

Aproximación a la Evaluación de la Calidad de Aplicaciones Web

Noelia Soledad Pinto¹, Nicolas Gabriel Tortosa¹, Liliana Raquel Cuenca Pletsch¹, César Javier Acuña¹, Cristina Greiner² y Marcelo Estayno³

Resumen

Lograr un alto nivel de calidad de un producto o servicio es el objetivo de la mayoría de las organizaciones que desarrollan software. En un esfuerzo por definir el concepto de calidad, algunos autores argumentan que un atributo de calidad puede contribuir a la obtención de mejoras en el funcionamiento y operación del software. En trabajos anteriores, (Cuenca Pletsch, Acuña, Pinto, Ibañez y Estayno, 2012) y (Cuenca Pletsch, Acuña, Pinto, Ibañez y Estayno, 2013), se desarrolló una aproximación a un framework, es decir un entorno de trabajo compuesto por un modelo de calidad y una herramienta de software para la evaluación de calidad en aplicaciones web. Ampliando esos aportes de investigación, en este artículo se describe detalladamente las características fundamentales de WQM, un modelo de calidad para evaluación de aplicaciones web y se presenta QUCO2, un software que gestiona los elementos del modelo de calidad propuesto y permite recolectar la información necesaria para evaluar productos web diversos y con características contrastantes.

Palabras Clave: Modelo de Calidad de Software, Métricas de Calidad, Calidad en Aplicaciones Web.

Fecha de recepción: julio 2013 | Fecha de aceptación: octubre 2013

¹ Universidad Tecnológica Nacional, Facultad Regional Resistencia, Departamento Ingeniería en Sistemas de Información.

² Universidad Nacional del Nordeste, Facultad de Ciencias Exactas y Naturales y Agrimensura, Departamento de Informática.

³ Universidad Nacional de Lomas de Zamora, Facultad de Ingeniería, Departamento de Informática.

Abstract

Achieving a high level of quality of a product or service is the goal of most software development organizations. In order to define the concept of quality, some authors argue that a quality attribute can contribute to achieve improvements in the software functioning and operation. In previous works, (Cuenca Pletsch , Acuña , Pinto , Ibanez and Estayno, 2012) and (Cuenca Pletsch , Acuña , Pinto , Ibanez and Estayno , 2013), a framework approach was developed, i.e. a working environment consisting of a quality model and a software tool for assessing quality in web applications. Expanding these research contributions, this paper describes the WQM key features, a quality model for web application assessment, and QUCO2 is introduces. The latter is a software that manages the elements of the proposed quality model and enables collecting the information needed to asses different web products and contrasting features.

Keywords: Software Quality Model, Quality Metrics, Web Applications Quality.

Introducción

Las cuestiones de calidad en los productos o servicios de cualquier índole se ha hecho algo cotidiano y cada vez más importante ya que se convirtió en un factor diferenciador cuando se está optando por adquirir un servicio o un bien. El concepto de calidad tiene muchas definiciones, pero una ampliamente aceptada es la establecida por la ISO 9000 (ISO, 2001) que define la calidad como el “grado en el que un conjunto de características inherentes cumple con los requisitos”. Estos requisitos son establecidos por los usuarios y por tanto se puede decir que la calidad es un tema subjetivo dependiente del nivel de satisfacción que el usuario sienta frente al producto utilizado. El software también está sujeto a la evaluación de su calidad, de tal manera que los usuarios puedan establecer el grado con el cual éste satisface sus necesidades.

Un tipo especial de productos software son las aplicaciones web, que exhiben factores en su proceso de desarrollo y en los resultados obtenidos que dificultan evaluar su calidad y por tanto obtener productos que satisfagan los requerimientos iniciales.

Actualmente ha aumentado la necesidad que los procesos de desarrollo web se enfoquen a mejorar aspectos relacionados a la calidad del producto. Este último concepto es uno de los más nuevos incorporado al estándar ISO/IEC 9126 2001 (ISO/IEC 9126-1, 2001) (ISO/IEC 9126-2, 2003) (ISO/IEC 9126-3, 2003) (ISO/IEC 9126-4, 2003), en el que se define al modelo de calidad como “la opinión que tiene un usuario al utilizar una aplicación de software”, que deriva, justamente, de los resultados obtenidos de la evaluación del uso del software.

