

# Tramea, a Tool to Support the Development of Final Year Projects. Usability Evaluation

Mikel Villamañe, Ainhoa Álvarez, Mikel Larrañaga, *Member, IEEE*, Begoña Ferrero

**Abstract**— Carrying out and monitoring Final Year Projects are challenging processes for all the actors involved: lecturers and students. Those projects exhibit problems regarding their evaluation, supervision and the lack of appropriate communication channels between the actors. The use of information and communication technologies can help to overcome those problems. This paper presents TraMeA, a system developed to support both lecturers and students and lighten the problems mentioned above. The usability evaluation conducted on the tool and its results are also described.

**Keywords**— Usability evaluation, educative software, Final Year Project

## I. INTRODUCCIÓN

CON la implantación del Espacio Europeo de Educación Superior (EEES), todos los planes de estudio conducentes a la obtención de un título universitario culminan con la realización y defensa de un Trabajo Fin de Grado (TFG) orientado a la evaluación de las competencias asociadas al título. El desarrollo de un TFG es un reto para todos los actores involucrados: estudiantes y docentes. Los autores de este artículo llevaron a cabo un estudio con estos actores con el objetivo de identificar y analizar los principales problemas que surgen en el proceso de realización, supervisión y evaluación del TFG [1]. A partir de la problemática detectada, identificaron las necesidades que debería cubrir una herramienta informática para dar un soporte adecuado al seguimiento, desarrollo y evaluación de los TFG [1].

Las necesidades identificadas incluían por un lado el establecimiento de criterios públicos y objetivos tanto para la aceptación de propuestas de TFG como para la posterior evaluación del trabajo realizado. Por otro lado, también incluían el establecimiento de un proceso de supervisión continuo con retroalimentación y de canales de comunicación adicionales entre alumnado y profesorado.

A continuación, se analizaron distintas herramientas ya existentes y, tras comprobar que ninguna daba una respuesta adecuada a todas estas necesidades [2], se decidió desarrollar el sistema TraMeA (Tracking, Mentoring and Assessment).

Dado que el éxito del software está directamente

relacionado con el grado de usabilidad del mismo [3], [4], se decidió evaluar la usabilidad del prototipo de TraMeA. Este tipo de evaluación permite evaluar la satisfacción de los usuarios [5] y observar cómo de intuitivo es el uso del sistema, los problemas que se encuentran los usuarios durante su utilización así como la formación previa requerida para usarlo correctamente

En este artículo primero se presenta el prototipo de TraMeA desarrollado. A continuación, se describe la evaluación de usabilidad realizada. Posteriormente, se analizan los resultados obtenidos en la evaluación y se finaliza presentando un conjunto de conclusiones y líneas futuras.

## II. TRAMEA

TraMeA es un sistema web que combina aspectos de gestión de proyectos, herramientas de comunicación y de evaluación para satisfacer todas las necesidades detectadas en el proceso de desarrollo, supervisión y evaluación de los TFG. Antes de proceder a su diseño y desarrollo, se analizaron distintas herramientas ya existentes en los ámbitos de gestión de proyectos, herramientas de comunicación y herramientas de evaluación [2].

Entre las herramientas de gestión de proyectos analizadas se encuentran Microsoft Project Server<sup>2</sup>, OpenProject<sup>3</sup> y Redmine<sup>4</sup> entre otras. Todas ellas ofrecen funcionalidades destinadas a gestionar proyectos, sin embargo carecen de funcionalidades orientadas a la evaluación del proceso de desarrollo de los mismos.

Las herramientas de comunicación estudiadas (correo electrónico, foros de comunicación, wikis y herramientas comerciales como Slack<sup>5</sup> y similares) se limitan a proporcionar canales de comunicación entre los actores sin dar respuesta al resto de necesidades identificadas.

En lo relativo a las herramientas de evaluación, se tuvieron en cuenta herramientas que permitieran el cálculo de una nota numérica en base a rúbricas de evaluación. Entre dichas herramientas se analizaron las hojas de cálculo de distintos paquetes ofimáticos, complementos de plataformas de aprendizaje como Evalcomix<sup>6</sup> o iRubric<sup>7</sup> y aplicaciones web como eRubrica<sup>8</sup> entre otras. Sin embargo, ninguna de las herramientas de evaluación estudiadas cubría los aspectos

---

M.V. Universidad del País Vasco UPV/EHU, Escuela de Ingeniería de Bilbao, [mikel.v@ehu.es](mailto:mikel.v@ehu.es)

