

Pautas para definir una política editorial de uso de la Inteligencia Artificial



El desarrollo de la Inteligencia Artificial es probablemente el desafío más importante que nos toque vivir como humanidad en este siglo, e implicará transformaciones en todos los órdenes de la vida. Podemos criticarla y tener una mirada escéptica sobre el futuro al que nos va a llevar, y hay argumentos y razones sobradas para ello. Pero su impacto (positivo y negativo) es inevitable, y debemos comprenderla y aprender a convivir con esta tecnología. Los seres humanos hemos evolucionado de la mano de las tecnologías que hemos inventado. Pero la IA es especial, porque nos obliga a cuestionarnos cuál es nuestra esencia y qué nos distingue de nuestro entorno.

Considerando que el año 2025 será un año bisagra para el despliegue de la IA (no falta más que ver los acontecimientos de estos últimos días), resulta clave para cualquier empresa u organización sentar una política que sirva de base para su funcionamiento, y que tenga al mismo tiempo la flexibilidad para ajustarse a medida que los eventos se desarrollen.

Esta política deberá estar centrada siempre en la misión de la empresa. En el caso de una editorial, comprender cuál es el valor sustancial de las publicaciones que realiza y del trabajo humano que desarrolla en torno a los textos. Y con base en esto, definir así en qué medida la Inteligencia Artificial se convertirá en un asistente que realmente optimice las tareas y potencie la creatividad (o bien, establecer sus límites y restricciones).

Por este motivo, hemos decidido publicar esta guía, que revisa y aclara los principales aspectos afectados en la edición y publicación de contenidos. Para que, a partir de este conocimiento, las empresas y organizaciones puedan definir su mejor hoja de ruta. Ya sea, en una política que integre estas tecnologías, o bien que las limite al extremo.

Esperamos que este modesto aporte sea de utilidad para todo tipo de organizaciones que habitan en nuestra industria.

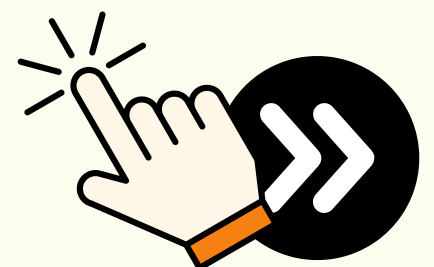
Daniel Benchimol
Director de Proyecto451
danielbenchimol@proyecto451.com



1. ¿Cuál es el riesgo de trabajar una obra literaria con una herramienta de IA?

Lo primero: haya o no riesgo, si no contamos con la autorización expresa del autor del contenido, no debemos emplear herramientas de IA para trabajar con su obra.

Aunque en la práctica, **evitar esto será cada vez más difícil**, ya que los modelos de IA habitan hoy en las aplicaciones que usamos cotidianamente (desde WhatsApp hasta GMAIL, Google Docs, Word o Photoshop).

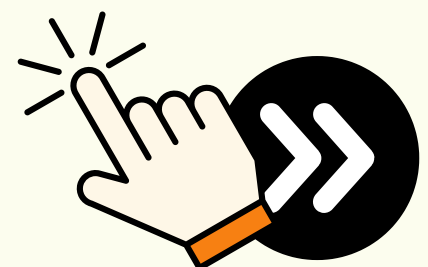


1. ¿Cuál es el riesgo de trabajar una obra literaria con una herramienta de IA?

Lo importante: en la interacción con una herramienta de IA, el modelo puede almacenar y procesar los contenidos que estamos compartiendo con ella.

Pero ojo: la IA **no funciona como una base de datos**. Si, por ejemplo, le compartimos un documento con una obra completa, no habría posibilidad que un usuario en otra parte del mundo le pida acceso a ese mismo documento.

