

Construcción y Validación del cuestionario: Métrica de Calidad de Credibilidad e Interacción de Cursos de Teleformación

Marcelo Pérez Ibarra, Nilda Pérez Otero, Sandra Mendez, Adelina García, Ma. del Pilar Gálvez Díaz, Viviana E. Quincoces, Héctor Liberatori, Beatriz Fiorito, Cecilia María Lasserre.

Facultad de Ingeniería - Universidad Nacional de Jujuy - S.S. de Jujuy, Argentina

Resumen

La Ingeniería Web es el proceso con el que se crean aplicaciones Web de alta calidad. Al igual que cualquier producto software el desarrollo de WebApps necesita incluir actividades de garantía de calidad. Para asegurar la calidad se debe planificar la evaluación y el control de los productos intermedios hasta los productos finales.

Las WebApps aplicadas en educación a distancia dieron lugar a lo que se denomina e-learning. Considerando, que gran parte de las propuestas sobre evaluación de software educativo, son de índole cualitativo o necesitan por su complejidad de evaluadores expertos surgió desde el Grupo de Ingeniería de Software (GIS) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy, Argentina, el desarrollo de una Métrica para determinar Calidad de Credibilidad e Interacción de Cursos de Teleformación (MECACIN), de fácil aplicación por personas con formación equivalente a un nivel de instrucción medio finalizado, no necesariamente especializados en cuestiones informáticas ni de teleformación en forma específica, pero sí familiarizados con la navegación en INTERNET

El objetivo del presente trabajo es exponer el proceso de desarrollo seguido por el GIS para el diseño y evaluación de MECACIN y el modelo que la sustenta.

Palabras clave: Calidad del software - Teleformación – Métrica – Confiabilidad del cuestionario

1. Introducción

La Ingeniería Web (IWeb) es el proceso con el que se crean aplicaciones Web (WebApps) de alta calidad [1]. Si bien la Ingeniería Web incorpora muchos de los conceptos y principios fundamentales de la Ingeniería del Software, los sitios y WebApps como producto o producto en uso tienen sus propias características que

son sensiblemente diferentes del software tradicional [2]. En base a estas características se identifica una gran variedad de aplicaciones Web: informativas, de descarga, personalizables, de interacción, de entrada de usuario, orientada a transacciones, orientada a servicios, portal, acceso a una base de datos y almacén de datos [3].

La calidad del producto software es un tema que se ha venido tratando desde hace varios años pero, por lo general, para cada proyecto se adoptaban modelos de calidad diferentes [4].

En 1991 se publicó la primera edición del estándar de calidad ISO/IEC 9126 [5], con la intención de que se convirtiera en referencia para todos los desarrollos posteriores. Este estándar, en sus cuatro partes [6, 7, 8, 9], brinda modelos para evaluar la calidad externa, interna y en uso de un producto software.

Al igual que cualquier producto software el desarrollo de WebApps necesita incluir actividades de garantía de calidad. Para asegurar la calidad se debe planificar la evaluación y el control de los productos intermedios hasta los productos finales. ISO/IEC 9126 es utilizado para la formulación de modelos de calidad para diferentes aplicaciones Web [10, 11, 2].

Las WebApps aplicadas en educación a distancia dieron lugar a lo que se denomina e-learning. Según [12], e-learning se define como el uso sistemático de tecnologías multimediales basadas en computadoras para potenciar a los estudiantes, mejorar el aprendizaje, conectar a los estudiantes con personas y recursos que ayuden a sus necesidades, y a conjugar aprendizaje con rendimiento y metas individuales con metas organizacionales.

Los estudios de calidad de aplicaciones Web, en su mayoría, están orientados al comercio electrónico u otros tipos de sitios como revistas, canales de televisión o sitios de entretenimiento; por lo general financiados

por grandes organizaciones. En particular el tema de la evaluación de software educativo ha sido estudiado y documentado por diversos autores del ámbito educacional, proporcionando medidas de evaluación en el área [13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. Sin embargo, en vista de que gran parte de las propuestas sobre evaluación de software educativo, son de índole cualitativo o necesitan por su complejidad de evaluadores expertos surgió desde el Grupo de Ingeniería de Software (GIS) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy, Argentina, el desarrollo de una Métrica para determinar Calidad de Credibilidad e Interacción de Cursos de Teleformación* (MECACIN), de fácil aplicación por personas con formación equivalente a un nivel de instrucción medio finalizado, no necesariamente especializados en cuestiones informáticas ni de teleformación en forma específica [23, 24, 25], pero sí familiarizados con la navegación en INTERNET.

El objetivo del presente trabajo es exponer el proceso de desarrollo seguido por el GIS para el diseño y evaluación de MECACIN y el modelo que la sustenta.

El trabajo se estructura de la siguiente manera: en el apartado 2 se presenta el Modelo de Calidad de Credibilidad e Interacción de Cursos de Teleformación (MOCACIN) que sustenta la métrica; en el apartado 3 se presenta el diseño y refinamiento del instrumento de medición; en el apartado 4 se presentan los resultados de la evaluación de Consistencia Interna del instrumento de medición. Por último las conclusiones, referencias y el anexo se consignan.

