

Análisis de métricas de calidad en foros de discusión técnicos

Nadina Martínez Carod, Pamela Faraci, Gabriela Aranda

Grupo GIISCo

Facultad de Informática, Universidad Nacional del Comahue

Buenos Aires 1400 (8300) Neuquén, Argentina

{nadina.martinez|gabriela.aranda}@fi.uncoma.edu.ar,

pamelafaraci@yahoo.com.ar

Resumen Si bien los foros de discusión son consultados asiduamente en la Web para resolver problemas, muchas veces la información contenida en ellos no está bien organizada o no es la adecuada en términos de solución, razón por la cual se vuelve indispensable poder definir algunos criterios de calidad de la información existente en dichos foros. La propuesta es utilizar los foros como sistemas de recupero de información (SRI), con capacidad para almacenar, mantener y recuperar información. La información resultante estará ordenada de acuerdo a alguna prioridad generada considerando varias características de calidad. En esta propuesta se definen algunas métricas para medir la calidad de la información contenida en hilos de discusión técnicos, y se analizan las métricas propuestas para especificar un ranking en la información recuperada, se ejemplifica a través de un caso de estudio.

1. Introducción

Con el transcurso del tiempo la Web se ha convertido en un gran repositorio de datos donde se intercambia conocimiento, ideas y se producen debates. Las herramientas colaborativas son aplicaciones dinámicas que se caracterizan por tener comunidades donde el mayor énfasis se da a la contribución y participación de los usuarios; dentro de las herramientas existentes, los foros de discusión se destacan por ser generalmente utilizados para resolver inconvenientes que pudieran surgir en cualquier momento. No sólo son empleados por los usuarios que pertenecen a la misma comunidad como elemento de debate, sino que muchas veces pueden ser accedidos por usuarios extenos, como fuente de conocimiento. El crecimiento desmedido de la Web trajo como consecuencia algunas situaciones no deseables, en particular la existencia de información similar repetida en diferentes sitios, incluso muchas veces mal organizada. De esta manera, para poder acceder a la información almacenada en los foros de discusión, a menudo es necesario navegar por varios hilos hasta dar con una solución factible o con la solución adecuada, razón por la cual, se consideró la utilización de las características de calidad para evaluar las soluciones encontradas en diferentes foros de discusión [1].

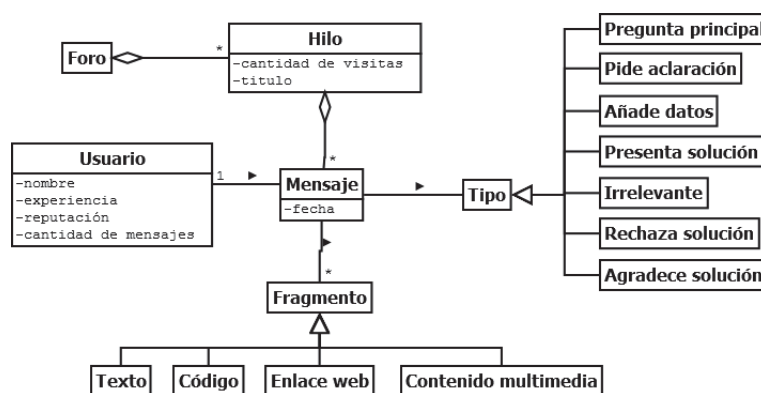


Figura 1: Modelo Conceptual de foros de discusión técnicos

De la revisión formal realizada sobre un grupo de foros de discusión, se desprende la metodología de búsqueda, donde una persona, perteneciente a la comunidad del mismo, realiza una pregunta, la cual llamamos pregunta principal, a partir de la misma se genera un hilo de discusión. Dicho hilo está formado por mensajes, los cuales a su vez están constituidos por fragmentos. Al ser un foro de discusión técnico, los fragmentos además de ser texto, gráficos, sonidos, videos, pueden contener código de lenguajes de programación. La Figura 1 muestra el modelo conceptual rectificado, diferenciado con el modelo presentado en [2], en el cual se identifican las componentes de un foro de discusión así como sus relaciones.

