

WQF EN ACCIÓN: UNA EXPERIENCIA DE EVALUACIÓN DE LA CALIDAD EN APLICACIONES WEB EN LA INDUSTRIA DEL SOFTWARE DEL NEA

Noelia Pinto, GICS UTN FRRe, ns.pinto@gmail.com

Nicolás Tortosa, GICS UTN FRRe, nicotortosa@gmail.com

Blas Cabas Geat, GICS UTN FRRe, blasc147@gmail.com

Maximiliano Ulibarrie, GICS UTN FRRe, agustin.ulibarrie@gmail.com

Liliana Cuenca Pletsch, GICS UTN FRRe, lilianacp@gmail.com

Resumen— Actualmente existen en el mercado diversas herramientas que permiten evaluar la calidad de aplicaciones Web teniendo en cuenta sólo algunas cuestiones particulares como por ejemplo: accesibilidad, experiencia de usuario, tráfico de visitas diarias, entre otras. Sin embargo, se evidencia ausencia de herramientas que puedan ser usadas para evaluar la calidad de forma integrada en este tipo de aplicaciones.

En trabajos anteriores, se han presentado las particularidades de diseño y desarrollo de WQF (Web Quality Framework), un framework compuesto por un modelo de calidad (WQM) y una suite de herramientas (QUCO2 y QUVI) orientado a la evaluación de calidad de aplicaciones Web.

En este artículo se presenta los resultados de la experiencia de validación del framework en un contexto colaborativo junto al Polo IT Chaco, una organización que nuclea a empresas PYMES de la Industria del Software pertenecientes a la región NEA.

La configuración del modelo de calidad y la gestión de sus componentes a través de aplicaciones del framework, constituyen uno de los aportes significativos de este estudio junto a la obtención de metadatos que servirán de línea base para futuras experiencias similares.

Palabras clave— *Calidad del Producto de Software, Evaluación de Calidad de Aplicaciones Web, Ingeniería de Software.*

1. Introducción

El concepto de calidad tiene muchas definiciones, pero una ampliamente aceptada es la establecida por la ISO 9000 [1] 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 [2]. Por ello el desarrollo de productos de software con calidad es uno de los factores críticos para el éxito de las organizaciones.

Las aplicaciones Web son un tipo especial de productos software, pues tienen características que las diferencian de los sistemas tradicionales tales como: el tamaño y complejidad de las

aplicaciones, el carácter multidisciplinar del equipo de desarrollo, la tasa apresurada de entrega del proyecto, entre otras. Las características antes mencionadas traen consigo que los procesos, modelos y métricas existentes para evaluar la calidad tengan que ser adaptados para considerar los cambios impuestos por las nuevas tecnologías [3].

En trabajos anteriores se ha presentado WQF[4], un framework que incluye un modelo de calidad orientado a la evaluación de calidad de aplicaciones web y que se ha denominado modelo WQM. El framework incluye además dos herramientas:

- QUCO2, un software que permite gestionar todos los elementos del modelo WQM e implementarlo a través de una encuesta automatizada que es dirigida a usuarios finales de aplicaciones Web [4].
- QuVi, una aplicación cuyo objetivo principal es la obtención de informes estadísticos en base a los datos recolectados a través de QUCO2, que colaboren en la toma de decisiones a nivel gerencial de las PYMES [5].

Por tanto, en este trabajo se presentan los resultados del proceso de validación del framework WQF, tomando como referencia un caso particular de unas de las empresas participantes de la experiencia.

El artículo se estructura de la siguiente manera: en la sección 2 se describen los detalles técnicos de implementación de ambas herramientas del framework junto a las particularidades surgidas en el proceso de integración de las mismas y, seguidamente, en la sección 3 se presenta la experiencia del caso de estudio junto a los resultados obtenidos en cada etapa. Finalmente en la última sección se describen las conclusiones obtenidas y se expresan las futuras líneas de trabajo vinculadas.