2. Modelo de Calidad de Credibilidad e Interacción de Cursos de Teleformación

En el contexto de la Articulación Educación Media-Universidad 2004-2005 en la provincia de Jujuy, Argentina; un grupo de docentes de enseñanza media plantea a la Universidad Nacional de Jujuy, la siguiente situación: que al tratar de incluir cursos de teleformación en su currícula, no contaba con medios para determinar con certeza ciertos aspectos como:

- calificaciones de los autores del contenido,
- veracidad de los contenidos,

- actualidad de los contenidos,
- secuencia de aprendizaje para el usuario,
- facilidad de navegación y
- correctitud de la edición.

Aspectos ellos, que le permitirían hacer una rápida selección de cursos de fácil y accesible navegación y de procedencia fiable; a fin de analizar luego en profundidad si son adecuados o no para su inclusión en la currícula. A partir de este planteo la Universidad solicita al GIS dé solución a la situación expuesta.

Para comenzar la formulación del constructo, el GIS recopiló información acerca de aplicaciones e-learning, educación a distancia, estándares de calidad existentes, publicaciones sobre Ingeniería Web (e-commerce, entertainment, etc.) y técnicas de formulación y evaluación de cuestionarios.

El material estudiado permitió determinar:

1. que las aplicaciones e-learning consideradas de interés por los docentes mencionados se clasifican dentro de la categoría WebApps Informacionales, debido a que únicamente presentan contenido de sólo lectura con navegación y enlaces simples [3],
2. que un curso de e-learning es en definitiva un producto software al que se le puede aplicar los modelos de calidad establecidos en el estándar ISO/IEC 9126, adaptados al proceso IWeb, y
3. que el estándar ISO/IEC 14598 [26], presenta un conjunto de pasos orientados a la evaluación de calidad.

El estándar ISO/IEC 9126 no contempla en su modelo de calidad la categoría *contenido* que se considera esencial para lo que se pretende medir, por lo tanto, siguiendo los pasos indicados en el estándar ISO/IEC 14598, el GIS formuló un modelo de calidad mixto (MOCACIN) que contempla:

1. algunas categorías y subcategorías de calidad externa definidas en el estándar ISO/IEC 9126,
2. la categoría *contenido* que incluyen Olsina [27] y Fernández Nodarse [11] al trabajar sobre las WebApps que, en el contexto de este trabajo, hace referencia sólo a texto, imágenes (gráficos, figuras, fotografías, etc.) y audio, y
3. una subcategoría accesibilidad que se considera imprescindible dentro del tipo específico de WebApp apuntado (cursos de teleformación).

La Tabla 1 presenta las categorías y subcategorías consideradas en la formulación del modelo o árbol de requisitos.

* La **Teleformación** es un sistema de formación a distancia apoyado en las Tecnologías de la Información y las Comunicaciones (tecnologías, redes de telecomunicación, videoconferencias, TV digital, materiales multimedia), que combina distintos elementos pedagógicos: instrucción clásica (presencial o autoestudio), las prácticas, los contactos en tiempo real (presenciales, videoconferencias o chats) y los contactos diferidos (tutoriales, foros, correo electrónico) (FUNDESCO, 1998).

Tabla 1. Categorías y subcategorías de MOCACIN.

CATEGORÍA	SUBCATEGORÍA
FUNCIONALIDAD	<i>EXACTITUD</i>
	<i>IDONEIDAD/ADECUACIÓN</i>
	<i>SEGURIDAD DE LAS TAREAS</i>
USABILIDAD	<i>COMPRESIBILIDAD</i>
	<i>OPERABILIDAD</i>
EFICIENCIA	<i>COMPORTAMIENTO EN EL TIEMPO</i>
	<i>UTILIZACIÓN DE RECURSOS</i>
	<i>ACCESIBILIDAD</i>
CONTENIDO	<i>EXACTITUD DE INFORMACIÓN</i>
	<i>ADECUACIÓN DE LA INFORMACIÓN</i>
	<i>LEGALIDAD</i>

A continuación se define cada una de las categorías y subcategorías consideradas:

Funcionalidad: la capacidad del producto software de proveer funciones establecidas que responden a las necesidades del usuario cuando el software se usa bajo condiciones especificadas.

Exactitud: la capacidad del producto software para proporcionar los resultados o efectos correctos o esperados con el grado necesario de precisión.

Idoneidad/Adecuación: la capacidad del producto software para proporcionar un conjunto apropiado de funciones para tareas y objetivos de usuario especificados.

Seguridad de las Tareas: la capacidad del producto software para proteger información y datos de personas o sistemas no autorizados a leerlos o modificarlos y no negar el acceso a personas o sistemas autorizados.

Usabilidad: la capacidad del software para ser comprendido, aprendido, usado y atractivo para el usuario cuando se usa bajo condiciones especificadas.

Comprensibilidad: la capacidad del producto software que permite al usuario comprender si el software es adecuado y cómo puede utilizarse para tareas y condiciones particulares de uso.