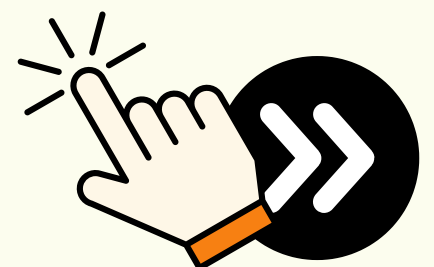
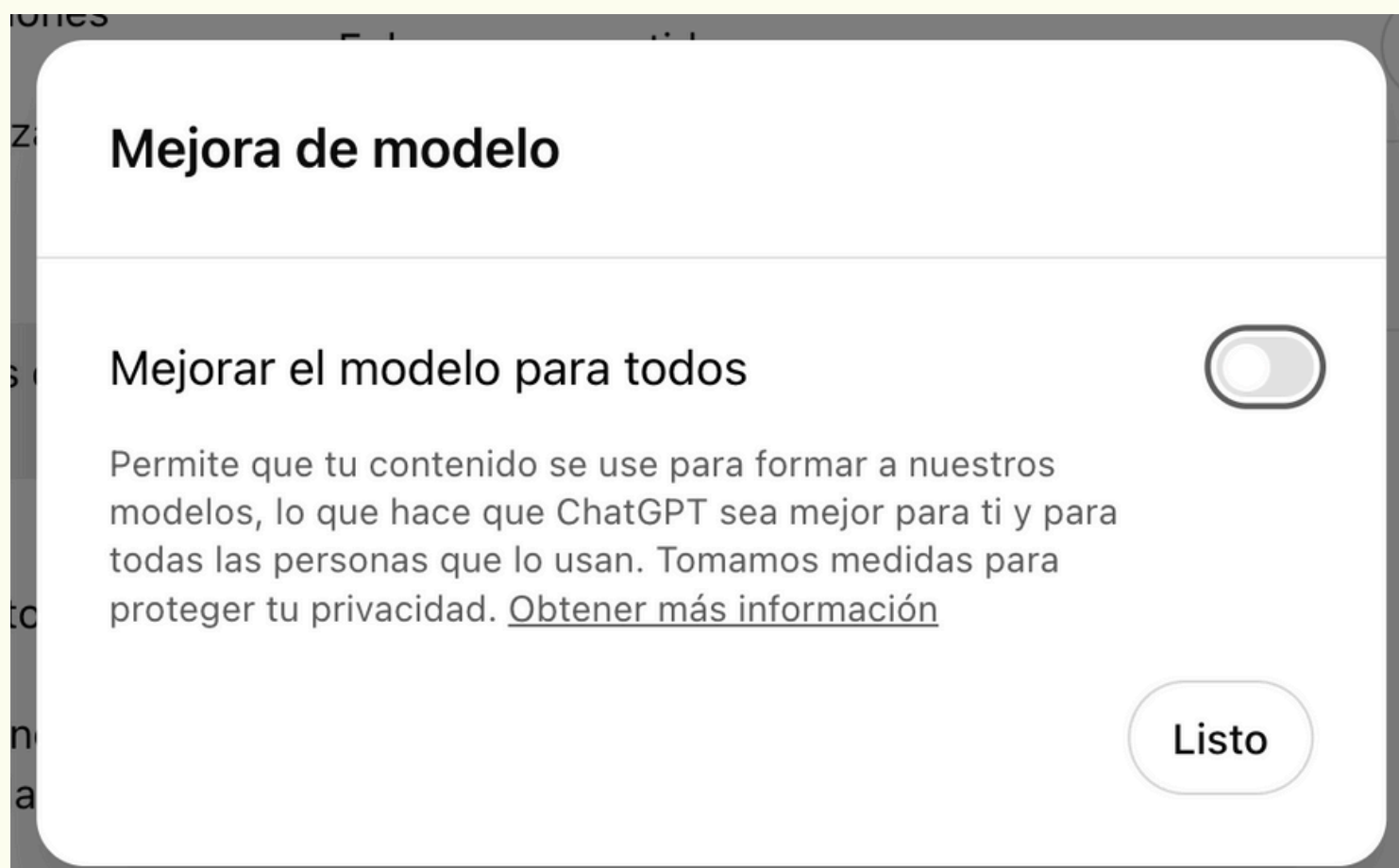
A pesar de que la IA puede entrenarse con nuestros contenidos e interacciones cuando la usamos, el riesgo de que modelos de IA como GPT generen contenido igual o similar al original en otros contextos es **muy bajo o prácticamente nulo** (esto lo aclaramos en la pregunta 5 sobre el plagio).



1. ¿Cuál es el riesgo de trabajar una obra literaria con una herramienta de IA?