El trabajo de investigación que aquí se presenta se desarrolla en el marco de un proyecto de investigación interinstitucional, “Modelos y Métricas para la evaluación de la calidad del software”, que llevan adelante investigadores de las Facultades Regional Resistencia de la UTN y de Ciencias Exactas y Naturales y Agrimensura de la UNNE, bajo la dirección de un investigador de la Universidad Nacional de Lomas de Zamora.

Particularmente en este artículo se describe un framework que incluye un modelo de calidad orientado a la evaluación de calidad en aplicaciones web y que se ha denominado modelo WQM. Además, se presenta la herramienta QUCO2, un software que permite gestionar todos los elementos del modelo WQM y obtener resultados que permitan validar preliminarmente dicho modelo.

En la sección 2, se incluye el Estado del Arte en cuanto a modelos de calidad existentes en la literatura y herramientas de software disponibles para medir la calidad de aplicaciones web. Luego, en la sección 3, se presenta el framework junto con la descripción del modelo de calidad propuesto y la implementación del software que lo completa. En la sección 4, se muestran los resultados luego de realizar una primera validación del framework mediante un caso de estudio.

Y finalmente, en la última sección se presentan conclusiones y trabajos a futuro.

Estado del Arte

Cuando se hace referencia a la calidad de software, debe ser analizada desde el punto de vista del Proceso o del Producto. En base a esta clasificación, se aprecia que la literatura ofrece diversos modelos de calidad entre los cuales se encuentra el más antiguo McCall (Mc Call, Richards y Walters, 1977), del cual derivan otros, como el de Boehm. Ambos, definen a la calidad como un conjunto de elementos organizados jerárquicamente. Esta jerarquía se presenta en la Fig.1:

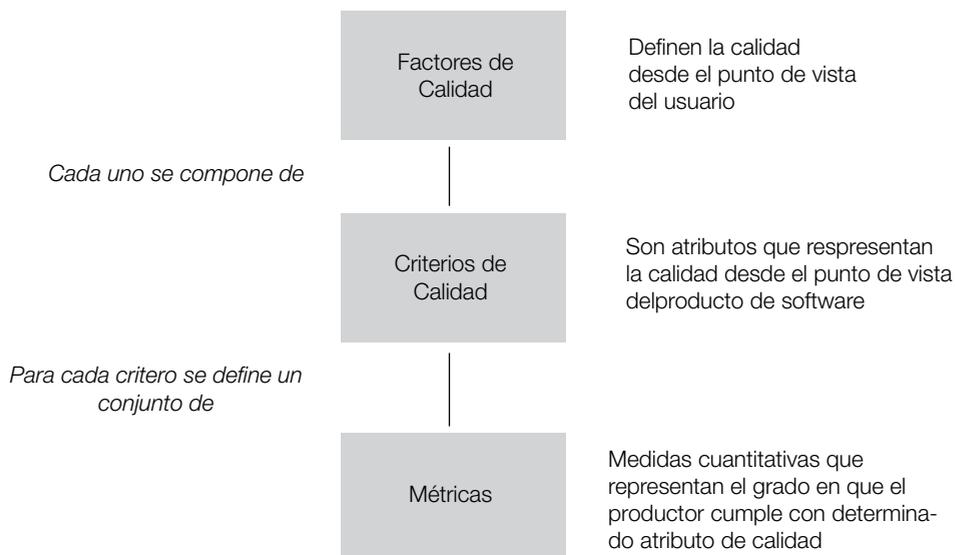


Fig. 1 Organización jerárquica

Sin embargo, el modelo que se presenta en este artículo, a diferencia de los enfoques existentes, se caracteriza por obtener, luego de la evaluación y mediante una fórmula definida, un valor que indica el nivel de calidad del sitio evaluado. Además, integra un framework que incluye también una aplicación usada para la realización de validaciones parciales de lo expuesto en WQM.