Operabilidad: la capacidad del producto software para permitir al usuario operarlo y controlarlo.

Eficiencia: la capacidad del producto software para proveer una performance apropiada en relación a la cantidad de recursos usados bajo condiciones establecidas.

Comportamiento en el Tiempo: la capacidad del producto software de proporcionar tasas de

rendimiento efectivo, tiempos de procesamiento y respuestas apropiadas cuando realiza una función bajo condiciones establecidas.

Utilización de Recursos: la capacidad del producto software para usar cantidades y tipos de recursos apropiados cuando el software ejecuta su función bajo condiciones establecidas.

Accesibilidad: enfatiza la importancia de los aspectos técnicos de la WebApp para hacer a los contenidos más accesibles a usuarios con características tecnológicas y/o personales diferentes.

Contenido: la capacidad del producto software para proveer información exacta, adecuada, accesible y legal.

Exactitud de la Información: hace referencia a que la información es correcta, no ambigua, autorizada, objetiva y verificable.

Adecuación de la Información: hace referencia a que la información es apropiada (cobertura apropiada para la audiencia objetivo), completa (cantidad relevante), concisa (breve en el modo de expresar las ideas) y actual.

Legalidad: la capacidad de la información para adherir a estándares, convenciones y normas legales relacionados a contenidos y derechos de propiedad intelectual.

Una vez planteadas las categorías y subcategorías del árbol de requisitos se explicitan los siguientes conceptos a aplicar:

Necesidad de Información: determinar en cursos de teleformación las calificaciones de los autores del contenido, la veracidad de los contenidos, la actualidad de los contenidos, la secuencia de aprendizaje para el usuario, la facilidad de navegación y la correctitud de la edición.

Concepto Calculable: credibilidad e interacción de cursos de teleformación. Entendiendo por calidad de credibilidad a la posibilidad determinar con certeza quien acredita el curso, quien es su responsable y si se siguieron normas de edición (ortografía, gramática, dicción, legibilidad), y calidad de interacción a la facilidad de navegación en el curso (velocidad de acceso a las páginas, indicación de cómo volver a la página anterior, ir a la siguiente, vincularse con otras páginas, fondo de páginas) y comprensión de su estructura (títulos, objetivos, audiencia).

Métrica: se utilizan métricas directas para determinar la presencia o ausencia de atributos y métricas indirectas para determinar la calidad global.

Para cada subcategoría del árbol de requisitos se identifican los atributos que serán medidos durante la evaluación de calidad de los cursos de teleformación.

Definida la necesidad de información, el concepto calculable e identificados los atributos de interés, se eligió como instrumento de medición un cuestionario con las siguientes características: directo, precodificado, de preguntas cerradas y policotómicas (1, 0 o X), multidimensional (cinco dimensiones) y de puntuación simple o no ponderado.

3. Instrumento de Medición

En base a opiniones de expertos en desarrollo de cursos de educación a distancia, estándares disponibles y la guía de diseño de comercio electrónico elaborada por Alexander y Tate [28] se diseñó inicialmente un cuestionario compuesto por 71 preguntas, basadas en heurísticas [24].

A continuación se muestran (Tabla 2) dos de los atributos correspondientes a la subcategoría *Operabilidad*, categoría *Usabilidad* y sus heurísticas asociadas.

Tabla 2. Correspondencia Categoría-Subcategoría Atributo-Heurística.

CATEGORÍA	SUBCATEGORÍA	ATRIBUTO	PREGUNTA
Usabilidad	Operabilidad	Vínculo ATRÁS y ADELANTE	Para sitios estructurados jerárquicamente, ¿se incluye en la página un vínculo a la página que se encuentra en un nivel superior y/o inferior en la jerarquía?
		Vínculo a Mapa de Sitio, Índice o Tabla de Contenido	¿Incluye la página un vínculo a un mapa de sitio, índice o tabla de contenido?

El cuestionario inicial se sometió a un test preliminar o investigación de ensayo para identificar lo siguiente:

- tipos de preguntas más adecuados,
- si el enunciado es correcto y comprensible,
- si las preguntas tienen la extensión adecuada,
- si es correcta la categorización de las respuestas,
- si existen resistencias psicológicas o rechazo hacia algunas preguntas,

- si el ordenamiento interno es lógico y
- si la duración está dentro de lo aceptable por los encuestados.

El test preliminar se llevó a cabo mediante la aplicación del cuestionario a un grupo de expertos en educación a distancia y semipresencial, y una encuesta sobre la comprensión de las preguntas.

Esta investigación de ensayo determinó que algunas de las preguntas requerían, por parte de los usuarios, conocimiento de la terminología informática. Además miembros del grupo consultado indicaron que la interpretación de algunas preguntas estaba asociada a apreciaciones personales (por ejemplo, ¿Está claro qué tópicos son incluidos en el curso?) o evaluaban más de un atributo (por ejemplo, Si hay gráficos, cuadros o tablas, ¿están etiquetados y fáciles de leer?).

