



**ESCUELA POLITÉCNICA NACIONAL  
FACULTAD DE INGENIERÍA DE SISTEMAS  
DEPARTAMENTO DE INFORMÁTICA  
Y CIENCIAS DE LA COMPUTACIÓN**



**PROYECTO INTERNO DE INVESTIGACION**

**Área del proyecto:** Ciencias de la Computación    Ciencias Básicas     Ciencias Aplicadas

**FACULTAD:** INGENIERÍA DE SISTEMAS

**DEPARTAMENTO:** INFORMÁTICA Y CIENCIAS DE LA COMPUTACIÓN

**LINEA DE INVESTIGACIÓN:**  
SISTEMAS INTELIGENTES

**Procesamiento de Lenguaje Natural – Análisis de contexto de citas bibliográficas**

**1 Proyecto de Investigación Interno**

**Título:**  
**PROPUESTA DE VALORACIÓN DEL IMPACTO PARA CITAS BIBLIOGRÁFICAS TOMANDO EN CUENTA CRITERIOS CUALITATIVOS**

**Resumen del proyecto**

La valoración de la influencia de una cita bibliográfica puede variar si se toma en cuenta el propósito y la disposición del autor para referenciar la cita, con respecto a si se usan métodos de simple conteo tomando a todas las citas como iguales. Una cita en un documento puede tener intención de crítica, puede ser punto de partida de un trabajo o, simplemente tener como objeto el reconocimiento de aportes de otros autores; la disposición de autor que menciona a otros también puede variar y ser positiva, negativa o neutral.

Con los métodos actuales de simple conteo, todas las citas tienen el mismo peso en la evaluación del impacto. Sin embargo, debido a que no todas las citas tienen la misma función o polaridad dentro del documento, se hace necesario que, para valorar el impacto, se considere estos criterios que marcan la influencia real de un artículo sobre otro más allá del número de veces que se lo menciona. En esta evaluación son importantes criterios de calidad como la polaridad de la cita y el papel o función que cumple dentro del documento.

La medición de impacto que se realiza actualmente, al no tomar en cuenta diferencias entre los distintos tipos de cita, provoca deformaciones que han sido analizadas por Hernández y Gómez (2015), por lo que es crucial proponer métricas más completas para enriquecer el análisis con criterios cualitativos como los sugeridos. Para lograr este objetivo es necesario realizar un análisis de contexto de las citas para obtener las características que son relevantes para su clasificación.

Para este análisis se usará el esquema de clasificación y el corpus generado por el proyecto interno anterior desarrollado por mi persona, entregado en abril 2015.

Palabras clave (3-5): Análisis, sentimientos, polaridad, WEKA, citas, bibliográficas, mediciones, índice, Hirsch

**2 Datos personales y académicos del Director del Proyecto**

Apellidos: HERNÁNDEZ ALVAREZ	Dirección particular: Rigoberto Heredia OE6-134 y Huachi
Nombres: MYRIAM BEATRIZ	
Lugar y fecha de nacimiento: QUITO, 21 de julio 1957	Teléfono casa: 2599846
Cargo actual en la EPN: PROFESOR A TIEMPO COMPLETO	Teléfono celular: 0992715620
Fecha nombramiento definitivo: noviembre 1987	Teléfono oficina: 2507144
Horas de dedicación al proyecto: 200 semestre	Ext. EPN: 2650
	Correo electrónico: myriam.hernandez@epn.edu.ec

**ESCUELA POLITÉCNICA NACIONAL  
CONSEJO ACADÉMICO**

- 2 -

<b>Formación de pregrado y postgrado</b>		
Ingeniero en Electrónica y Telecomunicaciones	Agosto 1982	Escuela Politécnica Nacional, Quito, Ecuador
Master in Computer Science	Julio 1987	Ohio University at Athens Ohio, Estados Unidos
Especialista en Dirección de Empresas	Agosto 1998	Universidad Andina Simón Bolívar, Quito, Ecuador

<b>3</b>	<p><b>Objetivos, hipótesis y resultados esperados de esta propuesta de investigación</b></p> <ul style="list-style-type: none"> <li>- <b>Objetivos</b> <ul style="list-style-type: none"> <li>○ Usar el esquema de clasificación desarrollado en el proyecto interno anterior a cargo de mi persona para definir un método integral de valoración del impacto tomando en cuenta criterios de calidad de las citas bibliográficas, relacionadas con función y polaridad.</li> <li>○ Realizar una investigación teniendo como fuente a autores de artículos científicos que proporcionen información sobre el nivel de influencia de las distintas citas referenciadas en sus documentos.</li> <li>○ Desarrollar una función para valorar el impacto y la influencia de las citas bibliográficas.</li> <li>○ Aplicar esta función para predecir el impacto en un corpus anotado especializado.</li> </ul> </li> <li>- <b>Hipótesis</b> Es posible mejorar los índices de impacto de citas bibliográficas usados actualmente para valorar la influencia de los autores en los distintos campos de la ciencia, usando información de función, polaridad, frecuencia y ubicación de las citas bibliográficas</li> <li>- <b>Resultados esperados</b> <ul style="list-style-type: none"> <li>○ Resultados primarios obtenidos de autores para la valoración del impacto de las citas referenciadas en sus propios artículos.</li> <li>○ Función para valoración del impacto basada en criterios de influencia de un artículo sobre otro.</li> <li>○ Clasificación automática para valoración de impacto en artículos dentro de un corpus especializado.</li> <li>○ Artículo aprobado para su publicación en un congreso o revista indexados en Latindex o Scopus</li> </ul> </li> <li>- <b>Potenciales Usuarios</b> <ul style="list-style-type: none"> <li>○ Científicos en el área de análisis de sentimientos para citas bibliográficas presentadas en publicaciones científicas.</li> <li>○ Científicos que estudian posibilidades de mejora en la metodología para el cálculo de índices de impacto.</li> <li>○ Comunidad científica que podría beneficiarse de índices de impacto mejorados que usen no solo criterios cuantitativos sino también de calidad de las citas bibliográficas.</li> </ul> </li> </ul>
<b>4</b>	<p><b>Relevancia de esta propuesta de investigación con los objetivos científicos del departamento y su Línea de Investigación.</b></p> <p>El presente proyecto de investigación está dentro del área de Ciencias de la Computación en la línea de investigación de Sistemas Inteligentes del DICC.</p>