Una actividad típica en la Web es la consulta a través de algún motor de búsqueda multipropósito, la cual trae como resultado una lista de objetos relacionados a dicha consulta, motivo por el cual se desea que la lista resultante contenga hilos de discusión referidos al tema en cuestión. Luego de que una persona realiza una consulta, ésta necesitará visitar cada uno de los elementos de la lista resultante del motor antes de encontrar una propuesta de solución para su problema, para luego analizar si es la correcta. De esta manera la determinación de criterios de evaluación de la información contenida en los foros podría influir en el orden óptimo de visita de un individuo sobre la lista resultante.

Dentro de la herramienta propuesta en [3], uno de los procesos que la componen, realiza el análisis de los hilos de discusión existentes en foros técnicos para, dado un problema particular, sugerir un conjunto de soluciones exitosas en base al análisis de atributos en los hilos de discusión clasificados por tema, ordenado de mayor a menor probabilidad de éxito.

Del modelo de calidad definido para foros de discusión se desprende un conjunto de métricas. Tanto el modelo como las métricas propuestas son detallados en la Sección 2. En la Sección 3 se detalla la especificación del caso de estudio, cuyos resultados serán analizados en la Sección 4, finalizando con las conclusiones y lineamientos de posibles trabajos futuros en la Sección 5.

		Dimensiones de Calidad																
		Exactitud	Complejidad	Consistencia	Credibilidad	Actualidad	Accesibilidad	Conformidad	Confidencialidad	Eficiencia	Precisión	Trasabilidad	Entendibilidad	Disponibilidad	Portabilidad	Recuperabilidad	Valor Agregado	Adec. Represent.
		Atributos																
HILO	título	●	●															
	cantidad visitas	●	●							●								
USUARIO	nombre																	
	cantidad mensajes	●	●															
	experiencia	●	●		●	●				●	●							
	reputación	●	●		●				●	●		●						
fragmento	texto	●	●									●					●	
	código	●	●					●			●				●		●	
	enlace web	●	●								●		●				●	
	figura	●									●		●				●	
	fecha					●												
MENSAJE	pregunta principal	●	●															
	añade datos	●	●							●	●	●	●	●			●	●
	irrelevante																	
	agradece solución	●	●		●					●	●	●	●	●	●		●	●
	presenta solución	●	●		●					●	●	●	●	●	●		●	●
	pide aclaración																	
	rechaza solución																●	

Figura 2: Modelo conceptual relacionado con características de calidad

2. Modelo de Calidad de Datos

Debido a que se desea adquirir información a partir de los foros de discusión, fue necesario el estudio de las características de calidad para clasificar los datos. En [1] se presentó la matriz que relaciona ciertas características de calidad de datos a atributos del modelo conceptual planteado. Las dimensiones de calidad de la matriz se basan en el estándar ISO/IEC 25012:2008 [4] extendiendo dicho modelo con dos características propuestas por Rafique [5]. De esta manera surge un modelo de calidad de información para foros de discusión, que incluyen las dimensiones de calidad del modelo 25012 con dos dimensiones de calidad agregadas, que son el valor agregado y la adecuación representacional. En la Figura 2 se puede observar con un círculo negro para cada uno de los atributos del modelo conceptual, su posible relación con cada uno de los atributos de calidad definidos, por ejemplo en la fila correspondiente a la reputación se encuentra marcada como posible la relación con el atributo de calidad credibilidad.

En base a los cuadros seleccionados se proponen algunas métricas de calidad correspondientes a las dimensiones de exactitud, credibilidad y portabilidad. Las métricas de exactitud sugieren cuan acertado es un hilo con respecto a la cadena de búsqueda, las métricas de portabilidad aluden a si los enlaces expresados en un hilo, están operativos, y algunas métricas de credibilidad están enfocadas en el grado de expertitud de las personas que intervienen en un hilo.