Esta situación condujo a la reformulación del cuestionario y la modificación de la redacción de las preguntas que presentaron los problemas consignados. Además se adicionaron notas al pie para aclarar el significado de términos técnicos. La tabla 3 muestra algunos ejemplos de los problemas reportados y las modificaciones realizadas al cuestionario inicial.

Tabla 3. Problemas y reformulaciones del cuestionario inicial.

Pregunta Original	Problema Reportado	Pregunta Reformulada/ Nota al pie
Si hay gráficos, cuadros o tablas, ¿están etiquetados y fáciles de leer?	Evaluación de más de un atributo. Interpretación sujeta a apreciación personal.	Si hay imágenes (gráficos, figuras, fotografías, etc.) ¿tienen su correspondiente epígrafe? Si hay imágenes. ¿se las observa con nitidez? Si hay tablas, ¿tienen su correspondiente epígrafe? Si hay tablas, ¿son legibles?
¿Incluyen todas las páginas subordinadas un vínculo a la página principal?	Término Informático: Página Principal	<i>Página Principal:</i> Es la página desde la cual se puede acceder a todos los ítems del curso (bibliografía, contenidos, actividades, evaluación, etc.). Puede estar precedida de una página de presentación.
¿Está claro qué tópicos son incluidos en el curso?	Interpretación sujeta a apreciación personal.	El temario incluido en el curso, ¿está indicado?

Como resultado de la reformulación anterior se obtuvo un cuestionario compuesto por 60 preguntas. Este cuestionario se sometió a pruebas para evaluar si posee las propiedades que aseguran su capacidad para medir. A continuación se definen dichas propiedades y se puntualiza sobre el Análisis de fiabilidad.

3.1. Propiedades de la Escala

Se entiende por *Escala* cualquier tipo de instrumento de medida, siendo uno de los más habituales el cuestionario. Un Instrumento de este tipo es un conjunto de preguntas que poseen una serie de propiedades métricas, ocupándose la medición de estudiar y cuantificar estas propiedades [29]. Toda escala debe ir acompañada de un método de corrección, de las instrucciones de aplicación y de la guía para la interpretación de las puntuaciones. Es decir, de propiedades que aseguren su capacidad para medir, las que se agrupan en: Fiabilidad, Validez y Factibilidad [29].

Fiabilidad es la capacidad de la escala para medir de forma consistente, precisa y sin error la característica que se desea medir. Cuando se aplica la escala a los mismos sujetos en dos situaciones diferentes ha de obtenerse la misma medición.

Validez es la capacidad del instrumento para medir lo que dice medir y no otros aspectos.

Factibilidad es la facilidad que tiene el instrumento para ser aplicado en diversas situaciones y grupos de sujetos.

3.2. Análisis de Fiabilidad

El Análisis de Fiabilidad permite determinar el grado en que los elementos del cuestionario se relacionan entre sí, obtener un índice global de la replicabilidad (consistencia interna de la escala en su conjunto) e identificar elementos problemáticos que deberían ser excluidos de la escala [29]. La fiabilidad de un instrumento de medida se valora a través de la Estabilidad, Consistencia interna y Confiabilidad Inter-Examinadores [30]

- Estabilidad: se pretende evaluar en qué grado el puntaje de un sujeto en un test está libre de errores de medición causado por fluctuaciones temporales azarosas. Son indicados para la medición el Método de Test-Retest y el Método de Formas Equivalentes [32].
- Consistencia interna: grado en que distintas partes o elementos del test miden la misma variable. Para este tipo de medición son indicados el método de Formas Equivalentes, la Partición en Mitades y el Método de Análisis de la Varianza de Ítems (Coeficiente Alfa de Cronbach o Kuder Richardson 20).
- Confiabilidad Inter-Examinadores: pretende evaluar el grado en que la medición de un rasgo a través de un instrumento es independiente de la subjetividad del evaluador. Es decir, que los sujetos examinados por un

test obtengan puntuaciones idénticas en sus ejecuciones independientemente de quién sea su examinador. El procedimiento indicado para este tipo de evaluación es el Método de Acuerdo entre Jueces.

Este trabajo presenta sólo el análisis de fiabilidad en lo que respecta a *consistencia interna*.

4. Evaluación de Consistencia Interna del Instrumento de Medición

Existen distintos modelos que permiten obtener diferentes coeficientes de fiabilidad, el GIS utilizó el Alfa de Cronbach que es un modelo de consistencia interna que se basa en la correlación inter-elementos promedio. Asume que la escala está compuesta por elementos homogéneos que miden la misma característica y que la consistencia interna de la escala puede evaluarse mediante la correlación existente entre todos sus elementos. Además asume que una escala es fiable cuando la variabilidad de las puntuaciones observadas es atribuible a las diferencias existentes entre los sujetos. Los valores del coeficiente oscilan entre 0 y 1, y se considera que existe una buena consistencia interna cuando el valor de alfa es superior a 0,7 (en algunos casos particulares es aceptable un valor inferior a 0,6 [31]). Este, como los otros estadísticos que se utilizan en el proceso de análisis de fiabilidad asume que los elementos de la escala se combinan aditivamente, es decir que la puntuación global de la escala se obtiene sumando las puntuaciones de sus elementos. En el caso de escalas multidimensionales (las preguntas están agrupadas por dimensiones de modo que unas preguntas miden una dimensión y otras una dimensión diferente) el cálculo del coeficiente de fiabilidad se realiza para cada una de las dimensiones. El cálculo del coeficiente Alfa de Cronbach se realizó con el software SPSS versión 12.

