

La influencia de la inteligencia artificial en la escritura

La influencia de la inteligencia artificial en la escritura está creciendo. Correcciones ortográficas, sugerencias gramaticales y de estilo, sugerencias de contenido, hasta máquinas que escriben por sí mismas. ¿Cómo es esto posible? ¿Quién domina todo esto? ¿Cuáles son los riesgos y cómo mitigarlos?

La influencia de la inteligencia artificial (IA) en la escritura está creciendo. Correcciones ortográficas, sugerencias gramaticales y de estilo, sugerencias para completar una frase, hasta máquinas que escriben por sí mismas. Y, en general, más allá de la escritura, la IA tiene cada vez más impacto en el uso de la lengua natural.

Traducimos entre una multitud de distintos idiomas que no dominamos, generamos automáticamente resúmenes de documentos largos, los filtros *antispam* nos mantienen el *inbox* limpio y conversamos por voz con asistentes digitales. Este último uso también permite a personas con alguna discapacidad o a los analfabetos acceder a contenido digital.

¿Cómo es posible que una máquina sea capaz de hacer todas estas tareas que requieren inteligencia como las hacemos las personas? El procesamiento de lenguaje natural (PLN) de hoy se basa en aprendizaje profundo, una rama del aprendizaje automático que se fundamenta en redes neuronales capaces de procesar muchísimos ejemplos y, a partir de ahí, generar patrones estadísticos complejos fuera del alcance de las personas. Simplificando mucho, un modelo de lenguaje consiste en millones de vectores (*word embeddings*) que representan cadenas de palabras y que reflejan la frecuencia de coocurrencia de palabras. Estos modelos se entrenan con cantidades ingentes de frases donde se quita una palabra y el modelo aprende a predecir la palabra eliminada: se llaman modelos de aprendizaje autosupervisado. El ejemplo más conocido que se puede usar a través de la web es GPT-3 de Open AI.

Las herramientas para corregir el idioma escrito se basan en entrenar un modelo de IA con muchísimos documentos

Otros métodos, más antiguos, usados para el PLN se basan en reglas gramaticales. De la misma forma que aprendemos en el colegio, el algoritmo de IA descompone una frase en sus componentes (verbos, sustantivos, adjetivos, adverbios, etcétera) y, a través de esta representación simbólica, puede hacer tareas de lenguaje natural (analizar y resumir documentos, responder a preguntas, etcétera). Estos modelos son más comprensibles para las personas porque entendemos las reglas en las que se basan.

El salto cuántico de los últimos años en la IA aplicada al lenguaje natural es debido al aprendizaje profundo que permite hacer tareas variadas una vez entrenado el modelo. Las tecnologías más tradicionales siguen aportando valor a los negocios, pero requieren mucho más trabajo manual y mantenimiento. No es de extrañar, entonces, que en la actualidad se apueste por los modelos híbridos que usan una combinación de ambas tecnologías.

Más allá de la escritura práctica, la misma tecnología de aprendizaje automático se está aplicando para explorar la escritura de noticias (*robo reporters*¹) y literatura², la composición de música clásica³ o la pintura de cuadros⁴. Además, con estos modelos se podría averiguar —estimar— la autoría de obras clásicas donde se desconoce el autor real como *El Quijote de Avellaneda* o *El Lazarillo de Tormes*.

Volvamos a la escritura del español en general, es decir, cuando las personas escribimos correos electrónicos, artículos, mensajes cortos o *posts* en *blogs* o redes sociales... Quizás no nos damos cuenta, pero en este proceso estamos vigilados por la inteligencia artificial que observa en tiempo real qué estamos escribiendo y detecta y señala errores ortográficos, gramaticales y de estilo, además de proporcionarnos sugerencias para mejorar o completar el texto. Pensemos por ejemplo en los correctores de Microsoft Word, Gmail, Google Docs, los teclados en los móviles de Android y IOS, etcétera. Cada vez más, estas herramientas se basan en redes neuronales y menos en reglas gramaticales. Veamos más en detalle cómo funcionan estas herramientas.

