

Entrevista con el Dr. Ludovic Lebart

Ponente magistral del pasado XVI Congreso de la AMAI

YESENIA GONZÁLEZ PEDRAZA



El Dr. Lebart es una autoridad a nivel mundial en el tema de minería de textos. Su investigación se centra en los métodos de análisis estadístico de datos cualitativos y textuales. Los datos provenientes de encuestas constituyen su principal ámbito de aplicación. El Dr. Lebart es autor de más de catorce libros especializados en el área de *text mining* y ha participado como coautor en decenas de publicaciones relacionadas.

Desde 1990 Ludovic Lebart es presidente del Comité Científico del CESP (*Centre d' Etude des Supports de Publicité*: www.cesp.org). También es director de investigación científica en la Universidad Telecom-Paris Tech. Página de internet: <http://ses.telecom-paristech.fr/lebart/>

Dr. Lebart, usted es un ingeniero civil en minas, ¿cuándo y cómo empezó a trabajar en el área de minería de textos?

Bueno, en realidad estaba estudiando ingeniería en minas y desde el principio trabajé analizando datos para mejorar las operaciones mineras en geología. Entonces surgió la necesidad de realizar una encuesta sobre la industria de la transformación y fui el encargado de analizarla; naturalmente se requirió realizar preguntas abiertas y fue lo que me condujo al área de minería de textos.

Mucha gente dice que *text mining* se reduce al conteo de palabras y que esto no es suficiente para analizar la información. ¿Qué opina acerca de esto?

Por supuesto que creo que el análisis de las palabras no es suficiente para completar un estudio, pero considero que es una parte interesante. Es más: ése es el objetivo del análisis de contenido que además proporciona nuevos elementos para el análisis. En este sentido, es mejor analizar las palabras espontáneas que la codificación.



Pero *text mining* no es un conteo de palabras; esta técnica nos permite analizar, además, el perfil de las personas que utilizan determinado vocabulario. No estudiamos las palabras aisladas, que es la diferencia entre el estudio de las estadísticas de textos. Podemos derivar de las respuestas a las preguntas abiertas algunos criterios que son capaces de discriminar entre los encuestados, y podemos aprender de los individuos para determinar, por ejemplo, si serán buenos pagadores o no. Podemos realizar análisis predictivo, no sólo descriptivo.

Cuéntenos acerca de sus actividades en la Universidad Telecom-Paris Tech y en el CESP.

De hecho, Telecom es una escuela con visión de futuro que apoya a los estudiantes para que actúen como consultores en temas de análisis de datos. Actualmente estoy implicado en la mejora de la medición en ciencias sociales mediante la promoción del uso de preguntas abiertas en las encuestas: descubriendo nuevos contenidos, comportamientos y opiniones. Para lograrlo, desarrollo nuevos procedimientos de análisis automatizados como los procedimientos multidimensionales, donde se analizan grandes conjuntos de preguntas abiertas y cerradas (a través de análisis de correspondencias múltiples, por ejemplo). Más en general, apoyo el desarrollo de *software* integrado para realizar todos estos procedimientos (DtmVic es un ejemplo de dicho *software*).

