

Proceedings of the

# **Workshop NLP in the Enterprise: Envisioning the Next 10 Years (PLN-E 2010)**

held at the

**XXVI Spanish Conference on Natural Language Processing  
(SEPLN 2010)**

7 September 2010

Valencia, Spain

José Carlos Cortizo<sup>1</sup>, José María Gómez<sup>2</sup>, Francisco Manuel Rangel<sup>3</sup>, Victor Peinado<sup>4</sup>,  
Hugo Zaragoza<sup>5</sup>, Francisco M. Carrero<sup>1</sup>

<sup>1</sup> BrainSINS, Madrid, Spain

<sup>2</sup> Optenet, Madrid, Spain

<sup>3</sup> Corex, Valencia, Spain

<sup>4</sup> Universidad Nacional de Educación a Distancia, Madrid, Spain

<sup>5</sup> Yahoo! Research, Barcelona, Spain



## **Prefacio**

¿Cuáles son las líneas de investigación (presentes y futuras) dentro del PLN que más interesan a las empresas? ¿Qué tecnologías dentro del NLP tendrán un mayor impacto empresarial o en nuestro día a día en los próximos 10 años? El objetivo de este Workshop se centra en responder estas preguntas y también servir como un punto de encuentro entre el mundo de la investigación universitaria y las empresas.

Si bien gracias a proyectos en cooperación entre Universidades/Centros de Investigación y Empresas, así como algunos consorcios como MAVIR que engloban a todo tipo de entidades trabajando conjuntamente en el desarrollo de tecnologías útiles con un importante foco en la transferencia tecnológica, la metodología de evaluación de los trabajos científicos sigue haciendo mucho foco en la evaluación por parte de pares, investigadores del mismo área con una visión algo más alejada de las necesidades tecnológicas de las empresas para desarrollar productos y servicios de interés para otras empresas y para usuarios finales.

Este Workshop trata de cambiar el paradigma de esta evaluación desde una perspectiva más científica a una perspectiva más aplicada y, para ello, el comité de programa encargado de evaluar los trabajos presentados se formó por científicos y expertos de empresas, centros tecnológicos y miembros de consorcios centrados en la transferencia de tecnología. De esta forma, se trató que los trabajos presentados tienen interés desde la perspectiva empresarial, sin importar tanto otros criterios más cercanos a la excelencia científica.

En el workshop PLN-E se presentaron 13 trabajos muy diversos, tanto de investigaciones realizadas desde empresas españolas dedicadas al PLN, o bien interesadas en el PLN como una herramienta para el desarrollo de productos que solventan problemas concretos, como de centros de investigación y Universidades, que muestran el potencial de estas investigaciones más de base en el desarrollo de tecnologías comerciales a medio/largo plazo.

*José Carlos Cortizo  
José María Gómez  
Francisco Manuel Rangel  
Victor Peinado  
Hugo Zaragoza  
Francisco Manuel Carrero*



## **Comité Organizador**

José Carlos Cortizo (BrainSINS)  
José María Gómez (Optenet)  
Francisco Manuel Rangel (Corex)  
Victor Peinado (Universidad Nacional de Educación a Distancia)  
Hugo Zaragoza (Yahoo! Research)  
Francisco Manuel Carrero (BrainSINS)

## **Comité de Programa**

Rodrigo Agerri (Vicomtech Research Centre)  
Matxalen Alfaro (Sarenet)  
Claudio Baccigalupo (VLEX)  
Sergio Berná (ExperienceOn)  
Enric Castellón (Thera)  
Juan Manuel Cigarrán (UNED, Consorcio MAVIR)  
Jesús Contreras (ISOCO)  
Javier Cuervo (Redepyme, EOI)  
Luis Ignacio Díaz (Acciona I+D)  
José Gregorio Escalada (Telefónica I+D)  
Diego Expósito (Answare Technologies)  
Angel Faus (VLEX)  
Jorge García Betanzos (Sarenet)  
Ricardo Farreres (Thera)  
Anabel Fraga (UC3M/REUSE)  
Juan Antonio Garrido (i2factory)  
Galo Gimenez (HP)  
Francisco Gomez Molinero (Visual Tools)  
José Carlos González (Daedalus)  
Carlos Gonzalez (ExperienceOn)  
Alberto Gragera (Tuenti)  
Francesc Grau (Conzentra)  
Carlos Lamas (T-Systems Iberia)  
Didac Lee (Inspirit, Spamina)  
Juan Llorens (UC3M/REUSE)  
Miguel Lucas (Acteo Soluciones)  
Diego Martín (Stratebi Business Solutions)  
Javier Martín (Loogic)  
Juan Carlos Martínez (Corex)  
Daniel Martinez (Indra Software Labs)  
Borja Monsalve (Social Gaming Platform)  
John Paul Moore (ATOS Research)  
César de Pablo (UC3M, Consorcio MAVIR)  
Enrique Puertas (UEM, Consorcio MAVIR)  
Joaquín Rieta (TICSinergies)  
Federico Rodríguez (Stratebi Business Solutions)  
Miguel Angel Rodríguez (Telefónica I+D)

Antonio Sánchez Valderrábanos (Bitext)  
Estela Saquete (UA, Consorcio MAVIR)  
Isabel Segura (UC3M, Consorcio MAVIR)  
Jim Shur (Strands)  
Jose Luis Suárez (Corex)  
Yaiza Temprado (Telefónica I+D)  
Marc Torrens (Strands)  
Paulo Villegas (Telefónica I+D)  
Pedro Vivancos (Vócali)  
David Zaragoza (Avanzis)

## Tabla de Contenidos

Lexisla: a Legislative Information Retrieval System.....	1
<i>Ismael Hasan, Javier Parapar, Alvaro Barreiro</i>	
mOCRA: Mobile OCR Application.....	5
<i>Xose R. De La Puente, Ismael Hasan</i>	
Empresa 2.0: Detección de Plagio y Análisis de Opiniones.....	9
<i>Enrique Vallés Balaguer, Paolo Rosso</i>	
PLN aplicado a Internet; en Búsqueda de la Subjetividad y Valoración Automática de los Textos.....	13
<i>Jorge García Betanzos</i>	
NaturalOpinions: Extracción de Opinión Basada en PLN para Contenidos Generados por los Usuarios.....	16
<i>Antonio S. Valderrábanos, Enrique Torrejón</i>	
Babxel: Búsqueda Multilingüe.....	21
<i>Andrés Velasco</i>	
Sistema de Información Aumentada para Dispositivos Móviles.....	26
<i>Luis Ignacio Díaz</i>	
Recomendaciones basadas en confianza para medios sociales.....	30
<i>Borja Monsalve Piqueras</i>	
Recuperación Aproximada de Direcciones Postales.....	34
<i>Yaiza Temprado</i>	
SIC+ Health 3.0 Sistema de Información Médica Personalizada.....	38
<i>Francisco M. Rangel, Enrique Vallés Juan C. Martínez, José L. Suárez, José Carlos Cortizo, Francisco M. Carrero</i>	
Sistemas Interactivos Multimodales de Procesamiento del Lenguaje Natural.....	42
<i>Elsa Cubel, Alejandro H. Toselli</i>	
Una aproximación a la mejora de los Motores de Búsqueda a través de la Recomendación Social de Contenido.....	46
<i>Luis Alberto Pérez García</i>	
Hacia la interacción en lenguaje natural.....	51
<i>Maria Fuentes y Meritxell González</i>	