2. Materiales y Métodos

2.1 Motivaciones

En una etapa previa al desarrollo e integración de WQF se han analizado el funcionamiento de herramientas que actualmente se encuentran disponibles en el mercado y evalúan diversos criterios asociados a la calidad en uso de software web.

Por ejemplo la aplicación desarrollada por Amazon, Alexa [6] que ofrece información sobre el tráfico de visitas diarias, calculando la posición que ocupa el sitio evaluado en un ranking mundial. Sin embargo, que una aplicación sea valorada en los mejores puestos no garantiza calidad del producto.

Otra aplicación estudiada fue Nibbler[7], patrocinada por SilkTide, 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 resulta poco útil porque si bien permite corregir errores y mejorar aspectos en todo el sitio, no permite al evaluador seleccionar las secciones que le interesan. Además, otro punto controversial está relacionado a que varios de los criterios que ofrece Nibbler para evaluar no se relacionan a cuestiones de calidad.

Así, y teniendo en cuenta la ausencia de herramientas que trabajen en base a modelos de calidad que integren varios aspectos a evaluar, se ha desarrollado WQF, cuya validación y resultados de implementación en ambientes reales se presentan en este artículo.

2.2 Características técnicas de WQF

Haciendo referencia brevemente al modelo de calidad que forma parte de WQF y denominado WQM (Web Quality Model) [8] se exponen a continuación las métricas que lo conforman:

- *Usabilidad*: Indica el grado de eficacia, eficiencia y satisfacción con la que usuarios específicos pueden lograr objetivos específicos, en contextos de uso específicos al utilizar un producto. Los criterios a evaluar son: Facilidad de aprendizaje, Consistencia, Recuperabilidad, Retención en el tiempo y Flexibilidad.
- *Confiabilidad*: Esta métrica se relaciona con la capacidad del software de mantener su nivel de performance bajo las condiciones establecidas por un período de tiempo. En este caso se evalúa: Frecuencia de fallos, Exactitud en las salidas, Capacidad de recuperación ante fallas.
- *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, Seguridad, Cumplimiento.

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, que se presenta en (1), ha sido pensada para calcular el nivel de calidad general que se obtendrá en cada evaluación, utilizando WQF:

$$NO = \frac{\sum_{i=1}^n (VC \cdot PC)}{\sum_{i=1}^n (PC)} \quad (1)$$

dónde NO es el Nivel Obtenido, VC es el Valor Calculado para la métrica i y PC es el Peso de la Característica i . 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.

Durante su desarrollo el framework pasó por diversas fases, para lograr obtener la mejor aproximación en las funcionalidades que lo caracterizan. Así luego de la definición del modelo WQM, base fundamental de su implementación, se continuó con el desarrollo de las herramientas de apoyo.

De esta forma, en una primera instancia se trabajó sobre QUCO2, una aplicación que permite a los usuarios evaluar un software web de manera automatizada, teniendo en cuenta los componentes propuestos en el modelo de calidad WQM. La tecnología utilizada para el desarrollo fue Django [9], un framework de código abierto para el desarrollo rápido de aplicaciones web. Posteriormente se desarrolló la aplicación QUVI, que toma los datos recolectados a través de QUCO2 y permite elaborar informes estadísticos.

Al finalizar el desarrollo de ambas herramientas, y luego de haberse validado cada una por separado, se dio inicio a la etapa de integración funcional bajo tecnología REST (Representational State Transfer), un estilo de arquitectura que permite desarrollar servicios, favoreciendo la escalabilidad en el desarrollo y la información compartida entre ambas aplicaciones de una manera segura.

2.3 Implementación Ágil de WQF

Para la fase de integración de WQF, y con el objetivo de utilizar metodología ágil, el equipo de desarrollo decidió que el proyecto se implemente bajo Scrum, definiéndose todos los elementos necesarios a tal fin.