A continuación se define un grupo de métricas de calidad.

Tabla 1: Plantilla de descripción de Métricas

Nombre (1)	Sigla (2)	Tipo (3)
Nombre en castellano (4)	Descripción de la métrica (5)	
Entidad relacionada (6)		
Atributo relacionado (7)	Descripción del atributo relacionado (8)	
Tipo de escala (9)	Tipo de valor (10)	Unidad (11)
Método de medición (12)	Tipo de método de medición (13)	

2.1. Métricas

Siguiendo la propuesta presentada por Olsina y Martín en [6], la Tabla 1 representa una plantilla para describir las características que posee una métrica de calidad para hilos de discusión técnicos. Una métrica recibe un nombre (1) que generalmente se presenta en inglés y en castellano (4).

Los ítems descriptivos de las métricas son:

- Sigla (2): Forma abreviada de mencionar a la métrica. Por convención se usa el símbolo # (hacen referencia a cantidad) y el símbolo % (porcentaje).
- Tipo (3): El tipo puede ser directa, si no depende de métricas de otros atributos, o indirecta en caso que sí dependa.
- Descripción (5): Explica lo que contabiliza la métrica.
- Entidad relacionada (6): Objeto sobre el cual se realiza la medición (por ejemplo, en nuestro caso la entidad relacionada es el hilo de discusión).
- Atributo relacionado (7): Propiedad física o abstracta medible de una entidad.
- Descripción del atributo relacionado (8): Describe las características del atributo.
- Tipo de escala (9): Puede ser *nominal* (con la especificación de la lista de valores), *ordinal*, *intervalo*, *proporción* o *absoluto*.
- Tipo de valor (10): En escalas nominales es la lista de valores posibles correspondientes a la categorización, en otros casos representará una lista o rango de valores numéricos.
- Unidad (11): Objeto referencia para dimensionar la medición (por ejemplo palabras, líneas de código, mensajes, etc.).
- Método de medición (12): Secuencia lógica de operaciones para realizar la medición que se describe.
- Tipo de método de medición (13): Puede ser de tipo subjetivo (implica juicio o criterio humano), u objetivo (basada en reglas numéricas).

2.2. Especificación de las Métricas

Se definieron 8 métricas, algunas correspondientes a la pertinencia de un hilo de discusión, otras sobre la disponibilidad de información y las últimas sobre expertitud dentro en un hilo de discusión. Todas ellas cumplen las mismas características con respecto a su entidad y a su escala, su entidad son los hilos de discusión y su escala es absoluta. El tipo de valor de las métricas directas

Tabla 2: Métricas directas

Nombre	Nombre castellano	Sigla	Descripción	Atributo relacionado	Descripción atributo relacionado	Unidad	Método de medición
Number of keywords in the title	Cantidad de palabras clave en el título	#KT	palabras clave que aparecen en el título	Pertinencia del hilo	cuan pertinente es el hilo de discusión para la cadena de búsqueda ingresada por un usuario externo en un buscador	palabra	Contar la cantidad de palabras clave que aparecen en el título
Number of keywords in the main question	Cantidad de palabras clave en la pregunta principal	#KM Q	palabras clave que aparecen en la pregunta principal	Pertinencia del hilo	cuan pertinente es el hilo de discusión para la cadena de búsqueda ingresada por un usuario externo en un buscador	palabra	Contar
Number of operational links	Cantidad de links operativos	#OL	links operativos presente en el hilo de discusión	Link operativo	Enlaces web que al referenciar no presentan errores del tipo 404, es decir su contenido puede ser observado	link	contar links operativos a lo largo de todo el hilo de discusión
Number of expert users	Cantidad de usuarios expertos	#EU	usuarios expertos que participan en el hilo de discusión	Expertitud del hilo	usuarios expertos pertenecientes a la comunidad del foro que participan en el hilo de discusión	mensaje	contar la cantidad de usuarios expertos que participan en el hilo de discusión