A.A. Universidad del País Vasco UPV/EHU, Escuela Universitaria de Ingeniería de Vitoria-Gasteiz, [ainhoa.alvarez@ehu.es](mailto:ainhoa.alvarez@ehu.es)

M.L. Universidad del País Vasco UPV/EHU, Escuela Universitaria de Ingeniería de Vitoria-Gasteiz, [mikel.larranaga@ehu.es](mailto:mikel.larranaga@ehu.es)

B.F. Universidad del País Vasco UPV/EHU, Escuela de Ingeniería de Bilbao, [bego.ferrero@ehu.es](mailto:bego.ferrero@ehu.es)

---

<sup>2</sup> <https://products.office.com/es-es/project/enterprise-project-server>

<sup>3</sup> <https://www.openproject.org/>

<sup>4</sup> <http://www.redmine.org/>

<sup>5</sup> <https://slack.com/>

<sup>6</sup> <http://evalcomix.uca.es/>

<sup>7</sup> <http://www.rcampus.com/indexrubric.cfm>

<sup>8</sup> <http://erubrica.uma.es/>

relacionados con la gestión de proyectos ni la comunicación entre los distintos actores.

Dado que ninguna de las herramientas analizadas daba una respuesta adecuada a todas las necesidades detectadas, se estudiaron las posibilidades de integración y ampliación que ofrecían las distintas herramientas. En este aspecto destacó Redmine que era la más completa de las herramientas analizadas y la que menos modificaciones requería. Por ello, se decidió tomar Redmine como base de TraMeA, completándolo con la implementación e integración de aquellas funcionalidades de las que carecía.

La funcionalidad específica que fue necesario incluir en Redmine para dar soporte al seguimiento, al desarrollo y a la evaluación de los TFG fue la relacionada con los distintos aspectos de la evaluación (ver Fig. 1).



Figura 1. Visualización general de la evaluación un TFG.

Por ejemplo, se han incluido en TraMeA las interfaces necesarias para evaluar un TFG usando rúbricas (ver Fig. 2).

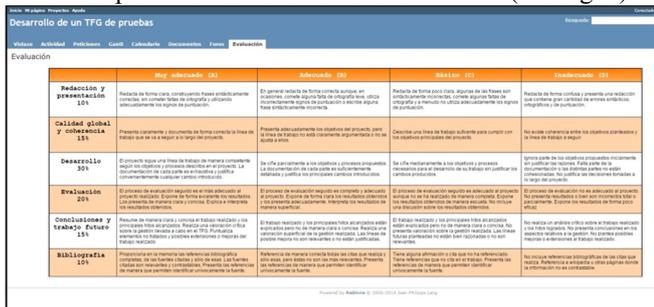


Figura 2. Interfaz de evaluación de un TFG usando rubricas.

Para la definición y posterior desarrollo de las funcionalidades necesarias, se siguió la metodología de desarrollo InterMod [6], [7]. InterMod es una metodología ágil para desarrollo de software interactivo de calidad, dirigida por modelos y centrada en el usuario. InterMod organiza el desarrollo de un proyecto mediante una serie de iteraciones donde se realizan diferentes actividades para alcanzar los Objetivos de Usuario, que se corresponden con los deseos o necesidades (funcionales y no funcionales) de los usuarios del sistema [8].

Durante la formalización y desarrollo de los Objetivos de Usuario, InterMod presta especial atención a la interfaz gráfica, su reacción a las acciones de los usuarios y al comportamiento del sistema en situaciones alternativas o erróneas. Estas características hacen de InterMod una metodología especialmente indicada para la elaboración de prototipos.

El prototipo de TraMeA fue desarrollado usando el

software Justinmind<sup>9</sup>, que permite generar prototipos interactivos funcionales basados en HTML y JavaScript. Una vez construido, el prototipo se desplegó en un servidor web para permitir el acceso remoto a los distintos usuarios.

A continuación, se detalla la evaluación de usabilidad que se ha realizado del prototipo de TraMeA, explicando el diseño de la misma y los resultados obtenidos.

### III. DISEÑO DE LA EVALUACIÓN

Dado que TraMeA está desarrollado incorporando nueva funcionalidad sobre Redmine, la evaluación del sistema se debe centrar en las nuevas funcionalidades añadidas tanto para el profesorado como para el alumnado. Sin embargo, la nueva funcionalidad asociada al alumnado coincide con parte de la añadida para el profesorado. Además, en sistemas educativos es imprescindible satisfacer las necesidades y deseos del profesorado para que el sistema tenga éxito [9]. Por estas razones, la evaluación del sistema que se presenta en este artículo se centra en el uso de TraMeA por parte del profesorado.

