal of Human Computer Studies, vol. 55, no. 4, pp. 453–483.
30. Munesh Chandra Trivedi and Mohammadi Akheela Khanum (2012): Role of Context in Usability Evaluations: A review. Advanced Computing: An International Journal (ACIJ), vol. 3, no. 2, pp. 69–78.
31. Lyons, K., Starner, T. (2001): Mobile Capture for Wearable Computer Usability Testing. ISWC '01, Proceedings of the 5th IEEE International Symposium on Wearable Computers, pp. 69–76.
32. Schusteritsch, R., Wei, C., LaRosa, M. (2007): Towards the Perfect Infrastructure for Usability Testing on Mobile Devices. CHI '07 Extended Abstracts on Human Factors in Computing Systems, pp. 1839–1844.