Tabla 3: Métricas indirectas

Nombre	Nombre castellano	Sigla	Descripción	Atributo relacionado	Descripción atributo relacionado	Método de cálculo
Percentage of keywords in the title	Porcentaje de palabras clave en el título	%KT	idem a descripción de #KT	Pertinencia del hilo	idem a descripción del atributo de #KT	(cantidad de palabras clave encontradas en el título/cantidad de palabras clave buscadas)*100
Percentage of keywords in the main question	Porcentaje de palabras clave en la pregunta principal	%KM Q	idem a descripción de #KM Q	Pertinencia del hilo	idem a descripción del atributo de #KM Q	(cantidad de palabras clave encontradas en la pregunta principal/cantidad de palabras clave buscadas)*100
Percentage of operational links	Porcentaje de links operativos	%OL	idem a descripción de #OL	Link operativo	idem a descripción del atributo de #OL	(cantidad de links operativos/cantidad de links en el hilo)*100
Percentage of expert users	Porcentaje de usuarios expertos	%EU	idem a descripción de #EU	Expertitud del hilo	idem a descripción del atributo de #EU	(cantidad de usuarios experto/cantidad de usuarios)*100

definidas es numérico, específicamente entero positivo, mientras que en todos los casos el tipo de medición es objetivo. Los otros detalles de las métricas directas visualizan en la Tabla 2, donde se puede observar por ejemplo que las siglas en las métricas directas comienzan con el símbolo # debido a que hacen referencia a cantidad de elementos.

Las métricas indirectas definidas, están en función de las directas por lo cual, tanto la descripción de la métrica como del atributo relacionado correspondiente son similares a las directas referenciadas en la Tabla 2. Se puede observar en la Tabla 3 que las siglas comienzan con el símbolo % y que en vez de método de medición se especifica el método de cálculo. Si bien comparten el tipo de valor numérico, a diferencia de las métricas directas, los métodos de cálculo corresponden a valores reales positivos, en estos casos la unidad es normalizada a porcentaje.

Existen algunos casos especiales en las métricas indirectas que son cuando la cantidad de lo que se busca es nula y la cantidad total de referencia del atributo relacionado también. En esos casos (%KT, %KM Q, %EU) se consideran con un valor de 0 %. Salvo en la métrica correspondiente a la disponibilidad de información (%OL) que corresponde a los casos en los cuales no tienen links relacionados; en estas ocasiones se puede decir que la información está disponible en su totalidad, por lo que se define un valor del 100 %.

La aplicación de las métricas sobre un conjunto de hilos de discusión ayudaría a establecer prioridades en la búsqueda de soluciones, para ello la propuesta es utilizar las cuatro métricas indirectas detalladas para determinar un orden de prioridad al adquirir información en debates.

2.3. Pesos relativos y Ranking

Con las métricas indirectas propuestas, se va a definir el ranking para cada hilo obtenido como resultado de una consulta; con el fin de lograr este objetivo se propone la asignación de un vector de pesos relativos, definido con un peso específico otorgado a cada una de las métricas indirectas. El vector se conforma por los siguientes pesos relativos:

- PRKT = Peso relativo otorgado a la métrica %KT
- PRKMQ = Peso relativo otorgado a la métrica %KMQ
- PROL = Peso relativo otorgado a la métrica %OL
- PREU = Peso relativo otorgado a la métrica %EU