Al centrar la evaluación de usabilidad en el uso del sistema que realiza el profesorado, las tareas que formaron parte del estudio se determinaron teniendo en cuenta los distintos roles que este colectivo puede ejercer en el marco de los TFG: persona administradora, persona tutora o componente de tribunal.

El rol de persona administradora es el correspondiente a la persona encargada de gestionar todos los datos necesarios para el correcto funcionamiento del sistema. El rol de persona tutora es el que ejercen los miembros del profesorado que supervisan el desarrollo de los diferentes TFG. Finalmente, el rol de componente de tribunal corresponde a las personas que forman parte de los tribunales de evaluación correspondientes.

Debido a que los roles persona tutora y componente de tribunal tienen el 70% de los Objetivos de Usuario en común, ambos roles se combinaron para la evaluación de la usabilidad. De esta manera, la evaluación se realizó desde dos perspectivas diferentes. Primero se analizó la perspectiva de administración del sistema (rol persona administradora). Posteriormente se analizó la perspectiva de tutorización/evaluación de los proyectos (roles persona tutora y componente de tribunal).

Para realizar la evaluación de usabilidad de las dos perspectivas, se optó por el método del paseo cognitivo, (*Cognitive Walkthrough*) entre los diversos métodos existentes [10]. Este método consiste en proponer al usuario la realización de una o varias tareas y dejarle navegar por el sistema intentado solucionarlas [11].

El proceso de evaluación para cada una de las perspectivas se dividió en tres fases (ver Fig. 3). En la primera fase («Fase1: preparación estudio»), se definieron las tareas a realizar por los participantes en el estudio siguiendo las pautas del método del paseo cognitivo.

En la segunda fase («Fase2: recogida de datos») se solicitó a los participantes la ejecución de las tareas definidas en la fase

<sup>9</sup> <http://www.justinmind.com>

anterior (Fase 2a en Fig. 3) y se les pidió que respondieran a un cuestionario para recoger sus impresiones sobre el sistema (Fase 2b en Fig. 3).

Los datos recogidos durante la ejecución de las tareas y los resultados del cuestionario fueron sometidos a distintos análisis (Fase 3 en Fig. 3) cuyos resultados y conclusiones se presentan en el apartado IV de este artículo.

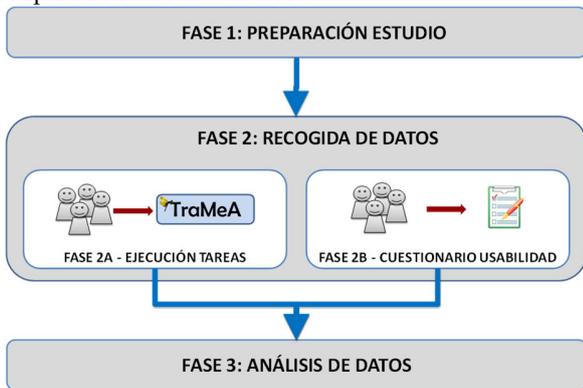


Figura 3. Proceso de evaluación.

### A. Preparación del estudio

El primer paso para una evaluación de usabilidad siguiendo el método del paseo cognitivo consiste en determinar el conjunto de tareas a evaluar.

Al utilizar InterMod como metodología de desarrollo, las tareas a evaluar se extrajeron de los Objetivos de Usuario identificados. En la TABLA I se recogen las tareas definidas junto con el código que las identifica y la perspectiva del estudio a la que pertenecen.

El código de las tareas está formado por las siglas CPU (Caso de Prueba de Usabilidad), seguido de un identificador numérico y unos caracteres que hacen referencia al rol de usuario encargado de su ejecución (A, persona administradora; T, persona tutora y Tr componente de tribunal).

TABLA I CÓDIGOS DE LAS TAREAS DEL ESTUDIO DE USABILIDAD

Código Tarea	Descripción Tarea	Perspectiva del estudio
CPU0_A	Añadir criterios de evaluación y alcance	Administración
CPU1_A	Asignar estudiante y tutorización	
CPU2_A	Finalizar proyecto	
CPU3_A	Asignar tribunal	
CPU4_T_Tr	Consultar planificación proyecto	Tutorización/ Evaluación
CPU5_T_Tr	Consultar actividad estudiantes	
CPU6_T_Tr	Consultar detalles tarea	
CPU7_T_Tr	Consultar documento	
CPU8_T	Dar retroalimentación	
CPU9_T	Consultar retroalimentación	
CPU10_T	Solicitar cierre	
CPU11_T_Tr	Evaluar un elemento	
CPU12_T_Tr	Consultar evaluación	

### B. Recogida de datos

En el estudio de usabilidad se recogieron dos tipos de datos: los relativos a la ejecución de las tareas y los obtenidos a partir de los cuestionarios de usabilidad (ver Fig. 3).