Así como hay disponibles diversos modelos de calidad, existen en el mercado herramientas que permiten evaluar la calidad de aplicaciones web. Entre estas, se pueden consultar las publicadas en el sitio de la W3C, llamadas Validadores (W3C, 2012) y que permiten verificar el código HTML y el CSS de una aplicación web, según un estándar definido por dicha Organización, para que pueda ejecutarse sobre cualquier plataforma. También está disponible LoadImpact (Load Impact AB, 2013) que simula el acceso de múltiples usuarios a una web, informando la

velocidad de carga de la web en situaciones de alta concurrencia de visitantes. Sin embargo, estas herramientas no evalúan calidad del software desde el punto de vista del usuario, sino que se enfocan en la calidad del producto desde el punto de vista del desarrollador.

Finalmente se ha analizado el funcionamiento de dos aplicaciones, que se promocionan en el mercado como herramientas para evaluar criterios asociados a la calidad en uso de software web. Una de ellas es Alexa (Alexa, 1996), que ofrece información sobre tráfico de visita diario, calculando la posición que ocupa, el sitio evaluado, en un ranking mundial. La otra aplicación analizada es la desarrollada por Silktide, Nibbler, (Silktide, 2012), que ofrece una puntuación desagregada en varios criterios, e indica cómo mejorar en cada aspecto. La evaluación puede variar en cada test realizado, debido a que elige cinco secciones al azar incluidas en la aplicación web, lo cual si bien permite corregir errores y mejorar aspectos en todo el sitio, no permite al evaluador seleccionar las secciones que le interesan.

Así, teniendo en cuenta la ausencia de un marco que englobe un modelo de calidad y software soporte, en este artículo se presenta un framework desarrollado en base a la propuesta de un nuevo modelo de calidad que se aplica específicamente a la evaluación de aplicaciones web y la implementación de una herramienta de software que permita validar parcialmente lo expuesto en este modelo.

WQF: Un framework para evaluar calidad de Aplicaciones Web

Para gestionar los elementos de calidad, el framework WQF incluye un modelo de calidad (WQM) y una herramienta de software (QUCO2), desarrollada en base a dicho modelo.

WQM: Modelo de Calidad para Aplicaciones Web

Debido a la importancia que la calidad de software en Internet ha adquirido en los últimos años, la Conferencia Internacional de la Ingeniería de Software del año 2002 (ICSE 2002) se centró en los aspectos de Calidad para los Sistemas en Internet (Dávila y Mejía, 2002). En esta conferencia se concluyó que las métricas más importantes son las siguientes: Fiabilidad, Usabilidad, Mantenibilidad, Seguridad, Disponibilidad y Escalabilidad (Covella, 2005).

En una primera instancia, y de acuerdo al alcance de este trabajo de investigación, el modelo de calidad propuesto queda conformado teniendo en cuenta las siguientes métricas (IEEE, 1990):

Métrica 1 - Usabilidad: todo sitio Web debe ser un producto atractivo, entendible y fácil de utilizar para los usuarios del mismo. Los criterios a evaluar son:

- **Facilidad de Aprendizaje:** Tiempo necesario que se requiere desde el desconocimiento de una aplicación hasta lograr su uso productivo.

- **Consistencia:** Verifica si todos los mecanismos mantienen el mismo circuito independientemente del momento.

- **Recuperabilidad:** Facilidad para corregir una acción una vez que el usuario reconoce un error en la operación.

- **Retención en el tiempo:** Sencillez, desde el punto de vista del usuario, para recordar la operación del sistema aun cuando hubiese transcurrido un tiempo considerable desde la última vez en que se utilizó.

- **Flexibilidad:** Evalúa las posibilidades de intercambio de información entre el usuario y el sistema.

