



Descifrando a Babel

Eneko Agirre

CATEDRÁTICO DE LENGUAJES Y SISTEMAS INFORMÁTICOS EN LA UPV-EHU



Desde su centro de investigación HiTZ, que dirigen en la UPV-EHU. Foto: Fundación BBVA

“Intentamos que las máquinas adquieran sentido común”

DONOSTIA – “Ahora estamos intentando que las máquinas adquieran sentido común, sobre todo los agentes conversacionales como Siri, los altavoces inteligentes”, sostiene el profesor Eneko Agirre, catedrático de Lenguajes y Sistemas Informáticos de la Universidad del País Vasco (UPV-EHU), en Donostia, distinguido con tres premios de Google relacionados con este área. “Lo ideal es que las máquinas aprendan por sí solas, porque enseñar a las máquinas es muy costoso; tiene que haber una persona dándole constantemente datos para que aprenda. Para que los agentes conversacionales sepan hay que facilitarles las informaciones preparadas por las personas. El reto actual, de hecho está funcionando, es hacer unos motores que entiendan el lenguaje de tal forma que con menos datos

funcionen mejor los sistemas conversacionales”, explica a DEIA.

Reconoce que hace cinco años hubiera sido impensable que Siri nos entendiera y mucho menos que ejecutara nuestras órdenes: “La posibilidad de hablar por las máquinas con voz se va a extender, no solo al altavoz inteligente y al móvil, sino también a otros aparatos que nos rodean en la industria”.

Al profesor Eneko Agirre, galardonado con el premio Arimel, por sus contribuciones de excepcional valor en el ámbito del procesamiento de lenguaje natural, le interesa sobre todo “cómo las personas nos entendemos y qué mecanismos hay que añadir a una máquina para que entienda de qué hablamos”.

¿Cómo se puede lograr que las máquinas entiendan a las personas, y en concreto que nos entiendan en

los muchos idiomas en los que hablamos? Es la pregunta que le sigue fascinando.

Distinguido por la “excepcional calidad de sus publicaciones científicas, Agirre ha pasado a ser un referente internacional en su campo, ejerciendo como profesor visitante en la Universidad de Stanford y colaborando con el *European Lab for Learning and Intelligent Systems*, la sociedad que aglutina a los mejores investigadores europeos en Inteligencia Artificial.

En el centro de investigación HiTZ que dirige en la UPV-EHU, Agirre supervisó la tesis de Mikel Artetxe –premiado en la categoría de Jóvenes Investigadores–, en la que se ha comprobado por primera vez que una máquina puede aprender a traducir sin basarse en ningún recurso bilingüe. “Hasta ahora, se creía que

La V edición de los de los premios de la SCIE y la Fundación BBVA han distinguido a dos informáticos vascos: Eneko Agirre y Mikel Artetxe

Reportaje de Nekane Lauzurika

LOS Premios de Informática reconocen a investigadores que han impulsado avances en visión artificial, realidad aumentada, traducción automática y ciberseguridad. Los galardones comprenden dos modalidades: Investigadores Jóvenes Informáticos, que reconoce a los autores de las tesis docto-

“La Inteligencia Artificial propiciará más puestos de valor añadido y los trabajos obsoletos los harán las máquinas”

“Se trata de que las máquinas sean capaces de entender cualquier idioma humano”

ENEKO AGIRRE
Dtor del Centro de investigación HiTZ

para traducir entre dos idiomas hacia falta una especie de piedra Rosetta con traslaciones entre ambas lenguas”, explica. “Pero con nuestro trabajo hemos demostrado que, por ejemplo, para traducir entre el árabe y el chino, basta que la máquina lea muchos textos en estos dos idiomas”, añade. Gracias a este avance, Agirre subraya que se ha ampliado “el campo de los traductores automáticos a idiomas para los que hay pocos recursos, como el euskera, el quechúa o el swahili”.

Su grupo también está investigando cómo mejorar la capacidad de las máquinas para comprender el lenguaje informal. “La posibilidad de conversar de manera natural con una máquina empezará a ocurrir poco a poco. Ahora con el móvil o un altavoz inteligente hay tareas muy sencillas que podemos hacer, como poner música o llamar a alguien y la máquina te entiende perfectamente. Pero el campo de comandos que pueda entender va a ir aumentando cada vez más”.

Además, Agirre también está liderando el desarrollo de sistemas para facilitar cualquier tarea que implique el procesamiento de un gran volumen de documentos. “Esto va a permitir”, asegura, “que en vez de tener que buscar documentos y luego leerlos, podamos preguntar directamente a la máquina lo que queremos y será capaz de encontrar la respuesta que buscamos”.

Reconoce que es el momento de la informática y la inteligencia artificial. “En nuestra área se está produciendo un auténtico boom, porque cada vez más las empresas y la administración van hacia la digitalización de sus datos para automatizar los procesos; con la IA se podrán hacer de forma más rápida y mejor”. ¿Se generarán más puestos de trabajo? “La idea es que haya más puestos de trabajo de más valor añadido y que las tareas repetitivas que se han quedado obsoletas la hagan las máquinas”. - N. Lauzurika

Mikel Artetxe

INVESTIGADOR CIENTÍFICO EN FACEE



El joven informático se encuentra ahora t

BILBAO – Hasta hace bien poco la inmensa mayoría de los sistemas de inteligencia artificial se desarrollaban para el inglés u otros pocos idiomas de muchos recursos, pero los sistemas multilingües que funcionan para un gran número de idiomas, van cogiendo más y más fuerza. “Esto es fundamental para que idiomas de menos recursos como el euskera no se queden atrás en este mundo. Y para ello necesitamos sistemas que aprendan sin una supervisión fuerte o grandes cantidades de textos paralelos, pues estos son escasos o incluso inexistentes para idiomas de pocos recursos. Nuestro trabajo se sitúa en este contexto, centrándose sobre todo en el tema de la traducción”, explica a DEIA, Mikel Artetxe desde Londres, donde se encuentra investigando.