Para recoger las opiniones de los usuarios sobre el sistema se usó el cuestionario con la Escala de Usabilidad del Sistema (*System Usability Scale, SUS*). Esta escala es una herramienta eficaz y eficiente a la hora de evaluar la usabilidad de un sistema [12], [13].

Para facilitar la recogida de datos referentes a la ejecución de las tareas, se optó por realizar una evaluación remota controlada (*instrumented remote evaluation*) [14]. Esta manera de realizar la evaluación, facilita la participación ya que cada usuario puede ejecutar la prueba de manera independiente, en cualquier momento y sin necesidad de contar con un evaluador.

Para poder realizar este tipo de evaluación, se incorporó al prototipo una interfaz inicial con el listado de las posibles tareas a realizar en función de la perspectiva del prototipo que se esté evaluando (ver Fig. 4a).

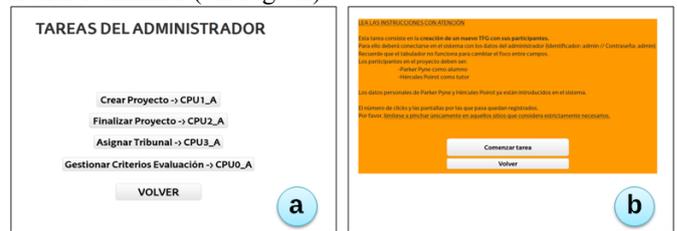


Figura 4. Interfaces gráficas de elección (a) y descripción (b) de tareas.

Para este estudio, se solicitó a cada participante que ejecutase las tareas en un orden determinado. Dicho orden se estableció de manera aleatoria para cada sujeto, con el ánimo de paliar la influencia que pudiera tener la ejecución previa de una tarea sobre otra a realizar a posteriori [15], [16] y el efecto que el cansancio o aburrimiento acumulado durante la prueba pudiera tener sobre las últimas tareas a realizar.

Una vez que el usuario selecciona una tarea, se le muestran las instrucciones para llevarla a cabo (Fig. 4b). Cuando considere que ha comprendido en qué consiste la tarea, puede comenzar su ejecución. A partir de ese momento, el usuario dispone de acceso completo al prototipo pudiendo navegar libremente e interactuar con todos los elementos del sistema que crea necesario para completar la tarea propuesta.

Para dar soporte a una evaluación remota controlada es necesario recoger los datos sobre la ejecución de las tareas de manera remota. En este caso, se incorporó al prototipo de TraMeA la capacidad de recoger de manera automática la información sobre la ejecución de las tareas.

Para cada tarea realizada por un usuario, el prototipo almacena la información referente a la ejecución de la misma, que incluye: para cada elemento de la interfaz el número de veces que el usuario lo ha pulsado, el momento de inicio y de finalización de la prueba para el cálculo del tiempo empleado y el camino seguido por el usuario durante la resolución de la tarea propuesta.

### C. Análisis de datos

La norma ISO 9241-11 define la usabilidad como «el grado en el que un producto puede ser utilizado por usuarios para lograr objetivos concretos con efectividad, eficiencia y satisfacción» [17]. Por lo tanto, los datos se analizaron a través de diversas métricas para evaluar la efectividad, eficiencia y el grado de satisfacción de los usuarios. Los datos de ejecución de las tareas permiten evaluar tanto la efectividad como la eficiencia del sistema; mientras que los de los cuestionarios permiten calcular el grado de satisfacción de los usuarios.

La *efectividad* del sistema se evaluó mediante la tasa de éxito de las tareas [16]. Se calculó la tasa de éxito muestral a partir de los datos recogidos de los participantes. Estos datos permiten estimar (mediante el método ajustado de Wald [18], [19]) el intervalo de confianza al 95% en el que se encontrará la tasa de éxito de la población general.