La escala de valores establecida para esta métrica se describe en la Tabla 1:

Criterios de calidad	Escala Descripción	Valor
Facilidad de Aprendizaje	Muy complicado	-2
	Complicado	-1
Retención en el tiempo	Fácil	1
	Muy fácil	2
Consistencia	Consistencia en todo el sitio	3
	Consistencia en la mayor parte del sitio	2
	Inconsistencia en todo el sitio	1
Recuperabilidad	Si	1
	No	-1
Flexibilidad	El Sitio presenta diferentes opciones para una misma tarea	3
	El Sitio presenta una única alternativa para una tarea. Pero no presenta inconvenientes para llegar hasta la misma	2
	Es difícil intercambiar información dentro del Sitio	1
	No se puede intercambiar información en el Sitio	-1

Tabla 1. Escalas aplicadas a la métrica Usabilidad

Métrica 2: Confiabilidad: se relaciona con la capacidad del software de mantener su nivel de performance bajo las condiciones establecidas por un período de tiempo. Los criterios a evaluar son:

- **Frecuencia y severidad de las fallas:** Mide cada cuánto tiempo ocurren fallas en el sistema, si es que ocurren, y la habilidad del Software para mantener determinado nivel de performance.

- **Exactitud de las salidas:** Indica la aproximación de la salida deseada respecto de la salida lograda por el software.

- **Capacidad de Recuperación ante fallas:** Incluye procesos requeridos para detectar y recuperarse de situaciones inesperadas.

- **Confiabilidad:** mide la ocurrencia de accesos no autorizados a información privada.

En la Tabla 2 se presentan las escalas definidas para esta métrica.

Criterios de calidad	Escala	Valor
Frecuencia y severidad de las fallas	Errores graves continuamente	-2
	Errores leves continuamente	-1
	Errores leves sólo en ciertas funciones	1
	No se observan errores	2
Exactitud de las salidas	Si	1
Capacidad de Recuperación ante fallas	No	-1
Confiabilidad		

Tabla 2. Escalas aplicadas a la métrica Confiabilidad

Métrica 3 - Funcionalidad: permite comprobar la relación entre las funciones de las aplicaciones, los resultados esperados y los resultados reales. Los criterios de calidad a evaluar son:

- Adecuación: Determina si el conjunto de funciones son apropiadas para las tareas especificadas.

- Seguridad: Mide la habilidad para prevenir accesos no autorizados, ya sea accidentales o deliberados.

- Cumplimiento: Especifica si el software adhiere a estándares relacionados con la aplicación, y convenciones o regulaciones legales.

- Confiabilidad: No ocurrencia de accesos no autorizados a información privada.

En la Tabla 3 se presentan las escalas definidas para esta métrica.

Criterios de calidad	Escala	Valor
Adecuación	Si	1
Cumplimiento	No	-1
Confiabilidad		
Seguridad	Completamente seguro	3
	Muy seguro	2
	Seguro	1
	Se observan zonas inseguras	-1
	Completamente inseguro	-2

Tabla 3. Escalas aplicadas a la métrica Funcionalidad

Cada métrica propuesta asocia a cada característica un peso (por ejemplo: No Aplica, Aplica, Aplica fuertemente), que se evalúa según una escala de medición (por ejemplo Malo, Regular, Bueno, Muy Bueno). La fórmula general derivada para calcular el nivel de calidad general que se obtendrá es:

$$NO = \sum_{i=1}^n (VC_i \cdot PC_i) \quad (1)$$

donde NO es el Nivel Obtenido, VC_i es el Valor Calculado para la métrica *i* y PC_i es el Peso de la Característica *i* con PC_i entre 0 y 1.

Las sumatorias se realizan en función de todos los componentes seleccionados para la evaluación. Básicamente se trata de un promedio entre los valores obtenidos para cada componente influido por el peso de ese componente en el estudio general.

Así, con este modelo es posible obtener un nivel de calidad de determinada aplicación web a partir de la evaluación de las métricas que incluye el modelo WQM.

QUCO2: Software para medir la calidad de aplicaciones web

Contar con un modelo de calidad, que permita medir el nivel de calidad de una aplicación web, no es suficiente si no se dispone de una herramienta que posibilite gestionar los elementos del modelo y analizar los resultados obtenidos a partir de diversos casos evaluados. Por eso, WQF incluye el desarrollo de una aplicación que se utiliza para la evaluación on-line de sitios web por parte de usuarios finales, para la obtención de un nivel de calidad global.