La ejecución de la evaluación de *fiabilidad* realizada al cuestionario de 60 preguntas, sobre el aspecto de *consistencia* reportó un alfa de Cronbach de 0,740 para un grupo de expertos. Al aplicarlo en el grupo de docentes disparador de la investigación, el alfa resultó inferior a 0,7.

Al analizar estos resultados se concluyó que los problemas encontrados podían deberse a que el cuestionario incluía preguntas sobre aspectos técnicos propios de INTERNET desconocidos por los usuarios especificados (por ejemplo, mapa de sitio, título de navegador, etc.). Como consecuencia de esto se decidió adoptar dos líneas de trabajo:

- Modelo de Calidad de Credibilidad e Interacción para usuarios de cursos de teleformación,

- Modelo de Calidad de Credibilidad e Interacción para desarrolladores de contenidos de cursos de teleformación.

Para la primera línea se optó por un cuestionario compuesto por 41 preguntas, destinado a personas con formación equivalente a un nivel de instrucción medio finalizado, no necesariamente especializadas en cuestiones informáticas ni de teleformación en forma específica, pero sí familiarizadas con la navegación en INTERNET.

Esta versión se sometió a tres evaluaciones, en grupos con las características mencionadas, mejorando sensiblemente el Alfa de Cronbach. Estas evaluaciones permitieron realizar un nuevo ajuste al cuestionario, finalmente éste quedó de 39 preguntas distribuidas en 5 (cinco) dimensiones (Anexo). La última versión se sometió a dos nuevas pruebas. La primera con un grupo de cuatro personas y la segunda con un grupo de once personas con las características mencionadas. La tabla 4 muestra los resultados obtenidos.

Tabla 4. Resultados obtenidos con el cuestionario de 39 heurísticas

Dimensión	Cantidad de Heurísticas	α de Cronbach	
		G1	G2
D1	22	0,967	0,770
D2	14	0,820	0,562
D3	1	*	*
D4	1	*	*
D5	1	*	*
Global	39	0,925	0,779

*como es una dimensión con una sola pregunta,

no se puede calcular el Alfa de Cronbach.

Conclusiones

Los resultados mostrados en la tabla 4, permiten asegurar que el cuestionario de MECACIN posee consistencia interna. Esto se debe a que el valor de Alfa de Cronbach global en ambas experiencias supera el umbral de 0,7, a pesar de que uno de los grupos presenta, en la dimensión D2, un Alfa menor a 0,7. Como las dimensiones D3, D4 y D5 poseen sólo una

heurística no fue posible aplicar un Alfa de Cronbach, sin embargo estas dimensiones presentan, en ambas experiencias, homogeneidad de respuesta para todos los evaluadores, lo cual permite asegurar que estas tres preguntas están bien formuladas

Probada la consistencia interna el GIS se abocará a determinar la Estabilidad y la Confiabilidad Inter-Examinadores para asegurar la fiabilidad de la escala. Posteriormente se enfocará el estudio de la validez y factibilidad. Los experimentos realizados permiten vislumbrar que este último aspecto se presentará como positivo.

Referencias

- [1] Pressman, Roger S. (2005), Ingeniería del Software. Un enfoque práctico. Sexta Edición. McGraw Hill Interamericana. México.
- [2] Covella, G. J. (2005). "Medición y Evaluación de Calidad en Uso de Aplicaciones Web". Presentación de Tesis de Maestría en la Facultad de Informática de la UNLP. La Plata.
- [3] Dart, S. (1999). "Containing the Web Crisis Using Configuration Management". En Proc. First ICSE Workshop on Web Engineering, ACM. Los Ángeles. <http://www.fistserv.macarthur.uws.edu.au/san/icse99-WebE/ICSE99-WebE-Proc/default.htm> (22/02/07)
- [4] Pflieger, S. L. (2002). Ingeniería de Software. Teoría y Práctica. Primera Edición. Prentice Hall y Pearson Educación. São Paulo SP.
- [5] ISO/IEC 9126 (1991). Information technology - Software product evaluation - Quality characteristics and guidelines for their use. Ginebra
- [6] ISO/IEC 9126-1 (2001), Software Engineering – Product quality. Part 1: Quality Model, Secretaría General de ISO, Ginebra.
- [7] ISO/IEC 9126-2 (2003). Software Engineering – Product quality. Part 2: External Metrics. Secretaría General de ISO. Ginebra.
- [8] ISO/IEC 9126-3 (2003). Software Engineering – Product quality. Part 3: Internal Metrics. Secretaría General de ISO. Ginebra.
- [9] ISO/IEC 9126-4 (2005). Software Engineering – Product Quality. Part 4: Quality In Use Metrics. Ginebra.
- [10] Olsina, L. A. (2000). Metodología Cuantitativa para la Evaluación y Comparación de la Calidad de Sitios Web. Presentación de Tesis Doctoral en la Facultad de Ciencias Exactas de la UNLP. La Plata.
- [11] Fernandez Nodarse, F.; N. Soubal y S. Lima Montenegro. (2002) "Experiencias en la Concepción de una metodología para el desarrollo y control de calidad de productos y servicios informáticos orientadas a la Educación a distancia y el Comercio electrónico en Internet" En Actas I Congreso Internacional de Tecnologías y Contenidos Multimediales en Ambientes Digitales. <http://espejos.unesco.or.uy/simplac2002/ad.html>.