Para evaluar la *eficiencia* del sistema se analizaron primero los tiempos de ejecución de las tareas por parte de los participantes [20]. Con estos datos se estimó el tiempo que tardarán los nuevos usuarios en resolver las tareas, comparándolos a continuación con el tiempo máximo admisible para cada tarea. Este tiempo máximo se ha calculado, de acuerdo al método propuesto por Sauro & Kindlund [21], como el percentil 95 de los tiempos de ejecución de los participantes que finalizaron la tarea y valoraron satisfactoriamente el sistema. Por otro lado, se calculó y analizó el nivel de pérdida [22] de cada tarea y se elaboró un grafo que representa el recorrido que los usuarios realizaron a través del sistema para resolverla. Esta información permite analizar qué partes del sistema han resultado más confusas para los usuarios y en qué zonas se han perdido con mayor facilidad [23].

Finalmente, para calcular el grado de *satisfacción* de los usuarios con el sistema, se calculó la puntuación obtenida en los cuestionarios SUS [12]. Con los valores obtenidos a través de los cuestionarios se computó la media aritmética de las puntuaciones obtenidas y se construyó un intervalo de confianza al 95% para estimar la puntuación poblacional que se obtendría.

### D. Participantes

Se formaron dos grupos de usuarios, el primero para evaluar la perspectiva de administración de TraMeA y el segundo para la perspectiva de tutorización/evaluación.

Todos los sujetos participantes en el estudio eran personal docente e investigador perteneciente a la UPV/EHU y relacionados con el campo de la Ingeniería Informática. Los participantes no tenían conocimientos previos del sistema, pero sí experiencia en la tutorización y evaluación de Trabajos Fin de Grado y en el desarrollo de proyectos en general.

El número de sujetos en cada grupo fue fijado en 12 ya que una población de ese tamaño proporciona una fiabilidad de entre el 90% y el 100% a los resultados obtenidos a través del SUS [24].

## IV. ANÁLISIS DE RESULTADOS

A continuación se detallan los resultados obtenidos en los tres aspectos relacionados con la usabilidad analizados.

### A. Efectividad

La efectividad se ha evaluado calculando las tasas de éxito muestrales y poblacionales. En ambos casos, el valor estándar para considerar una tasa de éxito satisfactoria es 0.7 [25].

La TABLA II recoge la tasa de éxito muestral de cada tarea, que en todos los casos es superior al 0.7. Además, el 75% de tareas presenta una tasa de éxito muestral de 1. Es decir, todos los participantes en el estudio finalizaron satisfactoriamente la tarea encomendada.

A partir de estos resultados se calcularon los límites inferior y superior del intervalo de confianza al 95% para la estimación de la tasa de éxito poblacional (ver TABLA II).

De acuerdo a las tasas de éxito estimadas, en 9 de las tareas el porcentaje mínimo de usuarios que se espera sea capaz de resolverlas satisfactoriamente es de un 78.4% (transformando el límite inferior del intervalo en un porcentaje).

Para dos de las tareas (CPU0\_A y CPU9\_T), reduciendo la confianza del intervalo al 85%, se obtiene un límite inferior de 0.7203, por lo que se puede afirmar con una confianza del 85% que dichas tareas también superarán el 0.7 establecido como mínimo aceptable.

Por lo tanto, sólo hay una tarea (CPU10\_T) que presenta una estimación de la tasa de éxito con un límite inferior por debajo del 0.7 establecido como referencia incluso reduciendo la confianza del intervalo. Estos resultados ocasionan que haya que tenerla especialmente en cuenta en los siguientes apartados para comprender las razones de los mismos.

TABLA II TASAS DE ÉXITO

Tarea	Muestral	Estimación Poblacional	
		Límite inferior	Límite superior
CPU0_A	0.917	0.625	0.999
CPU1_A	1	0.784	1
CPU2_A	1	0.784	1
CPU3_A	1	0.784	1
CPU4_T_Tr	1	0.784	1
CPU5_T_Tr	1	0.784	1
CPU6_T_Tr	1	0.784	1
CPU7_T_Tr	1	0.784	1
CPU8_T	1	0.784	1
CPU9_T	0.917	0.625	0.999
CPU10_T	0.75	0.461	0.917
CPU11_T_Tr	1	0.784	1
CPU12_T_Tr	1	0.784	1

Profundizando en los resultados obtenidos, se analizó si las tareas se realizaron a la perfección, sin el más mínimo error y se estimó el porcentaje de futuros usuarios que resolverán las tareas de forma óptima la primera vez que se enfrenten a ellas. En esta situación, los datos indican que como mínimo el 80% de los usuarios serán capaces de resolver sin errores 9 de las 13 tareas planteadas.