Para un primer análisis, se considerará el mismo peso a cada una de las métricas, esto es un valor de 0.25. Entonces, en base a las métricas indirectas presentadas y al vector de pesos relativos, se determina el ranking que representa la relevancia de un hilo de discusión respecto a una cadena de búsqueda dada, al cual llamaremos $Ranking_{Tool}$. La forma de obtener dicho ranking se define con la siguiente fórmula:

$$Ranking_{Tool} = \%KT * PRKT + \%KMQ * PRKMQ + \%OL * PROL + \%EU * PREU$$

Por lo tanto, el valor $Ranking_{Tool}$ es el valor relativo asignado a un hilo de un foro de discusión respecto a una cadena de búsqueda. De esta manera, para obtener la proporción del ranking con respecto a la sumatoria de los valores obtenidos para todos los hilos que forman parte del conjunto bajo estudio, se utiliza la siguiente ecuación:

$$\%Relativo\ Ranking_{Tool} = \frac{Ranking_{ForumMet}\ del\ hilo\ i}{\sum_{j=1}^n Ranking_{ForumMet}\ del\ hilo\ j}$$

Para analizar el ranking dentro de un conjunto de hilos utilizaremos un ejemplo obtenido a partir de un caso de estudio.

3. Caso de Estudio

Un caso de estudio se desarrolla con el objetivo de investigar una entidad o fenómeno particular en el contexto de la vida real [7]. Los resultados al aplicar las métricas sobre 3 hilos de discusión recuperados de la Web utilizando la cadena de búsqueda: “*error en java nullpointerexception*” fueron los siguientes:

Hilo 1

- Las métricas retornadas fueron $\#KT = 1$, $\%KT = 33,33\%$, $\#KMQ = 1$ y $\%KMQ = 33,33\%$. De las 3 palabras clave de la cadena de búsqueda (error, java y nullpointerexception), en el título aparece solo la palabra *nullpointerexception*
- Este hilo presenta un link a contenido externo $\#OL = 1$ y $\%OL = 100\%$, dado que dicho enlace está operativo.
- Los resultados para estas métricas fueron $\#EU = 0$ y $\%EU = 0\%$.

Hilo 2

- Se obtuvieron $\#KT = 2$ y $\%KT = 66,66\%$. En la pregunta principal sucede algo similar y los resultados obtenidos fueron $\#KMQ = 2$ y $\%KMQ = 66,66\%$.
- En la Respuesta 3 aparece un link a las preguntas frecuentes de Java (FAQ's de Java), el cual está operativo. Luego las métricas retornaron los siguientes resultados $\#OL = 1$ y $\%OL = 100\%$.
- Dado que el foro no define roles para sus miembros, las métricas obtenidas fueron $\#EU = 0$ y $\%EU = 0\%$.

Hilo 3

- Se obtuvieron las mediciones $\#KT = 3$, $\#KMQ = 3$, $\%KT = 100\%$, $\%KMQ = 100\%$.
- El hilo no contiene links, por lo tanto las métricas para enlaces operativos presentaron los valores $\#OL = 0$ y $\%OL = 100\%$. Esto significa que está operativo en su totalidad.
- El foro define roles para sus miembros, en las respuestas participa un *moderador global avanzado* y el usuario que inició la pregunta. Por lo tanto se obtuvo $\#EU = 1$ y $\%EU = 100\%$.

En la Figura 3 se presenta los gráficos comparativos de las métricas del caso de estudio, en la Figura 3 (a) se comparan para los 3 hilos las métricas directas o de cantidad (#) y las indirectas o de porcentaje (%) en la Figura 3b(b).

4. Validación

Con el objetivo de evaluar los resultados obtenidos en el casos de estudio se elaboró una encuesta donde se solicitó a 10 informáticos, con amplios conocimientos en programación Java y usuarios frecuentes de foros de discusión, que dada una cadena de búsqueda establecieran el orden de relevancia (1, 2 o 3) de los 3 hilos tomados como base del caso de estudio, según el siguiente criterio:

1. Es el hilo más relacionado a la cadena de búsqueda y ofrece soluciones muy pertinentes (sin dudas, lo visitaría primero).