Premiado en la modalidad de joven investigador informático, menor de 30 años, Artetxe que estudió ingeniería informática y se doctoró (2020) en la Universidad del País Vasco, explica que los traductores automáticos convencionales son sistemas supervisados. Esto significa

rales más innovadoras, y los Premios Nacionales de Informática, que distinguen a veteranos investigadores o entidades públicas y privadas que han dedicado su labor al estudio y fortalecimiento de la informática

Entre los premiados en esta edición se hallan Eneko Agirre y Mikel Artetxe. Estos investigadores de la UPV-EHU, maestro y discípulo, se han convertido en referentes mundiales en Inteligencia Artificial aplicada a la

traducción automática. Mikel Artetxe, bajo la supervisión de su profesor Eneko Aguirre, han demostrado por vez primera que es posible que una máquina "aprenda" sola a traducir dos idiomas sin haber sido expuesta a traducciones hechas por humanos. Basta que la máquina aprenda a leer muchos textos en esos idiomas, por ejemplo chino y árabe, para ser capaz de aprender a traducir entre ellos.

Agirre pone en valor el premio, ya

que muestra la importancia de la investigación básica en esta área, porque gran parte del valor añadido "viene de esa investigación y de su transferencia a la industria. De hecho en Europa y en el Estado se han dado cuenta de la importancia de las tecnologías del lenguaje y que se están quedando atrás con respecto a EE.UU. y China. En el Estado se invertirán este año cien millones más", reconoce Agirre. ●

OOK AL RESEARCH, EN LONDRES



abajando en Londres, aunque le gustaría volver a Euskadi. Foto: Fundación BBVA

"Estoy trabajando en Londres y en un futuro me gustaría volver a Euskadi, pero la situación allí es bastante precaria"

"Un traductor automático en el móvil nos permite ya hoy comunicarnos en cualquier idioma"

MIKEL ARTETXE
Premio a investigadores jóvenes

"Si automatizamos tareas, viviremos mejor trabajando menos horas"

que parten de millones de traducciones humanas ya existentes, de los que identifican patrones y aprenden a generalizar. ¿Pero qué ocurre si no tenemos ninguna traducción existente y partimos únicamente de textos independientes en dos idiomas?

"Esto es lo que hace un traductor no-supervisado, que es el paradigma que desarrollamos en mi tesis doctoral. Es como si a una persona le diésemos un montón de libros en un idioma y otro montón de libros distintos en otro idioma y le dijésemos que tiene que aprender a traducir. La máquina lo consigue no por magia, sino hallando patrones en los textos", explica sonriente.

En su opinión, el aprendizaje no supervisado es un área cada vez más

importante en la Inteligencia Artificial. "En gran medida, los seres humanos aprendemos observando nuestro entorno, sin necesidad de que nos den millones de ejemplos explícitos de cómo se hace cada una de las cosas que aprendemos a hacer. Y, de manera similar, tampoco es viable obtener millones de ejemplos para enseñarle cada tarea a una máquina. Es por eso que el aprendizaje no supervisado es cada vez mas importante", recalca.

Autor de una tesis doctoral que "ha revolucionado el campo de la traducción automática demostrando que es posible realizar traducciones de forma no supervisada, es decir, sin que la máquina disponga de textos bilingües de los que aprender",

destaca la Sociedad Española para el Procesamiento del Lenguaje Natural en su nominación. "Esto se consideraba un hito difícilmente franqueable", añaden desde la Sociedad.

El gran impacto de este trabajo en la comunidad internacional y una trayectoria laboral como investigador en DeepMind y Google Research, explican que Artetxe reciba el premio "por su alta internacionalización, reconocida a través de sus estancias en las empresas más importantes dentro del campo del procesamiento del lenguaje natural".

Piensa que todos estos avances beneficiarán a la sociedad. "De hecho no es una cosa futurista, es algo que ya está entre nosotros en gran medida. Por ejemplo, hoy en día podemos viajar a cualquier parte del mundo y comunicarnos en el idioma local con la ayuda de un traductor automático que llevamos en el teléfono. Puede que no sea perfecto ni mucho menos, pero a mí por lo menos me ha resultado muy útil cuando he viajado al extranjero", subraya.

En traducción automática en general, centrada hasta ahora sobre todo en textos periodísticos, se han logrado traducciones incluso mejores que las humanas. "Esto hay que cogerlo con pinzas en Literatura, porque no creo que el problema esté en absoluto resuelto", añade. "Estamos muy lejos, pero creo es posible, también con textos literarios".

¿Reemplazarán estos programas a los traductores humanos? "En la historia siempre ha habido trabajos que han quedado obsoletos y esto no es malo, así es como avanza la humanidad. Si en el futuro podemos automatizar más trabajos y esto nos permite vivir bien trabajando menos horas, pues sería fantástico. Eso traerá consigo cambios en la sociedad y en la economía, y será trabajo de todos luchar para que esos cambios sean beneficiosos para todo el mundo y no solo para unos pocos privilegiados". - N. Lauzirika