Las empresas e instituciones hispanohablantes deben tomar cuanto antes el liderazgo en el desarrollo y la evolución de la IA en español

Las herramientas para corregir el idioma escrito se basan, en gran medida, en entrenar un modelo de IA con muchísimos documentos, usando, como hemos comentado antes, vectores de *word embeddings* que reflejan el uso real del idioma. Es importante darse cuenta del bucle que implica esta tecnología aplicándola al procesamiento de lenguaje natural: la IA aprende de los documentos que hemos escrito las personas. El modelo así aprendido por la máquina se usa para corregir errores o sugerir mejoras a las mismas personas, que generan a su vez más documentos. Pero existen riesgos asociados a este bucle.

El riesgo del empobrecimiento de la lengua española. Cuando alguien escribe un texto en formato digital, un corrector sugiere correcciones (línea roja) y mejoras (línea azul). En el párrafo anterior explicamos que estos correctores se basan en IA, es decir, han aprendido las palabras correctas y su uso en frases, basado en los documentos de entrenamiento. Entonces, no es ninguna garantía que una palabra subrayada en rojo realmente no exista; quizás simplemente no se está usando con mucha frecuencia y el algoritmo no la ha priorizado. De hecho, una investigación interna de LEIA, realizada por Telefónica en 2019, demostró que alrededor de un 10 por ciento de las palabras no reconocidas por Microsoft Word y Google Docs sí aparecen en el Diccionario de la Lengua Española de la RAE.

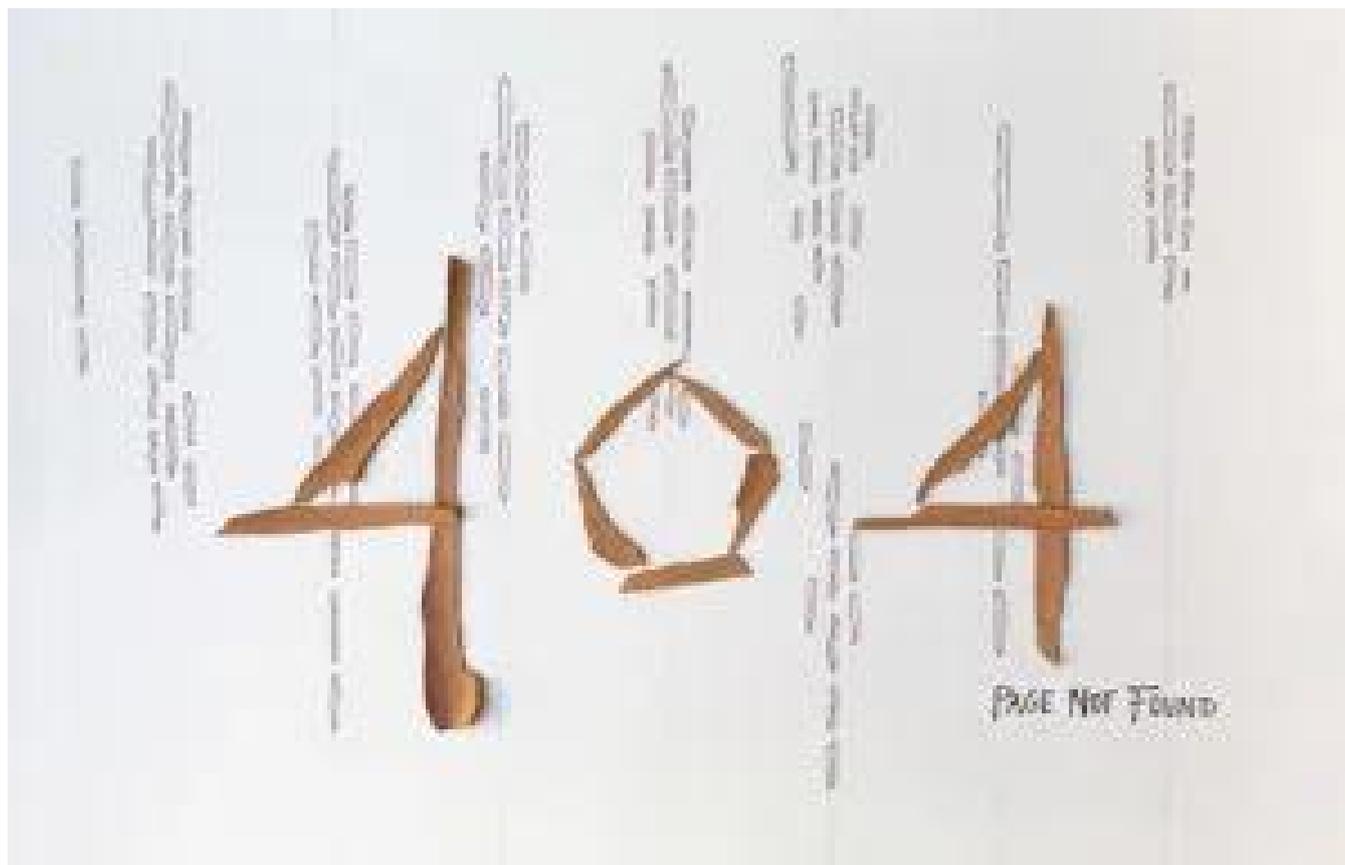
A muchos de nosotros no nos gusta que nuestro texto contenga palabras subrayadas en rojo y las cambiamos por otras palabras. Pero estamos quitando una palabra de nuestra elección por otra sugerida por un algoritmo. Y nuestro texto será usado después para entrenar otro algoritmo, que tampoco tendrá la oportunidad de