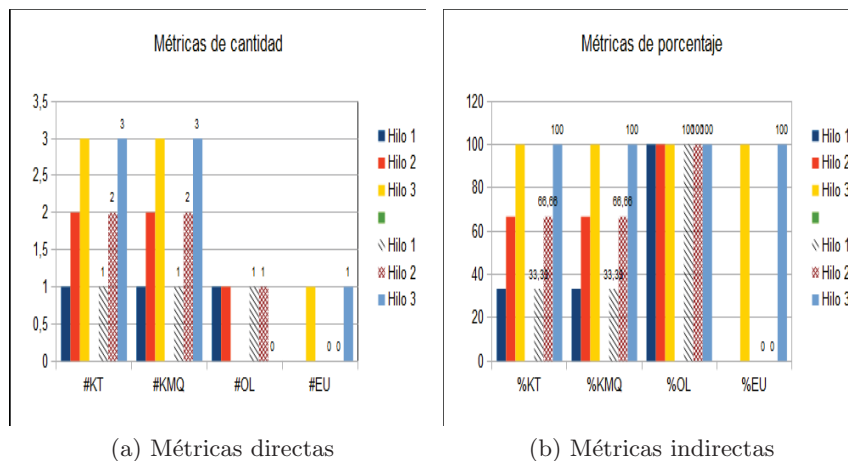


Figura 3: Métricas obtenidas en el caso de estudio

2. El hilo está muy relacionado a la cadena de búsqueda, pero hay otro que está más relacionado aún.
3. Es el hilo menos relacionado a la cadena de búsqueda de los tres.

Las opciones fueron ponderadas para poder calcular el orden de relevancia de cada hilo para cada uno de los expertos, dándole una puntuación de 20 a la opción 1, 10 puntos a la opción 2 y 0 puntos a la última opción.

En la Tabla 4 se presentan las respuestas que se obtuvieron por parte de los 10 expertos consultados. Por ejemplo en la Tabla 4 (a) se observa en la primera fila que el primer experto le dio el orden de mayor importancia al hilo 2, en segundo lugar el hilo 1 y por último el hilo 3 mientras que el segundo experto consideró con mayor importancia el hilo 3. Los valores ponderados para los órdenes de primero (1), segundo (2) y tercero (3) son 20, 10 y 0 respectivamente. En la Tabla 4 (b) se muestra los valores asignados ya ponderados, donde en la fila inferior se destaca la suma de los valores ponderados para cada hilo (por ejemplo, la suma de los valores ponderados para el Hilo 1 es 90, para el Hilo 2 es 60 y para el Hilo 3 es 150).

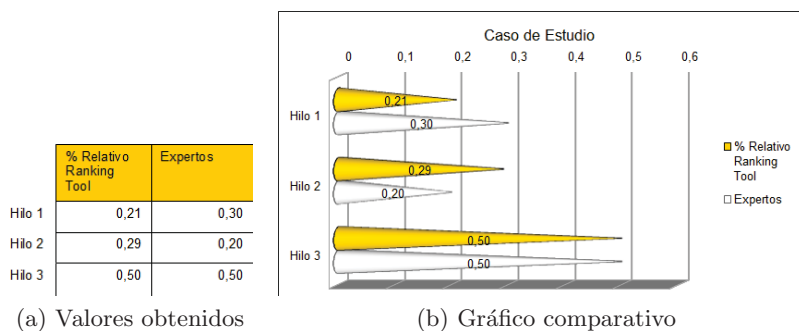
Orden hilo 1	Orden hilo2	Orden hilo3
2	1	3
2	3	1
2	3	1
3	2	1
3	2	1
1	2	3
3	2	1
1	3	2

Hilo 1	Hilo 2	Hilo 3
10	20	0
10	0	20
10	0	20
0	10	20
0	10	20
20	10	0
0	10	20
20	0	10
10	0	20
10	0	20
90	60	150

(a) Asignado (b) Ponderado

Tabla 4: Orden según encuesta a los expertos

Al analizar el caso de estudio, se destaca en la Tabla 4(a), la comparativa con los valores del %Relativo Ranking_{Tool} en la primer columna y el resultado de los encuestados en la columna titulada Expertos, mientras que en la Figura 4(b) se puede visualizar el gráfico comparativo de los datos de dicha tabla. Al comparar los resultados se puede observar que tanto el ranking relativo como los expertos sugieren el Hilo 3 en el primer lugar del ranking de visita. Aunque los Hilos 1 y 2 se encuentran en distinto orden de visita, la diferencia no sería significativa ($\leq 0,10$ en ambos casos).