- [12] Goodyear, P. (2000). "e-Learning, knowledge work and working knowledge". IST2000 Event, e-Learning Futures session, Nice.
- [13] Barroso, J.; J. L. Mendel; and J. Valdeverde. (1998). "Evaluación de los medios informáticos: una escala de evaluación para el software educativo". En Cebrián, M. et al. "Creación de materiales para la innovación con nuevas tecnologías": EDUTEC 97, 355-358. ICE Universidad. Málaga. <http://www.ice.uma.es/edutec97/edu97-c3/2-3-08.htm> (22/02/07).
- [14] Del Moral, E. (1998). "El desarrollo de la creatividad y las nuevas herramientas tecnológicas". En Comunicación Educativa y Nuevas Tecnologías". pp 51-66. Praxis. Barcelona. <http://www.eafit.edu.co/articulos/evalISE.htm> (22/02/07)
- [15] Galvis, A. (2000). "Evaluación de MECs por juicio de expertos", Capítulo 10 del libro: "Ingeniería de software educativo" 2da. reimpresión. Universidad de Los Andes. Colombia.
- [16] Gómez, M. T. (1997). "Un ejemplo de evaluación de software educativo multimedia". En Cebrián, M. et al. "Creación de materiales para la innovación con nuevas tecnologías": EDUTEC97. ICE Universidad. Málaga. http://www.ice.uma.es/edutec97/edu97_c3/2-3-03.htm (22/02/07)
- [17] González, M. (1999). "Evaluación de software educativo. Orientaciones para su uso" Proyecto Conexiones. Universidad de EAFIT.
- [18] Gros, B. (Coord.); A. Bernardo; M. Lizano; C. Martínez; M. Panadés y I. Ruiz. (1997) "Diseños y programas educativos, pautas pedagógicas para la elaboración de software". Editorial Ariel S.A. Barcelona
- [19] Marqués, P. (1998). "La evaluación de programas didácticos". En Comunicación y Pedagogía (149). pp 53-58. Barcelona. <http://www.xtec.es/~pmarques/tecnoedu.htm> (22/02/07)
- [20] MVU, Michigan Virtual University. (2002). "Standard Quality on-line courses". <http://ideos.mivu.org/standards> (22/02/07).
- [21] PEMGU. (1999). "Pedagogical evaluation methods & guidelines for multimedia applications" Partners: Epral (Portugal), Colegio Irabia (Spain) and Holbaek Technical College (Denmark), DEL (Denmark), the University of Cologne (Germany) and Olivetti (Italy). <http://www.irabia.org/pemgu/> (22/02/07).
- [22] Stephen, B. (1998). "Evaluating checklist. Evaluating training software". Lancaster University. http://www.keele.ac.uk/depts/cs/Stephen_Bostock/docs/evaluationch.ecklist2.html (22/02/07).
- [23] Lasserre, C.M. y V.E. Quincoces. (2005). Quality in E-Learning: a Heuristic Evaluation. En Proceeding Simposio ASSE 2005 (34JAIIO). Rosario.
- [24] Quincoces, V.E. y H. Liberatori. (2005). "Métrica para evaluación de Cursos de Tele training". Cuadernos 26, FHycS, Suplemento. VIII Jornadas Regionales de Investigación en Humanidades y Ciencias Sociales. ISSN 0327-1471. San Salvador de Jujuy.
- [25] Gálvez, M. P; H. Liberatori; V. Quincoces; A. García y C. Lasserre. (2006). "Evaluación de Sitios Educativos: Comparación de Heurísticas". En Investigaciones Docentes en Ingeniería. Facultad de Tecnología y Ciencias Aplicadas. Catamarca.
- [26] ISO/IEC 14598-1 (1999). Information Technology – Software Product Evaluation. Part 1: General Overview. Secretaría General de ISO. Ginebra.
- [27] Olsina, L. A.; G. Covella y G. Rossi. (2006). "Web Quality". En Web Engineering. E. Mendes & N. Mosley (Eds). pp 109-142.. Lecture Notes in Computer Science of Springer, ISBN 3-540-28-28-196-7.
- [28] Alexander, J. E. and M. A. Tate. (1999). WEB Wisdom. How to Evaluate and Create Information Quality on the WEB. Wolfram Memorial Library Widener University. Laurence Erlbaum Associates Publishens. Lodon. Mahwah. New Jersey. Apéndice A y B.
- [29] Pérez López, C. L. (2005). "Métodos estadísticos avanzados con SPSS". Editorial Paraninfo. Madrid.
- [30] Tornimbeni, S.; E. Pérez; F. Olaz y A. Fernández. (2004). Introducción a los Tests Psicológicos. 3º Ed. Editorial Brujas. Argentina. APA (American Psychological Association). (1999). Standards for psychological and educational tests. Washington, D.C.
- [31] Aron A. y E. N. Aron. (2001). Estadística para psicología. 1ª edición. Pearson Education. Buenos Aires.
- [32] APA (American Psychological Association). (1999). Standards for psychological and educational tests. Washington, D.C.