Se definieron, entonces, seis sprints que permitieron cubrir la totalidad de requerimientos incluidos en el product backlog. Luego se planificó cada sprint al inicio de la semana, estableciendo las prioridades de los requerimientos y el tiempo final de duración de la unidad de trabajo. Durante el proceso de desarrollo surgió la necesidad de refinar el product backlog, y clarificar algunos requerimientos que presentaron ambigüedad en su definición, por ejemplo la información a la que podría acceder cada perfil de usuario manejado en WQF.

Para la validación parcial de los componentes implementados en WQF se diseñaron y ejecutaron lotes de casos de prueba incrementales hasta llegar al testing de la herramienta integrada junto a QUCO2 y QUVI.

2.4 Primeros pasos

Una vez que se tuvo disponible una versión funcional de WQF, se inició la transferencia entre las empresas que integran el Polo IT Chaco. En primer lugar, y en reunión con los directivos de cada empresa, se presentó una propuesta de plan de acción con el objetivo de poner en conocimiento el conjunto de tareas y cronograma previsto para llevar adelante el proceso. La propuesta incluyó las siguientes secciones:

- *Propósito*: una breve descripción del motivo por el cual se propone realizar este proceso.
- *Alcance*: define el límite de la implementación.
- *Cronograma*: establecimiento de tiempos para implementar las diferentes etapas.
- *Equipo*: lista de los participantes y su rol correspondiente en el proyecto.

Luego de haber detallado los procedimientos, cada empresa seleccionó la aplicación web a evaluar y la población de usuarios que, luego, participaron de la experiencia. Este proceso se basó en las prioridades de la empresa respecto a qué aplicaciones deberían incluirse en ciclos de mejora y, en base a esto, se realizó un estudio sobre qué usuarios podrían participar de la experiencia de evaluación.

Asociado a esto es importante destacar que el equipo de desarrolladores fue el encargado de realizar las capacitaciones correspondientes para guiar en la utilización del framework y facilitó la documentación necesaria para que sea distribuida entre los usuarios finales que también utilizarían los componentes de WQF.

Además se configuró el servicio web de forma local en un servidor provisto por la Facultad Regional Resistencia, que permitió disponer de un dominio a través del cual se pueda acceder a QUCO2 y a QUVI. Finalmente se autorizó el inicio del lote de evaluación, estableciendo fechas límites para la obtención y análisis incremental de los resultados.

3. Resultados y Discusiones

Del caso de estudio participaron varias empresas y al momento siguen sumándose nuevos proyectos para evaluación. Por ello, y a los fines de este artículo, aquí se presentan los resultados obtenidos de la experiencia de una de las empresas participantes.

Así, en el caso que aquí se describe, la población total de la experiencia se compuso de 50 usuarios asiduos de la aplicación web propuesta por la empresa. Estos usuarios fueron consultados e informados acerca del objetivo del proceso y de la importancia de su rol en el proceso de evaluación.

Al inicio, se les entregó a cada uno la URL de QUCO2 junto a la cuenta de acceso para trabajar.

Así cada usuario realizó el proceso de evaluación de la aplicación asignada y a medida que iban finalizando, QUCO2 les informaba sobre una calificación final. Este valor numérico se encuentra comprendido en alguno de los rangos expresados en la tabla 1, donde se incluye la relación Valor-Nivel de calidad, de acuerdo al modelo WQM.

Tabla 1. Relación Nivel de Calidad-Valor

Nivel de calidad	Rango de Valor Final
Baja	0-7
Regular	8-13
Media	14-19
Máxima	20-22

Fuente: De elaboración propia

Una vez finalizado el lote de evaluación, en el período establecido, QUVI se sirvió de la información recolectada en las encuestas y se pudieron analizar estadísticamente los resultados. Así, de acuerdo a lo que se ve en la Figura 1, el 60% evaluó como Regular al producto, el 28% con calidad Baja y el 12% con una calidad Media.