### B. Eficiencia

Para analizar la eficiencia se han tenido en cuenta los tiempos de ejecución y los niveles de pérdida.

Para los tiempos de ejecución se calcularon la mediana de los tiempos de los participantes, el tiempo máximo admisible (el percentil 95 de los tiempos de los participantes que valoraron positivamente el sistema), y el intervalo de confianza de la población (ver TABLA III). En ambas perspectivas los límites superiores de los intervalos de confianza de todas las tareas son inferiores al tiempo máximo admisible de ejecución. Por lo tanto, se puede afirmar con una confianza del 95% que todos los usuarios serán capaces de completar cualquiera de las tareas en un tiempo razonable.

TABLA III TIEMPO DE EJECUCIÓN DE LAS TAREAS

Tarea	Muestral	Tiempo máximo admisible	Estimación Poblacional	
			Límite inferior	Límite superior
CPU0_A	148	187.25	118	186
CPU1_A	113	242.20	85	150
CPU2_A	39	63.30	27	56
CPU3_A	84	772.75	43	164
CPU4_T_Tr	11	58.8	6	19
CPU5_T_Tr	13	60.8	8	21
CPU6_T_Tr	46	167	26	83
CPU7_T_Tr	34	170	20	56
CPU8_T	85	216.2	48	149
CPU9_T	31	57	21	45
CPU10_T	47	161.9	20	110
CPU11_T_Tr	51	86.4	41	64
CPU12_T_Tr	31	63	21	43

A continuación se analizaron los niveles de pérdida muestrales y poblacionales para cada tarea (ver TABLA IV). Teniendo en cuenta que sólo se consideran problemáticos los niveles de pérdida superiores a 0.4 [22], a excepción de la tarea CPU10\_T (que está en el límite), todas las tareas presentan resultados excelentes en el nivel de pérdida muestral.

Se analizaron también los límites inferior y superior del intervalo de confianza calculado para estimar el nivel de pérdida de la población general. Como se puede observar (ver TABLA IV), los límites superiores del intervalo de confianza son inferiores a 0.4 para todas las tareas salvo CPU10\_T, por lo que se puede afirmar con un 95% de confianza que, incluso en el peor de los casos, los nuevos usuarios no se perderán para realizar las tareas contempladas en el estudio.

Teniendo en cuenta estos resultados y los de efectividad, se analizó el grafo con el recorrido que siguieron los participantes para resolver la tarea CPU10\_T. En este grafo se observó que el problema fue que los participantes tuvieron dificultades para localizar una de las opciones del menú y se dedicaron a navegar por el sistema intentando encontrarla a pesar de haber pasado en varias ocasiones por la página donde se encontraba dicha opción. De estos resultados se deriva que para mejorar la usabilidad de esta tarea se debe retocar la interfaz para ubicar mejor las opciones correspondientes.

TABLA IV NIVEL DE PÉRDIDA DE LAS TAREAS

Tarea	Muestral	Estimación Poblacional	
		Límite inferior	Límite superior
CPU0_A	0.13	0.054	0.215
CPU1_A	0.11	0.017	0.198
CPU2_A	0.06	-0.073 <sup>a</sup>	0.195
CPU3_A	0.07	-0.036 <sup>a</sup>	0.176
CPU4_T_Tr	0	0	0
CPU5_T_Tr	0.03	-0.04 <sup>a</sup>	0.107
CPU6_T_Tr	0.12	0.018	0.222
CPU7_T_Tr	0.02	0.169	0.264
CPU8_T	0.17	0.047	0.283
CPU9_T	0.1	-0.052 <sup>a</sup>	0.252
CPU10_T	0.4	0.155	0.644
CPU11_T_Tr	0.03	-0.03 <sup>a</sup>	0.08
CPU12_T_Tr	0	0	0

<sup>a</sup> Los valores negativos son interpretados como cero

### C. Satisfacción

La satisfacción de los usuarios se evaluó mediante los resultados obtenidos en el SUS.

Para categorizar los datos numéricos obtenidos se utilizó la escala definida por [26], según la cual a partir de una puntuación de 51, el grado de satisfacción con la usabilidad del sistema es *bueno*; a partir de 71 puntos se considera *excelente* y a partir de 85 la usabilidad es la *mejor imaginable* (ver Fig. 5).