Anexo.

Métrica para Evaluación de Credibilidad e Interacción de Cursos de Teleformación (MECACIN)

1. La métrica sólo puede aplicarse si la PC utilizada cumple con las características de equipo y programas especificados por los desarrolladores del curso.
2. Los/as potenciales usuarios/as de esta métrica deben ser personas con formación equivalente a un nivel de instrucción medio finalizado, no necesariamente especializados en cuestiones informáticas ni de teleformación en forma específica, pero sí familiarizados con la navegación en INTERNET.

La Teleformación es un sistema de formación a distancia apoyado en las Tecnologías de la Información y las Comunicaciones (tecnologías, redes de telecomunicación, videoconferencias, TV digital, materiales multimedia), que combina distintos elementos pedagógicos: instrucción clásica (presencial o autoestudio), las prácticas, los contactos en tiempo real (presenciales, videoconferencias o chats) y los contactos diferidos (tutoriales, foros, correo electrónico).

La Métrica para Evaluación de Interacción y Credibilidad de Cursos de Teleformación proporciona una herramienta técnico conceptual que permite evaluar la calidad en uso de cursos de teleformación. Ésta establece una indicación de la calidad de Credibilidad e Interacción de esos cursos. Se entiende por calidad de credibilidad a la posibilidad de determinar con certeza quién acredita el curso, quién es su responsable y si se siguieron normas de edición (ortografía, gramática, dicción, legibilidad). Se entiende por calidad de interacción a la facilidad de navegación en el curso (velocidad de acceso a las páginas, indicación de cómo ir hacia arriba y hacia abajo en la página, volver a la página anterior, ir a la siguiente, vincularse con otras páginas, fondo de páginas) y de comprensión de su estructura (Títulos, objetivos, audiencia). La métrica NO evalúa la corrección conceptual de los contenidos de un curso.

Para obtener el valor de la métrica se debe responder una serie de preguntas que evalúan diferentes características del curso de teleformación. Al contestar estas preguntas debe considerarse lo siguiente:

- se consigna 1 si la respuesta es afirmativa (existe el objeto por el que se pregunta),
- se consigna 0 si la respuesta es negativa (no está presente el objeto por el que se pregunta) o
- se consigna X cuando la pregunta no se aplica (no está presente el objeto al que apunta – son las preguntas que comienzan con el condicional “Si”, o sea que el objeto puede faltar).

Una respuesta afirmativa indica la presencia de la característica evaluada, una respuesta negativa señala la ausencia o aparición parcial de la característica. Si la pregunta no es de aplicación, en un caso particular, no incide en el puntaje total de evaluación.

Cuanto mayor sea el número de respuestas afirmativas mayor es el grado de calidad de credibilidad e interacción del curso. Se propone como indicador de calidad el valor MECACIN (cálculo descrito en el ítem MODO DE USO). El valor umbral para MECACIN debe ser mayor o igual que 70%.

MODO DE USO:

Leer detenidamente cada pregunta, teniendo en cuenta las Notas al Pie, y asignarle 1, 0 o X según corresponda: 1(uno) si se presenta el objeto, 0 (cero) si no está y X si no lo encuentra y puede faltar.

Contestadas las preguntas el valor de MECACIN se calcula

$$\text{MECACIN} = \frac{\text{Cantidad de 1 (uno)}}{N} \times 100$$

Donde:

Cantidad de 1Uno): valor obtenido de sumar todas las preguntas respondidas afirmativamente (1)

N: cantidad de preguntas consideradas (no se tienen en cuenta las NO SE APLICA)

Si MECACIN es mayor o igual que 70% se puede afirmar que el curso de teleformación en operación ha sido desarrollado por una persona u organización confiable y que su usabilidad es aceptable.

Nº	Tipo	Heurística	Ptje
1	D1	¿Está/n indicada/s la/s organización/es o compañía/s que respalda/n el curso?	
2	D1	¿Existe un medio para contactar a la/s organización/es o compañía/s que respalda/n el curso? (Un número de teléfono, una dirección Web, un correo electrónico oficial, o una dirección postal)	
3	D1	¿Se indican nombre/s y apellido/s del/los autor/es del curso?	
4	D1	¿Se indican los antecedentes del/los autor/es del curso? Se considera suficiente, al menos, la presencia de los antecedentes de un autor.	
5	D1	¿El/los autor/es es/son idóneo/s en el/los tema/s tratado/s en el curso? Se considera suficiente que, al menos, uno de los autores sea idóneo.	
6	D1	El curso, ¿está protegido por derechos de autor?	
7	D1	¿Existe un medio para contactar al/los autor/es? (Un número de teléfono, una dirección Web, un correo electrónico oficial, o una dirección postal). Se considera suficiente que, al menos, se indique la vía de contacto de uno de los autores.	