La herramienta se denomina QUCO₂, y para su desarrollo se planteó el diseño de un contenedor para pequeños componentes reutilizables que permitan evaluar una característica en particular. Por cada característica se define una escala asociada a la misma, con la posibilidad de definir a su vez un peso relativo según el proyecto. Es decir que se obtiene información genérica y consolidada sobre la calidad del producto evaluado. El software provee diferentes roles de usuario: Desarrolladores, quienes registran información referida al proceso de desarrollo y el diseño de las autoevaluaciones para monitorear la evolución de la calidad del producto; Evaluadores de Calidad, quienes realizan las evaluaciones de calidad desde su óptica, y Clientes, quienes manejan la información generada por el sistema para tomar decisiones en base a la misma.

En cuanto al diseño del software, cada uno de los componentes se implementa en la forma de plugins, los cuales se desarrollaron completamente con tecnologías web.

En la figura 2, se muestra la arquitectura del front-end. El mismo se implementó mediante el proyecto Twitter Bootstrap v2 (Twitter, 2011) que a su vez hace uso del proyecto Less. La combinación de ambas tecnologías provee una interfaz de usuario agradable a la vista y que gracias al “responsive web design” permite su uso en dispositivos con características de pantalla limitadas (por ejemplo smartphones) sin necesidad de rediseñar la interfaz.

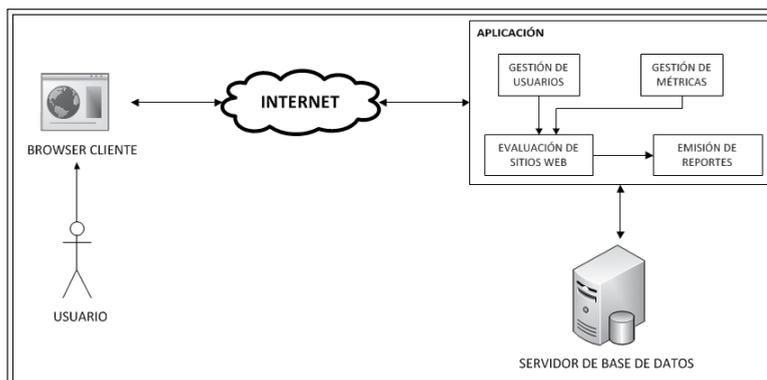


Fig. 2. Diagrama del diseño general del Front End

La herramienta consta, de acuerdo a la figura 3, de los siguientes módulos:

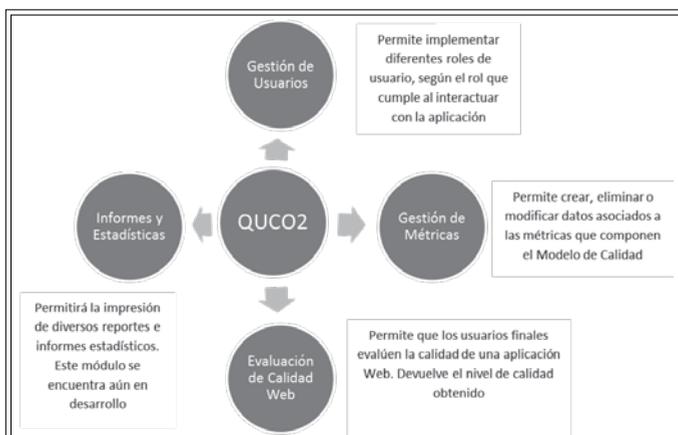


Fig. 3. Módulos funcionales de QUOCO2

Para el desarrollo del back-end del sistema de la figura 4, se utilizó el framework de desarrollo web Symfony (SensioLabs, 2005).