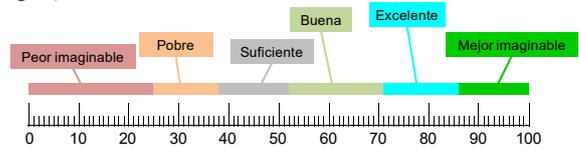


Figura 5. Escala de usabilidad. Adaptada de [26].

En la TABLA V se muestran, para cada perspectiva analizada, los resultados obtenidos en base a los participantes en el estudio (la satisfacción muestral) y los límites del intervalo de confianza para la estimación de la satisfacción de la población.

TABLA V SATISFACCIÓN DE LOS USUARIOS

Perspectiva de TraMeA	Muestral	Estimación Poblacional	
		Límite inferior	Límite superior
Administración	80.21	70.096	89.453
Tutorización / Evaluación	79.17	66.657	91.676

Como se puede observar, los resultados obtenidos muestran que la satisfacción de los futuros usuarios en cualquiera de las perspectivas de TraMeA se enmarcará entre las categorías *excelente* y *mejor imaginable* de acuerdo a la escala representada en la Fig. 5.

## V. CONCLUSIONES

En este artículo se ha presentado TraMeA, una herramienta desarrollada para dar respuesta a las necesidades detectadas en la supervisión y evaluación de los TFG. Para su desarrollo, se

ha tomado como base el sistema Redmine y se han ampliado sus funcionalidades mediante el diseño y desarrollo de los complementos necesarios para poder realizar la evaluación de los TFG mediante rúbricas. En este trabajo también se presenta la evaluación de usabilidad realizada de las perspectivas de administración y de tutorización/evaluación de TraMeA. Para llevar a cabo esta evaluación se ha utilizado el método del paseo cognitivo y se han evaluado los tres factores implicados en la usabilidad: efectividad, eficiencia y satisfacción. En los tres aspectos se han obtenido resultados excelentes, tal y como se resume a continuación.

En cuanto a la efectividad, más del 92% de las tareas superan la tasa de éxito estándar (0.7) [25].

En cuanto a la eficiencia del sistema, todas las tareas presentan unos tiempos de ejecución inferiores al tiempo máximo admisible de ejecución. Además, los niveles de pérdida son muy bajos, lo que augura que los futuros usuarios del sistema serán capaces de utilizarlo sin perderse por el mismo.

La satisfacción de los usuarios respecto a TraMeA demuestra que su usabilidad es como mínimo *excelente* de acuerdo a la escala definida por [26].

Por lo tanto, a la vista de los resultados en todos los aspectos de usabilidad evaluados, se puede concluir que, salvo los pequeños ajustes de la interfaz gráfica mencionados, el diseño de TraMeA es bueno y los usuarios serán capaces de utilizarlo para lograr objetivos concretos con efectividad, eficiencia y satisfacción [17].

Actualmente se está trabajando en la implantación de TraMeA teniendo en cuenta los resultados obtenidos en la evaluación de usabilidad. De cara al futuro, se planea implantarlo en el Grado en Ingeniería Informática de Gestión y Sistemas de Información que se imparte en la UPV/EHU y en el Grado en Criminología de la misma Universidad.

Asimismo, también se está diseñando el estudio de usabilidad de la parte de TraMeA correspondiente al alumnado.

#### AGRADECIMENTOS

Los autores desean agradecer su participación a todas las personas implicadas en el estudio de usabilidad. Este trabajo ha sido subvencionado por la Universidad del País Vasco UPV/EHU (EHUA16\_12, GIU16/20 y UF111/45).

