



RICARDO BAEZA-YATES :

“Como sociedad, deberíamos decidir qué cosas queremos que sean 100% humanas, y cuáles no”

El investigador, reconocido internacionalmente como experto en inteligencia artificial, destaca que hay muchas tareas que aumentan la productividad, pero no se relacionan con lo cognitivo, y enfatiza la importancia de estudiar el impacto social y ambiental de cualquier aplicación de esta tecnología.

ANA MARÍA PEREIRA B.

“**M**i utopía sería que la inteligencia artificial (IA) hiciera todo lo que no queremos hacer como especie: un montón de cosas físicas, aburridas, etc., y que nosotros hagamos lo que nos gusta hacer: más gente que escriba, que pinte, que componga... Yo dejaría el campo creativo para la humanidad. Que todas las personas hagan lo que siempre han querido y no pueden porque están trabajando para subsistir. Y tener un salario mínimo global generado por la IA universal. Ese, para mí, es el verdadero renacimiento”.

“Sueno bonito, pero no vamos por ese camino”, advierte sobre su propio sueño Ricardo Baeza-Yates, exalumno del DCC, quien en marzo pasado recibió el Premio Nacional de Ciencias Aplicadas y Tecnológicas 2024, convirtiéndose en la primera persona ligada al área que recibe este galardón.

Baeza-Yates comenzó su carrera realizando estudios simultáneos de Bachiller en Computación e Ingeniería Civil Eléctrica en la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile (FCFM). Luego, alcanzó el grado de magister en Ciencias de la Computación e Ingeniería Eléctrica, y más tarde obtuvo el PhD en Ciencia de la Computación por la Universidad de Waterloo (Canadá). Volvió a Chile para continuar su carrera en el DCC, donde ha sido académico, coordinador del Magister en Ciencias de la Computación y director del Departamento dos veces.

Radicado en el extranjero desde hace 21 años, hoy vive en Palo Alto, California, desde donde realiza trabajos de investigación *part time* para el Instituto Real de Tecnología de Suecia, la Universidad Pompeu Fabra de España y el DCC de la Universidad de Chile, labor

que ocupa cerca del 40% de su tiempo. El resto lo dedica a dar charlas y consultorías a empresas y *startups* sobre IA responsable, materia en que se ha vuelto reconocido.

LEGITIMIDAD INTEGRAL

En 2022, participó en el Consejo de Políticas Tecnológicas de la Association for Computing Machinery (ACM), la sociedad de computación más grande del mundo. Fue uno de los dos autores principales de la declaración global sobre los nuevos Principios para sistemas algorítmicos responsables (definidos así para incluir a cualquier sistema que pueda tener problemas éticos, no solo la IA).

Son nueve principios, en orden de importancia (ver recuadro), y el primero, “legitimidad y competencia”, lo propuso él. “Legitimidad significa que sea legítimo socialmente, que entrete que más beneficios que daños a la sociedad; científicamente, que no sea usado como una pseudociencia, por ejemplo, y, por supuesto, legal y ética mente”, explica.

“La competencia se refiere a tener las competencias administrativa (los permisos respectivos); técnica (los conocimientos necesarios), y de dominio del problema; muchas veces, los informáticos están contestando preguntas que no son de su ámbito de experiencia”, agrega.

Aplicar este principio significa hacer un análisis de riesgo e impacto social y ambiental, que recomienda realizar a través del ejercicio de *red teaming*, consistente en reunir a un grupo de “abogados del diablo”, de todos los sectores involucrados, para que busquen posibles problemas de un proyecto. Obviamente, es difícil prever todos los ángulos en el vertiginoso contexto actual, pero las miradas externas pueden reducir lo que no se ve.

El Premio Nacional enfatiza que la falta de este análisis ya ha producido problemas, y le preocupan, sobre todo, los suicidios adolescentes producto de interacciones con la IA; ya se cuentan siete, cinco el último año.