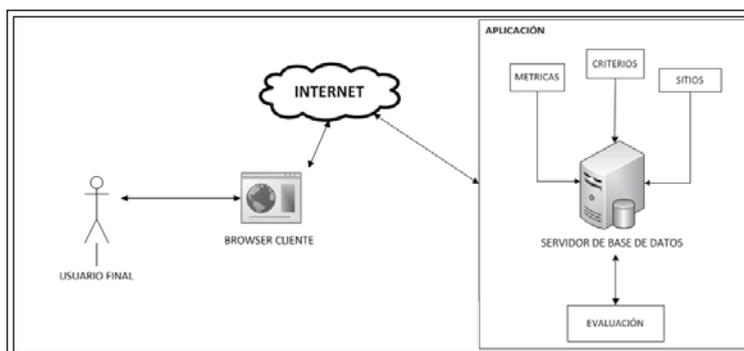


Fig. 4. Diagrama del diseño general del Back End

Validación

Para validar los resultados parciales de investigación, se tomó un caso de estudio que consistió en utilizar el framework para evaluar la calidad de tres aplicaciones web similares.

Para ello, se comparó QUCO2 con dos de las herramientas mencionadas en la sección 2:

- Nibbler: Mide la calidad de un sitio teniendo en cuenta varios aspectos, tales como: accesibilidad de personas con discapacidad, accesibilidad a partir de cualquier tipo de dispositivo, popularidad y aspectos vinculados con el desarrollo tecnológico del sitio. Asigna una nota a cada característica y obtiene un valor promedio como puntaje final del sitio. El valor del puntaje se establece entre 0 y 10.

- Alexa: Posiciona un sitio web, mediante un cálculo geométrico basado en los últimos 3 meses que promedia visitas de páginas del sitio y usuarios que han llegado hasta él. Un menor valor significa, por ejemplo, mayor permanencia del usuario en el sitio. Cuanto menor es el puntaje asignado, mayor es la popularidad del sitio, lo cual podría asociarse a una mayor calidad vinculada con los criterios de usabilidad antes descriptos.

Se definió un contexto de uso teniendo en cuenta el ámbito del comercio electrónico y gestiones administrativas. Para ello se eligieron tres sitios web con objetivos diferentes: un Portal gubernamental del NEA, un sitio de compra y venta on-line y un sitio de ventas de entradas on-line.

La población seleccionada la constituyó un grupo de 40 personas, caracterizados por realizar transacciones on-line periódicamente. Para la evaluación se tuvo en cuenta: información del sitio web, tiempo requerido para la evaluación y evaluación final obtenida con el uso de la herramienta.

Mediante el análisis de los resultados obtenidos con QUCO2, se demostró el correcto funcionamiento de la herramienta sobre un entorno real. Además se verificó que la curva de aprendizaje no presentó dificultades para los evaluadores.

Con la evaluación final, se procedió a calcular un promedio ponderado de los valores obtenidos para cada sitio propuesto. En base a una ponderación de los pesos de las métricas, se observó que el valor de máxima satisfacción ocupa el rango entre 20 y 22, al valor de calidad media le corresponde el rango entre 14 y 20, el valor de calidad regular se ubica entre 8 y 14 y el mínimo nivel registra un valor por debajo de 7. Teniendo en cuenta esto se presentan a continuación el resultado de la comparación:

- En el caso del sitio gubernamental, el valor promedio obtenido con QUCO2 fue 12, valorándose, entonces, como un sitio de calidad regular. Para Nibbler el resultado fue de 3.6, que podría calificarse como bajo, en tanto que Alexa ubica al sitio en la posición mundial número 692842. La siguiente figura incluye el gráfico

lineal del resultado de las evaluaciones, donde el eje horizontal indica el evaluador y el vertical la evaluación otorgada por cada uno.

