

¿Es la IA un reflejo de los prejuicios humanos y una amenaza para la privacidad, la creatividad y los procesos creativos?

ASIPI Panamá

2 diciembre 2024

Manuel DESANTES REAL

Universidad de Alicante

Principales desafíos legales en cuanto al uso de datos personales para el entrenamiento de sistemas de IA

1. **Discriminación y sesgos** (raza, género, religión, etc.).
2. **Anonimización** (ej. definición ambigua de datos anonimizados) y **riesgos de reidentificación** a partir de patrones combinados.
3. **Extracción de datos personales que son públicos en internet** (ej. cuando no son de libre uso para entrenar modelos).
4. **Contenido protegido por derechos de autor o por confidencialidad.** En quinto lugar,
5. **Normativas** (muy diversas en cada Estado) **sobre protección de dato**
 - a. ej. **consentimiento** explícito, informado y específico o
 - b. ej. aplicación del principio de que el uso de datos personales debe **limitarse a la finalidad** para la cual se recogieron originalmente
6. **Transparencia, explicabilidad** y dificultad de auditar y controlar los procesos de datos debido a la complejidad de los sistemas de IA.
7. **Vulneración** a lo largo del entrenamiento de derechos tan obvios como el **derecho de acceso y rectificación** y el **derecho al olvido**.
8. **Relevancia del Derecho internacional privado:** tribunales competentes, ley aplicable
9. **Equilibrio estable** entre el desarrollo de la IA y la protección de los derechos fundamentales en los que hemos cimentado nuestras sociedades democráticas.

¿Es la regulación el instrumento para ayudar a mitigar los sesgos de IA en sectores sensibles como la sanidad, la publicidad o las herramientas policiales predictivas?

- Una regulación inteligente es crucial en estas materias sensibles porque permite establecer estándares éticos, legales, técnicos para garantizar que los sistemas de IA van a ser justos y responsables y van a respetar los derechos fundamentales.
- Ejemplos:
 1. **Evaluaciones de impacto algorítmico** para analizar cómo un sistema de IA podría afectar a diferentes grupos sociales antes de su implementación
 2. **Calidad y representatividad de los datos** para minimizar el riesgo de sesgos por subrepresentación de ciertos grupos
 3. Prohibición de usar **datos sensibles como raza, etnia o religión**
 4. Respeto del respeto a derechos como **no ser discriminado** y a poder **impugnar** decisiones de la IA que sean injustas
 5. Necesidad de crear **organismos de supervisión** en sectores sensibles
 6. Exigir en estos sectores una obligación de **explicabilidad** de las decisiones
 7. Concretar la **responsabilidad del proveedor de servicios de IA**
- Problema: la exponencialmente acelerada evolución de la IA puede superar la capacidad de cualquier normativa y convertirla en papel mojado. De ahí la importancia de la colaboración internacional para evitar que los sesgos algorítmicos perjudiquen a las comunidades más vulnerables.

Desde una perspectiva ética y de buenas prácticas, ¿cuál es el papel ideal de la intervención humana para garantizar que la IA actúe de manera imparcial y transparente?

1. La intervención humana en los sistemas de inteligencia artificial (IA) es **esencial**.
2. El problema reside en acordar un **marco razonable** en el que el control humano no acabe ahogando el desarrollo de la propia IA en aquellos lugares donde se establezca una normativa para ello, como ocurre en Europa.
3. Consecuencias:
 - a. **Explicar** cómo se toman las decisiones
 - b. **Comunicarlas** de forma que se fomente la confianza en la tecnología
 - c. Garantizar la **supervisión humana** en decisiones que afecten a derechos fundamentales
 - d. Priorizar los **equipos multidisciplinares**
 - e. Integrar a personas, organizaciones sociales y comunidades potencialmente afectadas por la IA para entender mejor sus preocupaciones y expectativas
 - f. **Fomentar la educación y la formación permanente**, tanto en lo que respecta a la actualización en tecnologías híbridadas como en las implicaciones éticas y sociales de las decisiones.

Conclusión:

4. Cómo conseguir que los algoritmos trabajen para los “androritmos” y no al revés.
5. Revisión en profundidad de muchas cosas que hemos dado por sentadas, incluyendo todos nuestros sacrosantos postulados sobre la propiedad intelectual.