**ESCUELA POLITÉCNICA NACIONAL  
CONSEJO ACADÉMICO**

- 3 -

<b>5</b>	<p><b>Descripción del proyecto, metodología, cronograma de trabajo</b></p> <p><b>- Descripción del proyecto</b> En el proyecto propuesto se plantea realizar una investigación primaria de autores científicos para obtener información que nos sirva para definir una función integral para la valoración del impacto de un artículo sobre otro, tomando en cuenta criterios como función, polaridad, frecuencia y ubicación de las citas. Se validará esta función comparando resultados con los datos primarios. Una vez validada la función, se la usará para valorar el impacto de citas dentro de artículos científicos anotados en un corpus especializado usando herramientas de regresión en Weka.</p> <p><b>- Metodología y diseño de la investigación</b> En el proyecto de investigación se usará una metodología ágil, se generarán productos intermedios que serán el resultado de cada una de las actividades y llevarán a configurar los productos finales a través de revisiones e iteraciones cortas.</p> <p><b>Cronograma de trabajo anual</b></p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2">Actividad</th> <th colspan="6">MESES</th> </tr> <tr> <th>1-2</th> <th>3-4</th> <th>5-6</th> <th>7-8</th> <th>9-10</th> <th>11-12</th> </tr> </thead> <tbody> <tr> <td>Preparación de investigación primaria</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Envío de encuestas a autores</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Tabulación de resultados y obtención de una función que modele la evaluación del impacto</td> <td></td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td></td> <td></td> </tr> <tr> <td>Validación de la función de impacto que se genere</td> <td></td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td></td> </tr> <tr> <td>Aplicación de la función en corpus especializado</td> <td></td> <td></td> <td></td> <td></td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> </tr> <tr> <td>Documentación, conclusiones, recomendaciones</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> <td style="text-align: center;">X</td> </tr> </tbody> </table>	Actividad	MESES						1-2	3-4	5-6	7-8	9-10	11-12	Preparación de investigación primaria	X	X					Envío de encuestas a autores	X	X					Tabulación de resultados y obtención de una función que modele la evaluación del impacto		X	X	X			Validación de la función de impacto que se genere		X	X	X	X		Aplicación de la función en corpus especializado					X	X	Documentación, conclusiones, recomendaciones	X	X	X	X	X	X
Actividad	MESES																																																							
	1-2	3-4	5-6	7-8	9-10	11-12																																																		
Preparación de investigación primaria	X	X																																																						
Envío de encuestas a autores	X	X																																																						
Tabulación de resultados y obtención de una función que modele la evaluación del impacto		X	X	X																																																				
Validación de la función de impacto que se genere		X	X	X	X																																																			
Aplicación de la función en corpus especializado					X	X																																																		
Documentación, conclusiones, recomendaciones	X	X	X	X	X	X																																																		
<b>6</b>	<p><b>Fecha de inicio</b> 20 de mayo 2015</p>																																																							
<b>7</b>	<p><b>Tiempo dedicación docente y asistente técnico administrativo</b></p> <p>- Dedicación del Director del Proyecto: 200 horas – semestre.</p>																																																							

ESCUELA POLITÉCNICA NACIONAL  
CONSEJO ACADÉMICO

- 4 -

8	<b>Presupuesto estimado para la ejecución del presente proyecto</b>	
	Lista de ítems	<b>Cantidad solicitada (US \$)</b>
	Estancia de investigación en Universidad de Alicante para levantamiento de información primaria y trabajo colaborativo con Dr. José Gómez colaborador del proyecto.	5000,00
	<b>Subtotal</b>	5000,00
	<b>TOTAL</b>	5000,00
9	<b>Firma del aplicante</b>	<b>Lugar y Fecha</b>
	 Nombre: Ing. MSc. Myriam Hernández A. CC: 1705009304	Quito, 13 de mayo de 2015
<b>DECLARACION DEL JEFE DE DEPARTAMENTO</b>		
Esta propuesta ha sido aprobada por el Departamento de Informática y Ciencias de la Computación, y las instalaciones, incluyendo personal, edificios, equipo y recursos financieros están a disposición del aplicante de acuerdo con las especificaciones que se encuentran en esta aplicación.		
 JEFE DEL DEPARTAMENTO Nombre: Ing. MSc. Bolívar Palán Tamayo CC: 1801097427		Quito, 13 de mayo de 2015