

IASEAI hace un llamamiento a la acción para legisladores, académicos y la sociedad antes de la Cumbre de IA en París

La Asociación Internacional para una Inteligencia Artificial Segura y Ética (IASEAI) ha concluido hoy su conferencia inaugural con un llamamiento a la acción con diez medidas cruciales.

7 de febrero, PARÍS: La conferencia [IASEAI '25](#) ha sido la reunión global más grande hasta ahora sobre los últimos avances en seguridad y ética de la inteligencia artificial. Celebrada en la sede de la OCDE justo antes de la Cumbre de Acción sobre IA en París, la conferencia ha reunido a destacados representantes del mundo académico, la sociedad civil, la industria, los medios de comunicación y los gobiernos

El profesor **Stuart Russell**, presidente de la IASEAI y distinguido profesor de Informática en la Universidad de California, Berkeley, ha declarado:

"El desarrollo de una IA altamente capaz probablemente será el evento más importante en la historia de la humanidad. El mundo debe actuar decididamente para asegurar que no sea el último evento en la historia de la humanidad. Esta conferencia, y el espíritu de cooperación de la serie de cumbres sobre IA, me dan esperanza; pero debemos transformar la esperanza en acción, pronto, si queremos que haya un futuro en el que deseemos que vivan nuestros hijos."

Las siguientes acciones, derivadas de la investigación y los debates mantenidos durante la conferencia, ayudarán a avanzar en la misión de la IASEAI de asegurar que los sistemas de IA operen de manera segura y ética.

IASEAI pide:

- 1. Reconocimiento de la importancia de los nuevos desarrollos en IA.** Los responsables políticos deben actuar con una urgencia proporcional al potencial transformador de la IA, a la velocidad de su avance y a los riesgos crecientes que conlleva para la humanidad, especialmente a medida que sus capacidades comienzan a superar las humanas.
- 2. Evitar la disrupción institucional y social provocada por la IA.** El avance de la IA amenaza con desestabilizar el empleo, las estructuras sociales, agravar la desigualdad y comprometer gravemente el ecosistema informativo. Los responsables políticos deben tomar medidas proactivas para proteger las instituciones, las personas y los modos de vida, garantizando que la IA fortalezca la sociedad en lugar de debilitarla. Dado que estos problemas tienen formas similares en diferentes países, es necesario buscar soluciones colectivas.
- 3. Abordar la carrera hacia la IA General (AGI).** Los incentivos financieros y estratégicos para alcanzar una inteligencia artificial general (AGI) o "superinteligencia" llevan a empresas y países a reducir los estándares de seguridad para obtener una ventaja tecnológica y política. Los responsables políticos deben tomar acciones coordinadas en foros internacionales, como la ONU y la OCDE, para garantizar que la innovación y la competencia se desarrollen dentro de un marco de riguroso de seguridad.

4. **Unir los esfuerzos de las comunidades de investigación en torno al objetivo de una IA segura y ética.** Las amenazas crecientes para la prosperidad humana que plantean los sistemas de IA requieren que investigadores en IA, ética, ciencias sociales y política colaboren y unan sus esfuerzos. Las diversas perspectivas que ofrecen estas comunidades no están en tensión; al contrario, todas contribuyen de manera importante a garantizar que los sistemas de IA no perjudiquen a la sociedad humana.
5. **Adopción de requisitos de seguridad y ética obligatorios.** Aunque son loables, los compromisos voluntarios de las empresas deben hacerse más específicos y jurídicamente vinculantes. Estos compromisos podrían incluir el registro obligatorio de los sistemas de IA avanzados, incluyendo el autorregistro automático de las copias de código abierto; la instalación de interruptores de desactivación operables remotamente; la notificación obligatoria de incidentes; la formación en ética profesional para ingenieros; y los estándares para el entrenamiento, diseño, desarrollo, pruebas, auditoría y transparencia de los sistemas de IA avanzados. Los desarrolladores deben proporcionar un caso de seguridad científicamente convincente de que sus sistemas no cruzarán las llamadas líneas rojas de comportamiento, que delimitan conductas inaceptables.
6. **Fomento de la cooperación global.** Los desarrollos en IA y los riesgos asociados tienen un impacto transfronterizo, lo que requiere una colaboración internacional para la investigación y la regulación de la seguridad de la IA. Además, los beneficios de la IA deben distribuirse equitativamente. En este sentido, damos la bienvenida a la creación del Consejo Asesor de la ONU sobre IA y la red internacional de Institutos de Seguridad de la IA.
7. **Aumento de la financiación pública en investigación.** La escala del desafío requiere una mayor inversión pública en investigación sobre seguridad y ética de la IA. La investigación no puede continuar siendo dominada por empresas con conflictos de interés. Los beneficios potencialmente transformadores de la IA solo pueden materializarse si los avances en capacidades van acompañados de métodos para garantizar que los sistemas de IA sean seguros por diseño y alineados con los intereses humanos.
8. **Apoyo a la Fundación para la IA.** La futura Fundación para la IA, prevista por los anfitriones de la Cumbre de París, es un paso importante para garantizar que la IA beneficie a la humanidad y fortalezca las capacidades locales. Los gobiernos deben apoyar esta fundación.
9. **Apoyo al Convenio Marco sobre IA del Consejo de Europa.** Este tratado busca garantizar que las actividades en todo el ciclo de vida de los sistemas de IA sean coherentes con los derechos humanos, la democracia y el estado de derecho. Mientras que el Consejo administra el tratado, este está abierto a la firma de estados no miembros. 37 países lo han firmado, incluyendo Estados Unidos y el Reino Unido, y se espera que más países lo firmen en la Cumbre de París.
10. **Promoción de un diálogo informado.** Investigadores y medios de comunicación deben trabajar conjuntamente para informar de manera objetiva y precisa al público y a los responsables políticos sobre los desarrollos en IA y su importancia.

Acerca de IASEAI: La Asociación Internacional para una IA Segura y Ética (IASEAI, por sus siglas en inglés) es una organización sin ánimo de lucro fundada para abordar los riesgos y las oportunidades asociados con los rápidos avances en IA. Su misión es garantizar que los sistemas de IA operen de manera segura y ética, y también para dar forma a las políticas, promover la investigación, y construir entendimiento y comunidad en torno a este objetivo.