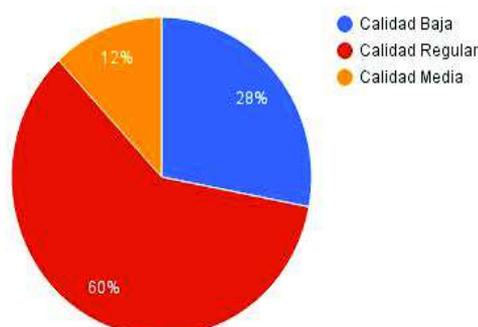


Figura 1. Porcentajes de usuario de acuerdo a la evaluación de calidad de la aplicación

Fuente: De elaboración propia

En promedio, de acuerdo a las evaluaciones obtenidas, la aplicación evaluada obtuvo un valor final de 11.14, aproximadamente, por lo que el framework la catalogó como un producto de Calidad Regular tal como se muestra en la Figura 2.



Figura 2. Nivel de calidad promedio de aplicación evaluada

Fuente: gics-quvi.frre.utn.edu.ar

Todas las evaluaciones almacenadas en QUCO, generan datos que luego son consumidos, a medida que son necesarios, por QUVI, la otra aplicación integrada en WQF.

A partir de esto, y en primer lugar, se ofrece a la empresa una vista resumida de las evaluaciones que se realizaron ofreciendo datos tales como: Cantidad de evaluadores que participaron del lote de evaluación, período en el que se realizó, etc. Un ejemplo de la vista inicial que presenta QUVI se muestra en la Figura 3.

QUVI Evaluaciones realizadas

	#	Titulo	Cantidad de evaluadores	Periodo de evaluación		Acciones
				Habilitación	Cierre	
+	1	EVALUACION 5 i	7	16/05/2016 21:59	-	Ver informe general
+	2	EVALUACION 4 i	4	16/05/2016 21:59	-	Ver informe general

Figura 3. Evaluaciones realizadas en QUVI

Fuente: gics-quvi.frre.utn.edu.ar

La principal potencia de QUVI, sin embargo, radica en los informes estadísticos que permite obtener. Uno de ellos, por ejemplo, es el gráfico de barras que informa, desde una perspectiva general, los valores promedios que se obtuvieron para cada una de las métricas, es decir, el cociente entre el puntaje total que obtuvo la métrica y la cantidad de usuarios que evaluaron la aplicación. En la Figura 4 se puede observar, teniendo en cuenta el lote de evaluación estudiado, que la Usabilidad obtuvo un mayor puntaje en comparación con la Funcionalidad y la Confiabilidad.



Figura 4. Gráfico de barra de valores promedio relacionados a las métricas en la evaluación
Fuente: gics-quvi.frre.utn.edu.ar

Otro informe estadístico se representa en la Figura 5, esta representa un ejemplo de gráfico de Caja Bigote obtenido a partir de las evaluaciones realizadas por los usuarios. El mismo nos permite identificar varias características importantes al mismo tiempo, tales como: valores atípicos, dispersión y simetría de los datos, valores centrales, etc.

