LITORALPRESS

Fecha: 03-06-2023
Medio: El Mercurio
Supl. : El Mercurio - Cuerpo A
Tipo: Noticia general
Título: ¿Es la inteligencia artificial una amenaza para la humanidad?

Pág. : 2
Cm2: 396,5
VPE: $ 5.208.750

Tiraje: 126.654
Lectoría: 320.543
Favorabilidad: Positiva

¿Es la inteligencia artificial una amenaza para la humanidad?

"...aplicaciones que buscan establecer lazos emocionales entre el usuario y la inteligencia artificial pueden conllevar riesgos imprevistos para personas psicológicamente vulnerables, como lo ejemplifica el caso del suicidio aparentemente inducido en un usuario por el Chatbot 'Eliza'...".

ABEL WAJNERMAN PAZ
Instituto de Éticas Aplicadas UC

PAULINA RAMOS VERGARA
Centro de Bioética UC

Estamos siendo testigos a diario de diferentes maneras en las que la inteligencia artificial (IA) aporta a los objetivos del desarrollo sostenible. El foco de atención está actualmente en las oportunidades que proporcionan las IAs que interpretan y producen lenguaje humano.

Si bien la discusión reciente se ha centrado en aplicaciones que conciernen a la educación e investigación, queremos detenernos en aplicaciones orientadas a la salud mental y el bienestar psicológico.

Se trata de *apps* que buscan establecer lazos emocionales entre el usuario y la IA. Esto incluye unas que ofrecen terapeutas artificiales, como Wysa o Woebot Health, y otras como Replika o Anima, donde la IA encarna un vínculo personal del usuario (una amiga, un hermano o una pareja romántica). Para cumplir con su función, estas IAs necesitan exhibir rasgos de tipo humano en sus interacciones, imitando nuestras reacciones emocionales, lo que les da a los usuarios la oportunidad de explorar y profundizar sus propias emociones.

Pero al mismo tiempo estas tecnologías pueden conllevar riesgos imprevistos para personas psicológicamente vulnerables, como lo ejemplifica el caso del suicidio aparentemente inducido en un usuario por el Chatbot "Eliza", basado en la tecnología GPT-J.

Por este motivo debemos llegar a un acuerdo sobre cómo restringir el uso de estas aplicaciones, lo que requiere tiempo para una discusión sustantiva que involucre a todos los actores relevantes.

Así, consideramos necesaria una moratoria o suspensión de estas *apps* hasta que se elaboren regulaciones efectivas. Este tipo de estrategias tiene precedente en el desarrollo de la investigación sobre el ADN recombinante. En

1974, biólogos moleculares se reunieron en California y acordaron una moratoria sobre muchos tipos de experimentos potencialmente riesgosos, lo que resultó en la redacción de una propuesta de normativa de seguridad para regular la ingeniería genética, que se convirtió en un modelo influyente en políticas sobre tecnologías de riesgo.

Ahora bien, la aplicación de un principio de precaución (esto es, pausa y revisión antes de avanzar con innovaciones que pueden potencialmente causar daños desastrosos) para este caso ha recibido críticas. Se ha argumentado que una moratoria podría conllevar una pérdida de control y transparencia al empujar el desarrollo de la IA a la clandestinidad, que podría introducir desigualdades si ciertos países no aceptan la moratoria, que no dura lo suficiente como para desarrollar una propuesta legislativa consistente, y que podría ralentizar el desarrollo de aplicaciones que beneficiarían a millones de personas.

Las primeras tres razones apuntan a que la medida es implausible por las limitaciones de las instituciones que deberían implementarla. Si bien somos conscientes de los retos que supone avanzar con una propuesta de este tipo, las consideraciones sobre qué es más factible hacer no responden a la pregunta de fondo, que es qué deberíamos hacer. Aún si constituye un verdadero desafío político, una moratoria es una medida éticamente necesaria cuando los potenciales beneficios de la aplicación desregulada de estas tecnologías podrían darse a expensas de la vulneración de principios no negociables, como la integridad psíquica y física de las personas.

Respecto de la última razón, coincidimos en que las medidas que detienen la investigación en estrategias terapéuticas producen un enorme daño a personas con patologías que requieren atención urgente. Creemos así que la moratoria solo debe afectar la aplicación de estas tecnologías y no la investigación sobre sus efectos en la salud mental, que son los que necesitamos entender mejor antes de que se sigan esparciendo en la sociedad. Por lo mismo, la moratoria podría estar restringida a IAs conversacionales con funciones sociales.

El llamado a una moratoria pretende interpelar a la comunidad científica a asumir su responsabilidad formalizando códigos de conductas (como lo ha hecho en el caso de la manipulación genética y, más recientemente, respecto de la neurotecnología), y a los legisladores a abordar los vacíos normativos sobre los efectos psicológicos de las IAs. Principios neuroéticos, como los de la integridad psicológica y la libertad cognitiva, podrían ser más eficientes para enfrentar estas problemáticas que los principios que se articulan en las recomendaciones vinculadas a la ética de la IA (como los de responsabilidad, transparencia y la explicabilidad).

Con este llamado buscamos así contribuir a informar al público e invitar a la ciudadanía a tomar posición sobre los peligros que representan para la humanidad las tecnologías sin regulaciones.