



> LENGUAJE

Detección automática de plagios de texto, incluso en traducciones

Expertos del Laboratorio de Ingeniería en Lenguaje Natural (Lab NLE), integrado en el Grupo de Ingeniería del Lenguaje Natural y Reconocimiento de Formas (ELiRF) de la Universidad Politécnica de Valencia, han desarrollado un nuevo método para la detección automática de textos plagiados y, en concreto, de plagio traducido. Este trabajo ha sido publicado recientemente en *Languages Re-*

sources and Evaluation Journal y se enmarca dentro del proyecto Text-Enterprise 2.0, financiado por el Ministerio de Ciencia e Innovación.

«En nuestros días, el problema del plagio, en particular el de texto, se ha incrementado debido al fácil acceso a grandes fuentes de información a través de medios electrónicos. Desafortunadamente, su detección es prácticamente imposible de forma manual. Por ello, es im-

portante desarrollar mecanismos automatizados que permitan realizar la tarea de detección de plagio y así combatir la creciente tentación de plagiar desde la Web», apunta Paolo Rosso, investigador del Lab NLE de la Politécnica de València.

Junto a Alberto Barrón-Cedeño, estudiante de doctorado de la UPV, Paolo Rosso trabaja actualmente en el desarrollo de tecnología y aplicaciones para la detección au-



Los creadores de la UPV de este método de detección. /E.M.

tomática de plagio y concretamente de plagio traducido, por ejemplo cuando la fuente está escrita en inglés y el plagio está en castellano o valenciano.

Los investigadores de la UPV han desarrollado un método que se basa

concretamente en modelos de traducción estadística. Estos modelos «aprenden» cuáles son las potenciales traducciones de un texto de un idioma a otro a partir de grandes colecciones de documentos.

«Con lo aprendido es posible calcular la similitud entre textos escritos en distintos idiomas y, si esta similitud es muy alta, se puede sospechar de la existencia de un caso de plagio traducido», apunta Alberto Barrón-Cedeño.