aprender esta palabra. Y así sigue el bucle con millones y millones de personas que escribimos millones de documentos. La consecuencia es que, poco a poco, las palabras menos frecuentes dejan de usarse y podrían caer en desuso, no porque la gente no las quiera usar, sino por una razón algorítmica.

El riesgo de la viralización de errores. El mismo bucle también genera otro impacto negativo: cada vez usamos más un lenguaje incorrecto para expresarnos. La generación más joven usa multitud de abreviaciones y combinaciones de letras fonéticas (hasta incluso iconos) para poder escribir más rápido en teclados pequeños (xq, tb, vd, esk⁵). Además, entran muchos anglicismos en la escritura. Como hemos visto, la IA aprende del texto que escribimos las personas. Si usamos muchas palabras incorrectas, el modelo acabará aprendiendo estos errores y puede sugerirnoslos. Por el uso masivo de estas herramientas, los errores se pueden hacer virales.

El riesgo de perder la unidad del español en el mundo digital. Es un hecho que, en la actualidad, el español en el mundo digital está dominado por algunas empresas estadounidenses que todos conocemos. Lo que la RAE ha conseguido en los últimos 300 años, con la colaboración de la Asociación de Academias de la Lengua Española (ASALE) —la unidad del idioma español respetando sus variedades locales— está actualmente en riesgo en el mundo digital.

No hay criterios lingüísticos unificados que usen Alexa, Siri o Cortana cuando hablan el español, ni los hay cuando los correctores sugieren correcciones o mejoras. Cada uno se basa en sus propios criterios, con las mejores intenciones, pero sin garantías para la unidad del español, y siempre al servicio de sus respectivos modelos de negocio.