#### REFERENCIAS

- [1] M. Villamañe, B. Ferrero, A. Álvarez, M. Larrañaga, A. Arruarte, y J. A. Elorriaga, «Dealing with common problems in engineering degrees' Final Year Projects», en *Proceedings of the 44th Annual IEEE Frontiers in Education (FIE) Conference*, Madrid, 2014, pp. 2663-2670.
- [2] M. Villamañe, B. Ferrero, y A. Álvarez, «Uso de las tic en el seguimiento y evaluación del trabajo fin de grado en el campo de las ingenierías», en *Actas del I Congreso Interuniversitario sobre el Trabajo de Fin de Grado. Retos y oportunidades del TFG en la sociedad del conocimiento*, Servicio Editorial UPV/EHU, 2014, pp. 222-233.
- [3] A. Hosseini-Khayat, T. D. Hellmann, y F. Maurer, «Distributed and Automated Usability Testing of Low-Fidelity Prototypes», 2010, pp. 59-66.
- [4] D. J. Mayhew, «The usability engineering lifecycle», 1999, p. 147.
- [5] K. E. Wiegers y J. Beatty, *Software requirements*, Third edition. Redmond, Washington: Microsoft Press, s division of Microsoft Corporation, 2013.
- [6] W.-S. Lin y C.-H. Wang, «Antecedences to continued intentions of adopting e-learning system in blended learning instruction: A contingency framework based on models of information system success and task-technology fit», *Comput. Educ.*, vol. 58, n.º 1, pp. 88-99, ene. 2012.
- [7] Losada, Begoña, «InterMod: Un enfoque ágil, dirigido por modelos y centrado en el usuario, para desarrollar aplicaciones interactivas», Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), 2014.
- [8] B. Losada, M. Urretavizcaya, y I. Fernández-Castro, «A guide to agile development of interactive software with a "User Objectives"-driven methodology», *Sci. Comput. Program.*, vol. 78, n.º 11, pp. 2268-2281, nov. 2013.
- [9] J. M. L. Gil, M. U. Loinaz, B. Losada, y I. F. Castro, «Field vs. Laboratory Usability Evaluations: a Study on a Context Dependent Mobile Application Developed with an Agile Methodology», *IEEE Lat. Am. Trans.*, vol. 14, n.º 1, pp. 339-348, Enero 2016.
- [10] J. Nielsen y R. L. Mack, Eds., *Usability inspection methods*. New York: Wiley, 1994.
- [11] C. Wharton, J. Rieman, C. Lewis, y P. Polson, «The cognitive walkthrough method: a practitioner's guide», en *Usability inspection methods*, J. Nielsen y R. L. Mack, Eds. New York, NY, USA: John Wiley & Sons, Inc., 1994, pp. 105-140.
- [12] J. Brooke, «SUS: A quick and dirty usability scale», en *Usability evaluation in industry*, P. W. Jordan, B. Weerdmeester, A. Thomas, y I. L. Mclelland, Eds. London: Taylor and Francis, 1996.
- [13] J. Brooke, «SUS: A Restrospective», *J. Usability Stud.*, vol. 8, n.º 2, pp. 29-40, 2013.
- [14] H. R. Hartson, J. C. Castillo, J. Kelso, y W. C. Neale, «Remote evaluation: the network as an extension of the usability laboratory», 1996, pp. 228-235.
- [15] B. Losada, M. Urretavizcaya, J.-M. López-Gil, y I. Fernández-Castro, «Combining InterMod agile methodology with usability engineering in a mobile application development», 2012, pp. 1-8.
- [16] T. Tullis y B. Albert, *Measuring the user experience: collecting, analyzing, and presenting usability metrics*, Second edition. Amsterdam ; Boston: Elsevier/Morgan Kaufmann, 2013.
- [17] ISO, «ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability». 1998.
- [18] A. Agresti y B. A. Coull, «Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions», *Am. Stat.*, vol. 52, n.º 2, pp. 119-126, may 1998.
- [19] J. Sauro y J. R. Lewis, «Estimating Completion Rates from Small Samples Using Binomial Confidence Intervals: Comparisons and Recommendations», *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 49, n.º 24, pp. 2100-2103, sep. 2005.
- [20] J. Sauro y J. R. Lewis, «Average task times in usability tests: what to report?», 2010, p. 2347.
- [21] J. Sauro y E. Kindlund, «How Long Should a Task Take? Identifying Specification Limits for Task Times in Usability Tests», presentado en 11th International Conference on Human-Computer Interaction, Las Vegas, Nevada, USA, 2005.
- [22] P. A. Smith, «Towards a practical measure of hypertext usability», *Interact. Comput.*, vol. 8, n.º 4, pp. 365-381, dic. 1996.
- [23] M. Y. Ivory y M. A. Hearst, «The state of the art in automating usability evaluation of user interfaces», *ACM Comput. Surv.*, vol. 33, n.º 4, pp. 470-516, dic. 2001.
- [24] T. Tullis y J. N. Stetson, «A comparison of questionnaires for assessing website usability», en *Proceedings of the 13th UPA Conference*, Minneapolis, 2004.
- [25] J. Rubin y D. Chisnell, *Handbook of usability testing: how to plan, design, and conduct effective tests*, 2nd ed. Indianapolis, IN: Wiley Pub, 2008.
- [26] A. Bangor, P. Kortum, y J. Miller, «Determining what individual SUS scores mean: Adding an adjective rating scale», *J. Usability Stud.*, vol. 4, n.º 3, pp. 114-123, may 2009.