



Premio por 'enseñar' a traducir a los ordenadores

Mikel Artetxe y Eneko Agirre, informáticos guipuzcoanos, logran que las máquinas detecten patrones comunes en las lenguas

LUIS ALFONSO GÁMEZ

Si Champollion hubiera tenido a mano miles de textos en griego y en escritura jeroglífica y los ordenadores con los que trabajan Mikel Artetxe y Eneko Agirre, no habría necesitado de la piedra de Rosetta para descifrar los jeroglíficos egipcios. Los investigadores guipuzcoanos han conseguido enseñar a las máquinas a traducir sin necesidad de textos bilingües, sin 'rosettas'. Solo haciendo que lean enormes cantidades de textos diferentes en cada idioma deseado. La creación de esos sistemas capaces de aprender solos a traducir hizo ayer a Artetxe y Agirre merecedores de dos de los Premios de Investigación de la Sociedad Científica Informática de España (SCIE) y la Fundación BBVA.

«Los humanos tendemos a hablar de lo mismo, independientemente del idioma. En euskera o en árabe, hablamos de lo que nos interesa a los humanos», indica Eneko Agirre (Eibar, 1968), catedrático de Lenguajes y Sistemas Informáticos en laUPV. El jurado le ha concedido el Premio Nacional de Informática Arimel «por sus contribuciones de excepcional valor en el ámbito del

procesamiento del lenguaje natural». El lenguaje natural es el que usamos en textos y conversaciones. Agirre, director de HiTZ Centro Vasco de Tecnología del Lenguaje, lleva 30 años investigando cómo hacer que las máquinas lo entiendan.

Traducir sin diccionario

Recuerda que a principios de siglo parecía algo lejano que una máquina nos entendiera. «Sin embargo, en los últimos cinco años, con las redes neuronales profundas, la gran capacidad de cómputo y los grandes volúmenes de texto y voz, han cambiado las reglas del juego y se pueden hacer cosas impensables hace veinte años», reconoce, sorprendido todavía de que nos hayamos acostumbrado a hablar con el móvil y con el coche.

«La idea de que los ordenadores podían aprender a traducir sin supervisión la tuvimos hace cinco o seis años», dice Mikel Artetxe (Zarautz, 1992), investigador actualmente en Facebook AI Research. Su tesis doctoral, dirigida por Agirre, «ha revolucionado el campo de la traducción automática, demostrando que es posible realizar traducciones



Mikel Artetxe, en Londres, donde trabaja para Facebook.



Eneko Agirre, en su despacho de la Universidad del País Vasco.

de forma no supervisada, es decir, sin que la máquina disponga de textos bilingües de los que aprender», aseguraba la Sociedad Española para el Procesamiento del Lenguaje Natural en su nominación a los premios en la modalidad de Investigadores Jóvenes Informáticos. Antes, cuenta el galardonado a este periódico desde Londres, a los ordenadores se les daba el equivalente a diccionarios para que aprendieran cómo se dice las palabras en cada idioma.

El enfoque de estos investiga-

dores es radicalmente diferente. «Es como si a una persona le das un montón de libros en un idioma y otro montón de libros distintos en otro, y le dices que tiene que aprender a traducir. La máquina lo consigue no por magia, sino encontrando patrones en los textos», explica Artetxe. «Con grandes masas de texto, la máquina es capaz de detectar que, por ejemplo, mesa va con silla y con ordenador, y aprende unos patrones que son iguales para el euskera, el árabe y el castellano».