El proyecto LEIA

El proyecto LEIA⁶ (Lengua Española e Inteligencia Artificial), liderado por la RAE, en colaboración con ASALE y con el apoyo de coordinación de Telefónica, intenta mitigar estos riesgos. Fue fundado en 2019 incluyendo también como socios a Microsoft, Google, Amazon, Twitter y Meta (anteriormente Facebook) con el objetivo de conseguir que las máquinas hablen un español correcto y que las herramientas de apoyo a la escritura, a su vez, respeten un español correcto. Así, se fomenta la unidad del español en el mundo digital. Y para garantizarlo, en el futuro, habrá un sello que fomente el uso correcto del español en el mundo digital, expedido por la RAE.

Algunos ejemplos de lo que ya podemos disfrutar, o lo que podemos esperar como resultado del proyecto LEIA, incluyen:

Telefónica. A través de un proceso formal, todas las frases que genera Aura, el asistente digital de Telefónica, cumplen con los criterios de la RAE según sus recursos lingüísticos. Además, a través del mando de Movistar+, es posible interactuar por voz con el diccionario de la RAE (DLE) preguntando por el significado de palabras españolas.

Microsoft. Consultando por el significado de una palabra en español en Bing, el buscador de Microsoft, la respuesta viene directamente del diccionario de la RAE. Microsoft también está trabajando para que todas sus herramientas editoriales (corrector en Word y otros productos) se basen en los recursos lingüísticos de la RAE, como el diccionario o el lexicon.

Google. El corrector del teclado Gboard de Google que se usa en todos los móviles Android hasta ahora solo usaba modelos de IA (los *word embeddings*) para sus sugerencias de correcciones o mejoras. Gracias a los recursos lingüísticos de la RAE, se añaden 200.000 palabras españolas al corrector, mejorando así la escritura de los usuarios de Android. Es un claro ejemplo de una aproximación híbrida de PLN. También esperamos ver en el futuro que el buscador de Google consulte el diccionario de la RAE en vez de un diccionario no oficial. De hecho, el diccionario que usa Google actualmente tiene unas 20.000 entradas menos que el DLE, además de tener menos acepciones y una peor cobertura geográfica (las variedades del español según geografía⁷).

Amazon. Cuando preguntas a Alexa por el significado de una palabra en español, responde con: “Según la Real Academia Española, la palabra significa...”. Asimismo, Amazon está trabajando en que todo lo que dice Alexa y que está bajo el control de Amazon, respete perfectamente el español oficial de la RAE. Amazon también hizo un estudio de los extranjerismos, errores más frecuentes y riqueza léxica según geografía⁸. Esta información es relevante para la investigación de la lengua española en todas sus geografías.

Twitter. Gracias a que Twitter se usa en todas las geografías donde se habla el español, está muy bien posicionado para investigar la evolución del español en términos de neologismos y anglicismos y cómo estos se van integrando en el uso diario del español en cada geografía y cuáles podrían ser las razones sociales que dan lugar a este nuevo uso.

Meta. Con la apuesta de Meta por el metaverso, no es de extrañar que en el futuro seguramente se podrá encontrar la RAE con su diccionario en el metaverso con un avatar, listo para resolver cualquier duda sobre las palabras en español.

Con todas estas iniciativas en el contexto de LEIA, es posible monitorizar la evolución del español a nivel mundial y mitigar los riesgos identificados anteriormente.

Aunque el proyecto LEIA es un paso muy importante para garantizar el uso correcto del español en el mundo digital, las empresas estadounidenses tratan el español, en el mejor de los casos, como segunda lengua, siendo el inglés siempre el primer foco.

Por eso es muy importante que, cuanto antes, las empresas e instituciones hispanohablantes tomen el liderazgo en el desarrollo y la evolución de la inteligencia artificial en español.

Bamler, R. y Mandt, S. (2017): *“Dynamic Word Embeddings” en 34th International Conference on Machine Learning* (380-389).

Costa-Jussà, M. R. et al. (2012): *“Study and comparison of rule-based and statistical catalan-spanish machine translation systems” en Computing and Informatics* (31 (2): 245

Tenney, I. et al. (2019): *“BERT Rediscovered the Classical NLP Pipeline” en Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (2019, 4593-4601).