El consejo: prácticamente todas las herramientas de IA permiten proteger el contenido que volcamos en ellas.

Las herramientas comerciales ofrecen versiones empresariales que garantizan mayor privacidad, encriptando la información. En ChatGPT y otras similares podremos encontrar opciones para **definir que el modelo no se entrene con nuestros contenidos** o establecer que las conversaciones sean privadas.

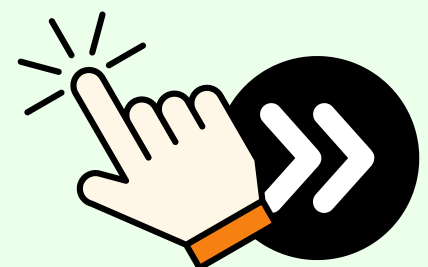


2. ¿Es posible proteger un contenido que ha sido intervenido con IA?

La mayoría de las legislaciones internacionales actuales **reconocen únicamente la autoría humana**. Por tanto, en los casos donde la intervención humana sea mínima o nula, es **casi imposible defender la originalidad de la obra** ante posibles disputas legales.

En la mayoría de los juicios que están ocurriendo en los Estados Unidos, se intenta determinar la **cuantía de trabajo humano**. Por esto, es recomendable documentar todo el proceso de edición/creación para demostrar el grado de intervención humana.

Caso DABUS: El 30 de julio de 2021, la Corte Federal Australiana determinó que un sistema de IA podía ser considerado como inventor en una solicitud de patente. Esta fue una decisión histórica, siendo la primera vez que un juez valoró favorablemente la posibilidad de que una IA fuera reconocida como inventora. Existen algunos otros casos similares en Japón.



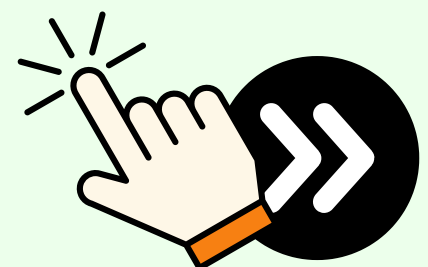
2. ¿Es posible proteger un contenido que ha sido intervenido con IA?

Esto no nos impide utilizar IA para intervenir los contenidos, pero definir cómo y para qué usarla:

- En piezas de comunicación, en la web, en redes sociales: trabajar con textos e imágenes es algo ya muy frecuente en empresas, y tiene todo el sentido ayudarnos con la IA para optimizar tiempos y procesos.
- En nuestros libros, deberemos tener más cuidado, y definir en la política si utilizaremos textos o imágenes parcial o totalmente generados con IA. Y establecer **cuál es la finalidad de utilizar estos modelos** y dónde está nuestro aporte como humanos en la tarea.

Podremos generar una imagen y aplicarla en una portada de un libro que comercialicemos, pero luego no podremos registrar la autoría sobre esa imagen, aunque sea una imagen única y original.

El dato: los contenidos generados con IA se denominan **contenidos sintéticos**.



3. ¿Es posible detectar la intervención de una IA en un contenido?

Detectar si un contenido ha sido generado o modificado por una IA es **realmente muy difícil**.

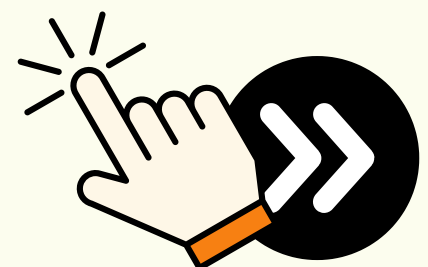
Existen herramientas y técnicas especializadas que analizan patrones textuales, coherencia narrativa y otras características que podrían delatar una intervención automatizada. Pero estas herramientas no son infalibles y podrían generar "**falsos positivos**".

Por caso, estas herramientas suelen señalar que la Constitución fue escrita con IA 😞.

La recomendación: establecer una **política de transparencia interna y hacia colaboradores externos**, que señalen cuándo y cómo han usado IA en sus procesos.