“¿Por qué OpenAI no puso control parental en ChatGPT antes? Creo que hay desidia en pensar las cosas más obvias”, dice. Pero no se trata solo del control parental. “Hay que ser drásticos: si empieza una interacción que pueda llegar a algo dañino, tóxico, etc., hay que terminarla. Eso debería ser como una ley”, afirma.

SABER DISTINGUIR

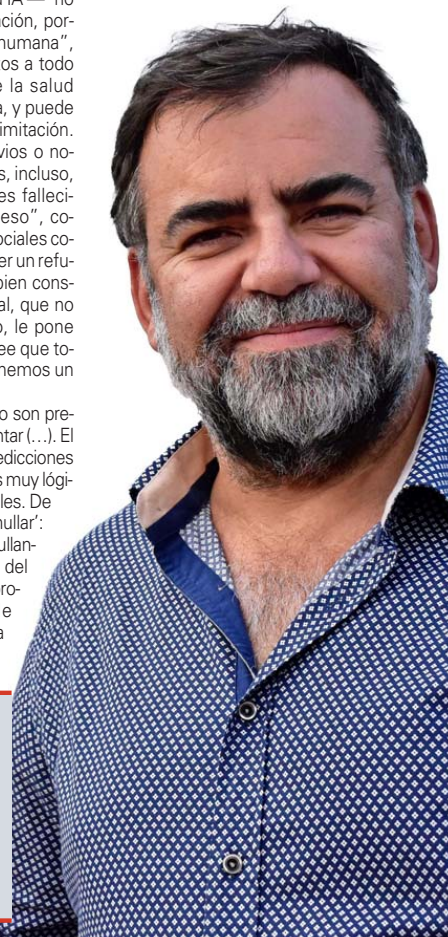
Porque la interacción con la IA —“no podemos hablar de conversación, porque no es una comunicación humana”, asegura— puede tener efectos a todo nivel, especialmente sobre la salud mental. La IA imita la empatía, y puede haber quienes no distinguen la imitación. “Hay gente que ya tiene novios o novias digitales, amigos digitales, incluso, conversaciones con parientes fallecidos, y vamos a ver más de eso”, comenta, debido a problemas sociales como la soledad. “La IA puede ser un refugio, pero si la gente no está bien consciente de que todo es artificial, que no hay una persona al otro lado, le pone nombre, le da las gracias y cree que todo lo que genera es cierto, tenemos un problema mayor”, subraya.

Recuerda que en la IA “todo son predicciones y puede llegar a inventar (...). El problema es que como las predicciones son lo más probable, son cosas muy lógicas y bien escritas, suenan reales. De ahí yo creé la palabra ‘chat-mullar’: por diseño, la IA está ‘chat-mullando’ siempre. La mayor parte del tiempo, eso es cierto, pero el problema ocurre cuando no le ‘achunta’ y la gente no se da cuenta. Aunque el 99% del

tiempo la IA esté correcta, el problema va a ocurrir con el 1% que no lo esté y que uno no verifica o muchas veces es muy difícil de verificar”.

Pero la IA no es negativa *per se*, advierte. Es neutra. “Hay muchas tareas que uno puede hacer con IA para aumentar la productividad, que ahorran mucho tiempo, pero no tienen que ver con la parte cognitiva”, manifiesta.

“Como sociedad, deberíamos decidir qué cosas queremos que sean 100% humanas, y cuáles no (...). Yo establecería que la IA no se pueda usar para hacer ninguna aplicación que reemplace a las personas, solo aquellas que las mejoren”, propone.



LOS NUEVE PRINCIPIOS BÁSICOS DE RESPONSABILIDAD

- 1) Legitimidad y competencia.
- 2) Minimización del daño.
- 3) Seguridad y privacidad.
- 4) Transparencia.
- 5) Interpretabilidad y explicabilidad.
- 6) Mantenibilidad.
- 7) Auditabilidad e impugnación.
- 8) Rendición de cuentas y responsabilidad.
- 9) Limitación de los impactos medioambientales.

Fuente: ACM Statement principles for responsible algorithmic systems, 2022.