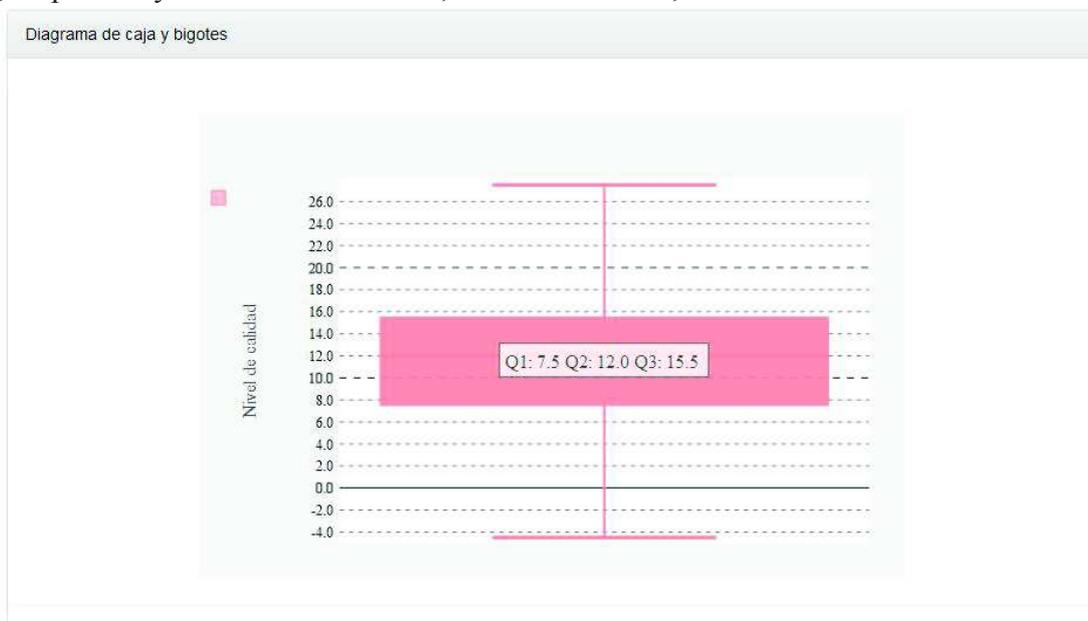


Figura 5. Diagrama de Caja y Bigote obtenido en la evaluación
Fuente: www.gics-quvi.frre.utn.edu.ar

Así para el caso de estudio se puede observar que entre los valores, 7.5 y 15.5 que forman la caja (el 1º y 3º cuartil de la distribución respectivamente), se concentran el 50% central de los

datos. La línea central en la caja representa el valor de la mediana. De este modo, si los resultados de las evaluaciones son simétricos, dicha línea se encontrará en el centro de la caja. Los extremos de los "bigotes" que salen de la caja son los valores que delimitan el 95% central de los datos, aunque en ocasiones coinciden con los valores extremos de la distribución, cualquier valor más allá de estos extremos representará un valor atípico que podría ser fácilmente detectado. Además podemos mencionar que entre el extremo inferior y el cuartil uno (segmento representado por el bigote inferior), muestra la distribución del primer 25% de los datos, ayudando a identificar de mejor manera la dispersión de los mismos. Esto significa que mientras más largo sea este tramo, mayormente dispersos se encontrarán los mismos. De igual forma ocurre con los últimos 25% de los datos, representados entre el cuartil tres y el extremo superior (segmento representado por el bigote que se encuentra más arriba).

QUVI, además, ofrece la posibilidad de obtener informes más detallados profundizando el análisis del comportamiento de los usuarios al evaluar cada métrica. Por ejemplo en la Figura 6 se muestra la caracterización de la métrica Usabilidad, cuyo valor fue el mejor evaluado en el conjunto de usuarios, de acuerdo a sus criterios, para la cual tuvo más impacto la Consistencia en el valor final obtenido.