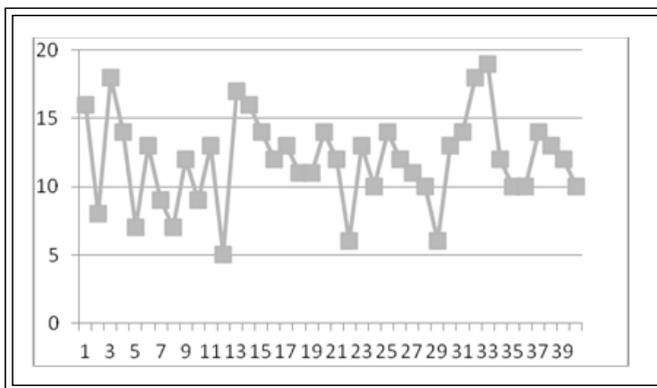


Fig. 5. Gráfico lineal de evaluación correspondiente al Portal gubernamental

- En el caso del sitio de compra y venta on-line se observa, según la figura 6, que la evaluación varía en el rango entre 15 y 20. El valor promedio obtenido de calidad fue 19, evaluándose así como un sitio de calidad ideal según el modelo de calidad propuesto en este artículo. Usando Nibbler el valor obtenido fue de 5.1, considerándose un puntaje medio, en tanto que Alexa lo posiciona en el puesto global número 2257.

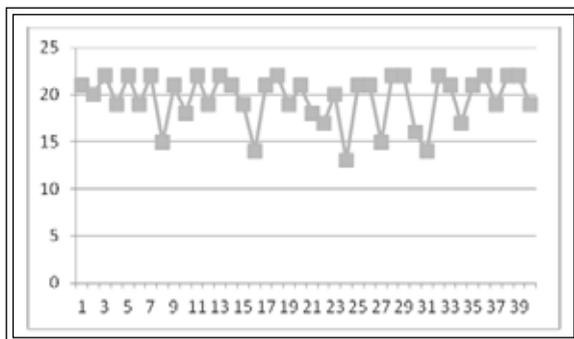


Fig. 6. Gráfico lineal de evaluación del Sitio de compra y venta on-line

- Los resultados obtenidos para el sitio de venta de entradas, según la figura 7, muestran que la evaluación se promedia en el rango entre 13 y 18. Efectivamente el valor promedio obtenido de calidad fue 16, un sitio de calidad muy buena. En el caso de Nibbler el valor final obtenido fue de 3.6, es decir un puntaje bajo igual al obtenido por el Portal gubernamental. Alexa lo ubica globalmente en la posición 225044, mejor que el portal gubernamental.

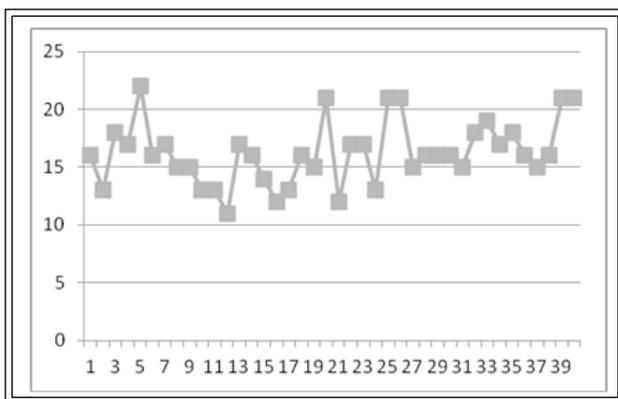


Fig. 7. Gráfico lineal de evaluación para Sitio de venta de entradas

Ante esta aparente disparidad de resultados se procedió a evaluar dos nuevos sitios, ubicados entre los mejor evaluados por Nibbler, para analizar los mismos con mayor detalle. Se procedió, entonces, a evaluar un sitio de servicios legales y otro vinculado con el turismo.

- En el primer caso, según la figura 8, el valor promedio obtenido de calidad fue 17, valorándose, entonces, como un sitio de calidad media, en tanto que para Nibbler el resultado fue de 9.4, considerado alto. Alexa ubica al sitio, en la posición mundial número 437715.