Herramientas de detección de IA:

- Originality
- Copyleaks
- Turnitin
- ZeroGPT

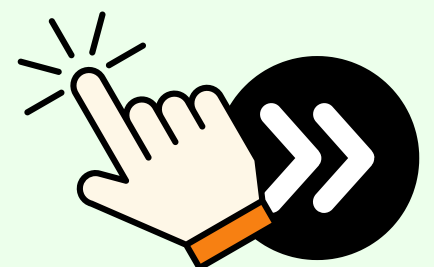


4. Pero... ¿existe alguna forma de trabajar con la IA en forma segura y privada?

¡Sí! A partir de emplear modelos OpenSource, que pueden instalarse en ordenadores personales o en servidores de la empresa.

Estos modelos **no se conectan a Internet, y no se entrenan o alimentan con la información que le compartimos.**

Importante: el modelo OpenSource no es privado o seguro en sí mismo. Todo depende de su implementación. Sí, por ejemplo, empleamos un modelo como DeepSeek en la propia web de la empresa, no será privado ni seguro.



4. Pero... ¿existe alguna forma de trabajar con la IA en forma segura y privada?

Mediante modelos OpenSource podemos trabajar con textos o imágenes que tengan alguna protección de derechos, o con datos privados o sensibles de nuestra empresa. **Nada de lo que compartamos o generemos con el modelo será almacenado** por ninguna otra empresa (cuando lo implementemos en nuestro ordenador).

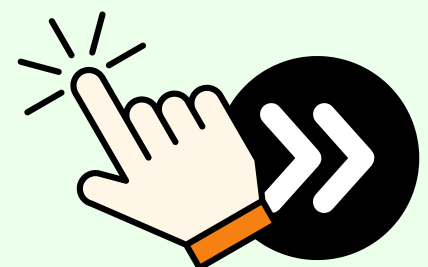
Aun así, **debemos siempre contar con el consentimiento de las personas** que son dueñas de estos datos e información.

El problema: estos modelos pueden exigirnos un poco más de trabajo técnico: debemos instalarlos, configurarlos, etc.

Las ironías de este mundo: Meta (Facebook👍) es quien nos ofrece uno de estos modelos.

Principales modelos Open Source:

- Llama
- Mistral
- DeepSeek
- Qwen
- Gemma



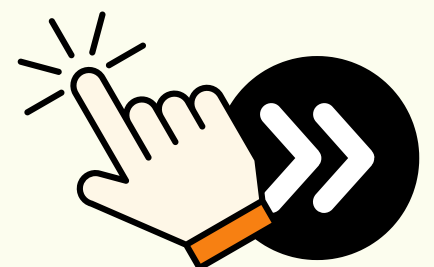
5. ¿Puede plagiar una IA?

La IA por sí sola no puede plagiar, porque no tiene intención. Quien lo hace siempre es el **humano que la utiliza**.

Así como la IA no puede tener autoría, tampoco tiene **responsabilidad**. Quienes generamos y difundimos la obra somos nosotros, los humanos.

La IA puede generar contenido que infrinja derechos de autor y plagiar **si ha sido entrenada previamente con material protegido**. Por ejemplo, podríamos entrenar una herramienta de generación de imágenes con las ilustraciones de un autor (algo muy sencillo hoy en día), y reproducir su estilo.

Estaríamos violando las leyes actuales en dos direcciones: subiendo ilustraciones de un artista sin tener la autorización para ello, y generando contenido a partir de su estilo desde una IA.

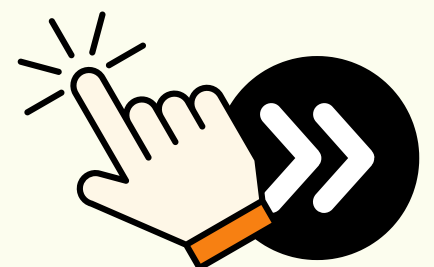


5. ¿Puede plagiar una IA?

Las IAs más ampliamente utilizadas hoy, como ChatGPT, solo pueden plagiar cuando se producen estas dos condiciones:

- El dato en su base de conocimiento está **sobrerrepresentado**. Es decir, se trata de una información con mucha presencia en su conocimiento. Sería el caso, por ejemplo, de Spider Man o Harry Potter.
- Se lo **induce** mediante el prompt a generar ese contenido que la IA ya conoce. Nuevamente, **la responsabilidad recae en la persona** que le solicita a la IA plagiar el contenido. Aunque los modelos de IA están preparados para evitar que esto ocurra, sigue siendo factible.