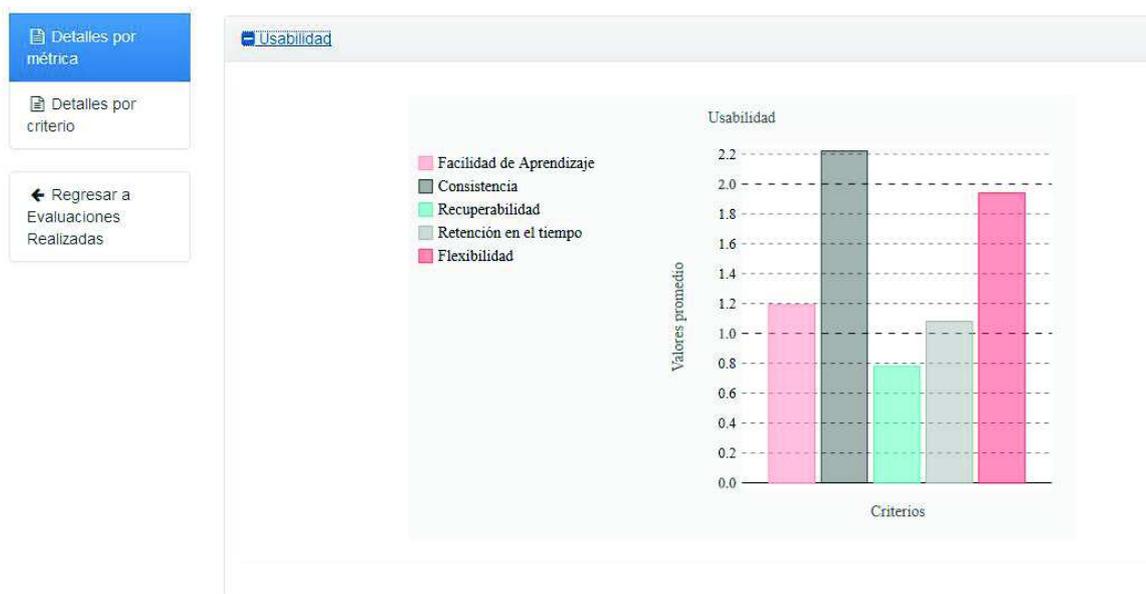


Figura 6. Detalle de evaluación de criterios para la métrica Usabilidad

Fuente: www.gics-quvi.frre.utn.edu.ar

En forma similar, la Figura 7 muestra cómo fue caracterizada la métrica de Confiabilidad en el lote de evaluación respecto a sus criterios. Por ejemplo se puede observar que el criterio mejor evaluado fue el de Capacidad de recuperación ante fallas y el peor Frecuencia y severidad de las fallas

EVALUACIÓN DE CALIDAD DE APLICACIONES WEB: EXPERIENCIA EN UNA PYME DEL POLO IT CHACO



Figura 7. Detalle de evaluación de criterios para la métrica Confiabilidad
Fuente: www.gics-quvi.frre.utn.edu.ar

En el caso de la métrica de Funcionalidad, se puede observar, de acuerdo a la Figura 8, que el criterio asociado al Cumplimiento fue el que obtuvo el valor más alto. En esta misma métrica, el criterio asociado a seguridad, fue el peor evaluado.

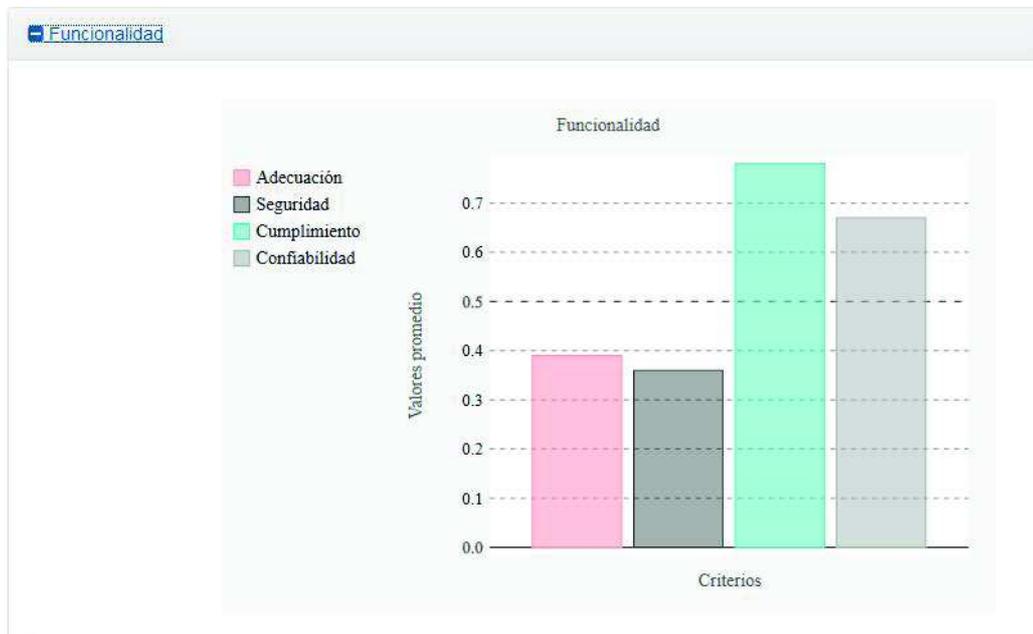


Figura 8. Detalle de evaluación de criterios para la métrica Funcionalidad
Fuente: www.gics-quvi.frre.utn.edu.ar

Además, se ofrece mayor detalle de información por criterio teniendo en cuenta el valor de cada métrica. Así por ejemplo si se quisiera analizar el caso de la Usabilidad, el criterio de Consistencia se caracteriza como se muestra en la Figura 9.