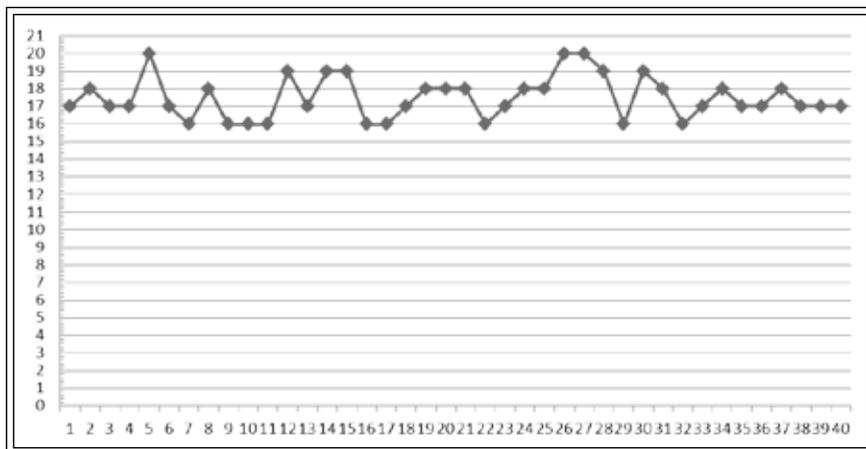


Fig. 8. Gráfico lineal de evaluación para sitio de servicios legales

- En el segundo caso, de acuerdo a la figura 11, el valor promedio obtenido de calidad fue 19, valorándose, entonces, como un sitio de calidad alta, coincidentemente con la valuación de Nibbler, 9.6. Alexa ubica al sitio, en la posición mundial número 23439.

Referencias

- Alexa. (1996). *Software Alexa*. Recuperado de <http://www.alexacom>
- Covella, G.J. (2005). *Medición y evaluación de calidad en uso de aplicaciones web*. Tesis de Maestría Facultad de Informática, Universidad Nacional de La Plata.
- Cuenca Pletsch, L., Acuña, C., Pinto, N., Ibañez, M., Estayno, M., Almirón, E. (2012). *Evaluación de Calidad de Aplicaciones Web*. Artículo publicado en CADI 2012. ISBN 978-987-1312-46-7
- Dávila, N., Mejía, A. (2002). *Evaluación de la Calidad de Software en Sistemas de Información en Internet*. Tesis de maestría no publicada, Centro de Investigación y de Estudios Avanzados del IPN, Departamento de Ingeniería Electrónica, Sección Computación.
- IEEE. (1990). *Standard 610. Compilation of IEEE Standard Computer Glossaries*.
- ISO. (2001). *Sistemas de gestión de la calidad-Conceptos y vocabulario. Norma Internacional ISO 9000*.
- ISO. (2001). *ISO/IEC 9126-1 – Software engineering–Product quality – Part 1: Quality Model*.
- ISO. (2003). *ISO/IEC 9126-2 – Software engineering– Product quality – Part 2: External Metrics*.
- ISO. (2003). *ISO/IEC 9126-3 – Software engineering– Product quality – Part 3: Internal Metrics*.
- ISO. (2003). *ISO/IEC 9126-4 – Software engineering– Product quality – Part 4: Quality in Use Metrics*.
- Load Impact AB. (2013). *Software LoadImpact*. Recuperado de <http://www.loadimpact.com>
- Mc Call, J.A., Richards, P.K., Walters, G.F.: Factors in Software Quality (1977). Vol. I, II, III: *Final Technical Report*, RAD-TR-77369. Rome Air Development Center, Air Force System Command, Griffith Air Force Base, Nueva York.
- Silktide. (2012). *Software Nibbler*. Recuperado de <http://www.nibbler.com>
- Pinto, N., Tortosa, N., Cuenca Pletsch, L., Acuña, C., Estayno, M. (2013). *QUCO2: Una herramienta para medir la calidad de aplicaciones Web*. Artículo publicado en WICC 2013. ISBN 978-987-28179-6-1.

- Sensio Labs. (2005). *Framework de Desarrollo Web Symfony*. Recuperado de <http://www.symfony-project.org>
- Twitter. (2011). *Proyecto Twitter Bootstrap V2*. Recuperado de <http://twitter.github.com/bootstrap>
- W3C. (2012). *Software Validador de CSS*. Recuperado de <http://jigsaw.w3.org/css-validator>
- W3C. (2012). *Software Validador de HTML*. Recuperado de <http://validator.w3.org>