Lo que debemos saber: los modelos de IA generativos, **generan contenido novedoso** todo el tiempo. Técnicamente, no realizan un proceso de *copy - paste*. Generan contenido, tal como lo hace un humano, a partir de sus conocimientos. El riesgo de generación de contenido idéntico es solo posible si la IA ha sido entrenada con material ampliamente distribuido en Internet.

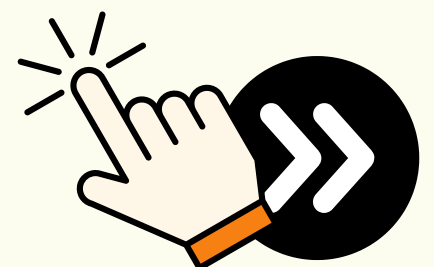


5. ¿Puede plagiar una IA?

El dato: A pesar de que los modelos de IA han sido generados inicialmente a base de **haber accedido a contenido con y sin copyright**, prácticamente **no existen juicios entre empresas hasta la fecha** (sí desde autores o artistas). El caso más relevante es el del The New York Times.

Esto señala que nos encontramos en una nueva era, sobre la que debemos sentar nuevas regulaciones. Es la primera vez que las máquinas generan contenido.

Ojo: en la actualidad, los modelos de IA se entrenan en gran parte con **contenido sintético**. Es decir, contenido que los mismos modelos generan.

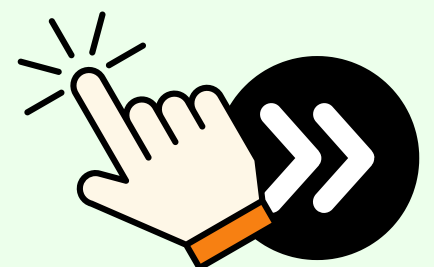


6. ¿Debería una editorial anunciar el uso de IA en sus publicaciones? ¿Cómo?

Sí, es recomendable que una editorial sea **transparente** sobre el uso de IA en sus procesos o publicaciones. Esto es además una medida que comienzan a exigir varias legislaciones, como la Ley de IA de Europa.

Una leyenda como "***Este contenido ha sido generado/optimizado con herramientas de IA bajo supervisión humana***", puede ser suficiente para informar a los lectores. Adicionalmente podríamos incorporar referencia a las herramientas empleadas.

Debemos saber igualmente que todavía no es una práctica muy extendida.

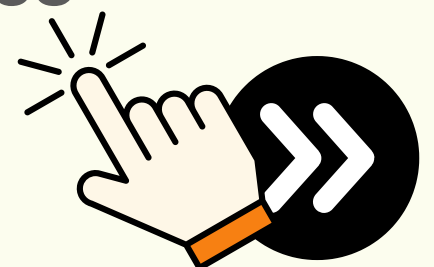


7. ¿Qué son los sesgos en la IA y cómo pueden impactar en los procesos editoriales?

Los **sesgos** en la IA son **prejuicios inherentes a los datos con los que ha sido entrenado un modelo**, lo que puede llevar a resultados desequilibrados o discriminatorios.

Los modelos de IA han sido **diseñados a imagen y semejanza de la humanidad**. Es decir, han sido entrenados con contenido que hemos producido y difundido los seres humanos, que estadísticamente está sesgado. Por lo tanto, el contenido que generan las IA **pueden reproducir estereotipos culturales o de género**.

Pero atención: a partir de las solicitudes que nosotros mismos le demos a los modelos, **podemos hacer que la IA adopte múltiples perspectivas**. Y esto es realmente muy interesante.



7. ¿Qué son los sesgos en la IA y cómo pueden impactar en los procesos editoriales?

Puedo, por ejemplo, solicitarle a la IA que escriba un ensayo desde la creencia de que la Tierra es plana. La IA **tiene un conocimiento amplio y profundo del mundo**, más que cualquier ser humano, y por tanto habitan en ella prácticamente **todas las perspectivas y miradas**.