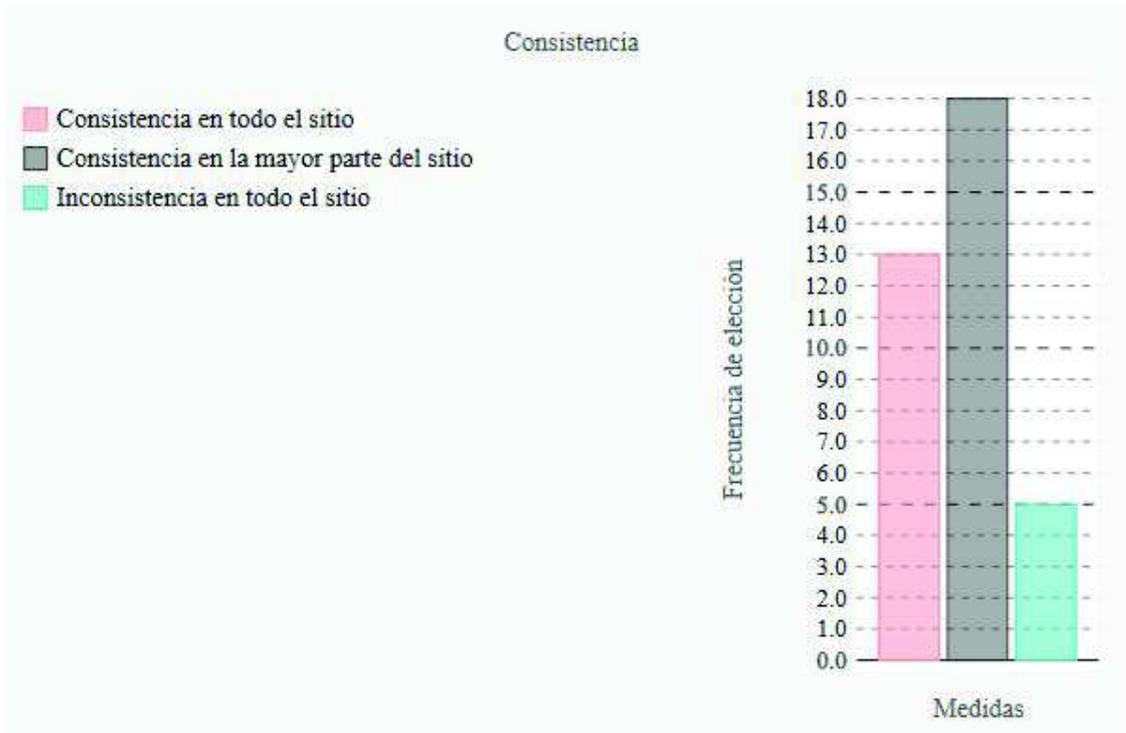


Figura 9. Detalle de evaluación del criterio Consistencia en Usabilidad

Fuente: www.gics-quvi.frre.utn.edu.ar

Una vez finalizada la etapa de recolección de datos y análisis de los informes estadísticos obtenidos, se realizó una reunión con la empresa participante para trabajar colaborativamente en el proceso de mejora continua respecto al framework propuesto. De esta reunión participaron los directivos de la PYME y los miembros del Grupo de Investigación abocados al estudio. Así a continuación se exponen las conclusiones y recomendaciones obtenidas al final de la experiencia.