(a) Valores obtenidos (b) Gráfico comparativo

Figura 4: Comparación Ranking vs. valoración de Expertos

5. Conclusiones y trabajo futuro

En este trabajo se presentan algunas métricas para medir la calidad de la información contenida en los hilos de discusión, se establece un criterio para determinar a partir de las métricas el orden de prioridad brindando un peso

específico a cada una de las métricas. Luego utilizando un caso de estudio enfocado en hilos de foros de discusión técnicos, y utilizando para validar criterios de expertos, se contrastaron los resultados obtenidos. Al comparar los datos de la encuesta realizada a un conjunto de expertos y el orden sugerido, se puede detectar que estos no difieren en gran medida. Por un lado, el orden de relevancia coincidió en el hilo más relevante y, aunque otorgó un orden distinto para el segundo y tercer hilo, la diferencia en ambos casos fue menor a 0,10 en una escala de 0 a 1, por lo que esta diferencia podría considerarse no significativa. Esta primer aproximación invita a contrastar resultados utilizando las mismas métricas en diferentes casos de estudio. A futuro se planea analizar bajo qué contexto una métrica podría ser más relevante que otra y ponderar de manera distinta el vector de pesos relativos. Si bien es un resultado parcial, en este caso existe coincidencia entre el ranking relativo y el otorgado por los expertos, es necesario entonces extender los casos de estudio y replicar los experimentos realizados.

Agradecimientos

Este trabajo está parcialmente soportado por el subproyecto “*Reuso de Conocimiento en Foros de Discusión, Parte II*”, correspondiente al Programa de Investigación 04/F009 “*Desarrollo Orientado a Reuso, Parte II*”, de la Universidad Nacional del Comahue (Neuquén, Argentina).

Referencias

1. G. Aranda, N. Martínez, P. Faraci, and A. Cechich, “Hacia un framework de evaluación de calidad de información en foros de discusión técnicos,” in *ASSE 2013-Simposio Argentino de Ingeniería de Software, JAIIO 42^o-Jornadas Argentinas de Informática*, (Córdoba, Argentina), SADIO, 2013.
2. Martínez Carod, Nadina, Aranda, Gabriela, Sagripanti, Mauro, Faraci, Pamela, and Cechich, Alejandra, “Análisis de la información presente en foros de discusión técnicos,” in *CACIC 2013, XIX Congreso Argentino de Ciencias de la Computación*, (Mar del Plata, Argentina), pp. 847–856, Oct. 2013.
3. Aranda, Gabriela, Martínez Carod, Nadina, Roger, Sandra, Faraci, Pamela, and Cechich, Alejandra, “Una herramienta para el análisis de hilos de discusión técnicos,” in *CACIC 2014, XX Congreso Argentino de Ciencias de la Computación*, (San Justo, Argentina), pp. 803 – 812, Oct. 2014.
4. “ISO/IEC 25012:2008, Software product quality requirements and evaluation (SQuaRE): Data quality model,” 2008.
5. I. Rafique, P. Lew, M. Q. Abbasi, and Z. Li, “Information quality evaluation framework: Extending ISO 25012 data quality model,” *International Journal of Computer and Information Sciences*, vol. 6, no. 1, pp. 1 – 6, 2012.
6. L. Olsina and M. d. l. A. Martín, “Ontology for software metrics and indicators,” *Journal of Web Engineering*, vol. 2, no. 4, pp. 262–281, 2004.
7. C. Wohlin, P. Runeson, M. Höst, M. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering*. Computer Science, Springer, 2012.