Nº	Tipo	Heurística	Ptje
8	D1	Si existe una vía de contacto, ¿se indica cuándo va a recibir respuesta?	
9	D1	¿Existe una indicación de que el curso ha sido revisado por un editor?	
10	D1	¿Existe una indicación de que el curso ha sido sometido a revisión por pares?	
11	D2	El título del curso, ¿está indicado?	
12	D2	Los objetivos del curso, ¿están enunciados?	
13	D3	La audiencia a la cual va dirigido el curso, ¿está indicada?	
14	D4	¿Se indica el orden para aprender los contenidos del curso?	
15	D5	El lenguaje utilizado para expresar los contenidos del curso, ¿es sobrio, conciso y concreto (no insinuante y ambiguo)?	
16	D1	Si hay imágenes (gráficos, figuras, fotografías, etc.) ¿tienen su correspondiente epígrafe?	
17	D1	Si hay imágenes. ¿se las observa con nitidez?	
18	D1	Si hay tablas, ¿tienen su correspondiente epígrafe?	
19	D1	Si hay tablas, ¿son legibles?	
20	D1	Si el curso incluye datos estadísticos. ¿se indican las fechas de recolección de éstos? La omisión de una fecha es suficiente para responder cero.	
21	D1	Los contenidos del curso, ¿se encuentran libres de errores ortográficos? La presencia de un error es suficiente para responder cero.	
22	D1	Los contenidos del curso, ¿se encuentran libres de errores de sintaxis? La presencia de un error es suficiente para responder cero.	
23	D1	Los contenidos del curso, ¿se encuentran libres de errores tipográficos? La presencia de un error es suficiente para responder cero.	
24	D1	Si el curso contiene texto hablado, ¿se encuentra libre de errores de dicción (pronunciación)? La presencia de un error es suficiente para responder cero.	
25	D1	Si existe bibliografía, ¿se indican nombre/s y apellidos/s del/los autor/es, título de la obra, editorial, año de publicación y ciudad de edición? (en caso de ciudades homónimas se indica el país). Deben consignarse TODOS los datos enunciados en la heurística para cada una de las entradas bibliográficas.	
26	D1	¿Se indica la fecha de creación y/o modificación (o fecha de actualización) del curso? Se considera suficiente la especificación del año.	
27	D2	El título de la página principal, ¿es conciso?	
28	D2	El título de cada una de las páginas subordinadas, ¿se corresponde con el material desarrollado en ellas?	
29	D2	¿Incluyen todas las páginas subordinadas un vínculo a la página principal?	

Nº	Tipo	Heurística	Ptje
30	D2	Si se utilizan los vínculos para ir a la página Anterior o Siguiente (Atrás y Adelante), ¿se mantiene la ubicación en todas las páginas que los presentan?	
31	D2	Si se utiliza una barra de vínculos y/o botonera de vínculos, ¿se mantiene la ubicación en todas las páginas que la presentan?	
32	D2	Si se utiliza una barra de vínculos y/o botonera de vínculos, ¿es igual en todas las páginas que la presentan?	
33	D2	¿Todos los vínculos acceden a las páginas (o secciones dentro una misma página) a las que hacen referencia?	
34	D2	El curso, ¿contiene un índice o tabla de contenidos?	
35	D2	¿Incluyen todas las páginas un vínculo al índice o tabla de contenido? Se considera suficiente al menos una de las alternativas indicadas.	
36	D1	La disposición de colores de fondo y fuentes (letras) utilizados en el curso, ¿permite la legibilidad de los contenidos?	
37	D2	¿Se mantiene igualdad de fondo para cada tipo de actividad propuesta en el curso? (teoría, práctica, ejercitación, evaluación, etc.).	
38	D2	¿Se indican las características del equipo y programas necesarios para un correcto funcionamiento del curso?	
39	D2	Dentro del curso, ¿el desplegado de una página no supera los 20 (veinte) segundos? No se tendrá en cuenta este tiempo cuando se hace un enlace a temas complementarios o a sitios afines.	

Dirección de Contacto del Autor/es:

Cecilia María Lasserre
 Patricias Argentinas 478
 4to piso – Dpto a
 San Salvador de Jujuy
 Argentina
 lasserre@imagine.com.ar

GRUPO DE INGENIERIA DE SOFTWARE.
FI - Univ. Nac. de Jujuy Directora: MSc. C. M. Lasserre
 Creado en 1995 - Proyecto actual: "Definición y Validación de Métricas y Competencias en Cursos de Teleformación (Código Prog. Inc. 08-D0059) perteneciente al Programa SecTER-UNJu "Educación, TIC's e IWeb".