4. Conclusiones y Recomendaciones

Como ya se mencionó, WQF es un framework basado en la definición de un modelo, de forma tal que todos los artefactos involucrados en la gestión de calidad de productos de software web puedan ser evaluados de forma automática. Así WQF fue validada como una herramienta integral diseñada para analizar, evaluar y mejorar la calidad de aplicaciones web. En este artículo se han presentado los resultados obtenidos a partir del proceso de validación con una de las PYMES del Polo IT Chaco, la cual participó de la experiencia de utilización del framework para evaluar la calidad de determinada aplicación, desarrollada y comercializada por dicha empresa.

Entre los resultados finales se definieron, en conjunto con la empresa, que algunos de los beneficios que proporciona la utilización del framework son: la rapidez en obtener el valor

final de evaluación de calidad de una aplicación web, la reducción de errores en la captura de información de la evaluación y mayor exactitud en el cálculo de valor de calidad asociado gracias a la automatización de los procesos. Los informes obtenidos fueron validados por la empresa, quienes consideraron de mucha utilidad el detalle que ofrece QUVI para analizar con mayor precisión cada métrica evaluada y expresaron su conformidad.

Como trabajos futuros surgió la necesidad de modificar el framework de forma tal que permita configurar el modelo WQM de acuerdo a las características que la empresa desee evaluar por aplicación, pues se pudo concluir, a partir del caso de estudio, que no siempre el peso de las métricas se mantiene uniforme sino que este varía de acuerdo al ambiente de producción.

Además, actualmente se ha iniciado el proceso de extender la funcionalidad del framework a la evaluación de calidad en ambientes ágiles incorporando un nuevo modelo y herramientas de soporte que automaticen el proceso de medición y evaluación, dando soporte a la gestión de elementos del modelo de calidad.

Agradecimientos

Este trabajo se ha realizado en el marco del proyecto de investigación acreditado “Framework para la evaluación de Calidad del Software”, EIUTIRE0002205TC de la Universidad Tecnológica Nacional.

Cabe destacar también que el artículo se enmarca en las actividades planificadas en el Proyecto de Desarrollo Tecnológico y Social (PDTS) presentado, “Aporte a la competitividad de las empresas de desarrollo de Software del NEA”, IP253, evaluado y aprobado por el Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina (CONICET).

Referencias

[1] ISO. (2001). *Sistemas de gestión de la calidad-Conceptos y vocabulario. Norma Internacional ISO 9000*.

[2] Pinto, N. S., Tortosa, N. G., Pletsch, L. R. C., Acuña, C. J., Greiner, C., & Estayno, M. (2013). Aproximación a la Evaluación de la Calidad de Aplicaciones Web. *Ciencia y Tecnología*, 1(13).

[3] Abrahão, S. M., Pastor, O., Olsina, L., & Fons, J. (2001, November). Un Método para Medir el Tamaño Funcional y Evaluar la Calidad de Sitios Web. In *JISBD* (pp. 477-490).

[4] Tortosa, N., Pinto, N., Acuña, C., Cuenca Pletsch, L. R., & Estayno, M. G. (2013). QUCO2: development of a tool for measuring the quality of web applications. In *XVIII Congreso Argentino de Ciencias de la Computación*.

[5] Pinto, N., Tortosa, N., Cuenca Pletsch, L., Acuña, C., Seba, Sebastián. QuVi: Un enfoque estadístico para la obtención de sistemas Web de Calidad. CONAISI 2014. ISSN N° 2346-9927.

[6] Información disponible en <http://www.alexa.com/> Último acceso 05/2016

[7] Información disponible en <http://nibbler.silktide.com/> Último acceso 05/2016

EVALUACIÓN DE CALIDAD DE APLICACIONES WEB: EXPERIENCIA EN UNA PYME DEL POLO IT CHACO

[8] Pinto, Noelia; Tortosa, Nicolás; Acuña, César; Cuenca Pletsch, Lilian; Estayno, Marcelo. “Evaluación de Calidad de Aplicaciones Web asistida por herramientas tecnológicas”. WICC 2013. ISBN 978-987-28179-6-1.

[9] Información disponible en <https://www.djangoproject.com/> Último acceso 05/2016