Introducción a la Perspectiva Terraplanista

Como terraplanista, mi análisis se basa en la crítica a la interpretación convencional de las evidencias y la exploración de alternativas que considero han sido ignoradas o malinterpretadas por la ciencia establecida. Argumento que la Tierra es plana basándome en una serie de observaciones y teorías que cuestionan la noción de la esfericidad terrestre.

Análisis Crítico de la Evidencia Empírica

1. **La Observación del Horizonte:** El horizonte se presenta como una línea recta y plana a la vista, independientemente de la altitud del observador. Según la geometría esférica, la curvatura de la Tierra debería ser visible y medible desde ciertas altitudes. Sin embargo, experimentos como el realizado por el inventor y terraplanista Samuel Rowbotham, sugieren que no hay tal curvatura. Rowbotham realizó observaciones a lo largo de un canal recto de seis millas, concluyendo que no había evidencia de curvatura.
2. **La Física de los Fluidos:** Desde una perspectiva física, los líquidos en reposo tienden a asumir una superficie nivelada y plana. En grandes cuerpos de agua como los océanos, no se observa ni se mide empíricamente una curvatura conforme a lo que cabría esperar en una Tierra esférica.
3. **Problemas con la Fotografía Espacial:** Los terraplanistas argumentan que las fotografías de la

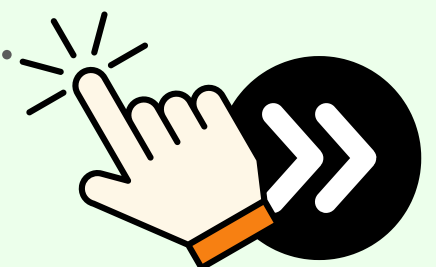


8. ¿Cuál es la diferencia entre los sesgos y la alineación de los modelos?

La **alineación** se refiere al proceso de ajustar el modelo para que sus resultados sean coherentes con los valores y principios establecidos por quienes lo han entrenado. Es decir, los modelos de IA vienen **cargados de ética y moral con la que no necesariamente acordamos**.

Muchos modelos como ChatGPT aplican restricciones fuertes a la hora de generar contenido relacionado con **violencia explícita, sexualidad, temáticas extremistas**, etc.

Esto podría llevar a que, por ejemplo, la IA no identifique el “contexto artístico” de manera profunda: es decir, no discierne si el uso de lenguaje violento o sexual es meramente decorativo o tiene un **propósito literario**.



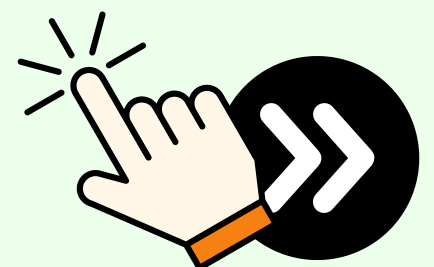
8. ¿Cuál es la diferencia entre los sesgos y la alineación de los modelos?

Como no controlamos directamente la **programación ni las políticas internas de los sistemas de IA**, podemos tener desencuentros entre lo que el modelo considera “apropiado” o “inapropiado” y lo que entendemos que es necesario o aceptable en un contexto creativo, literario o cultural.

Esta es una de las principales razones por las que muchos gobiernos del mundo están entrenando sus propias IAs: para que puedan estar basadas en la moral, ética y valores de su propio país.

Debemos entender esto como un **límite serio, problemático y concreto**, difícil de franquear, en el uso de estos sistemas.

Los modelos que provienen de China, como DeepSeek, han sido criticados por sus limitaciones para hablar de ciertos tópicos. Pero esto es algo presente en todos los modelos, no importa su origen.



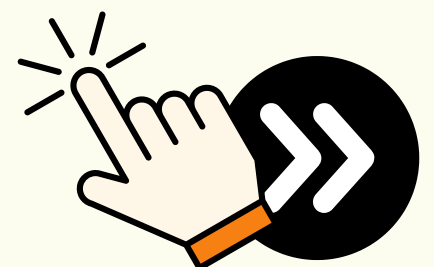
9. ¿Y qué es exactamente la alucinación en los modelos?

La **alucinación** ocurre cuando una IA genera información que **parece verosímil pero que en realidad es falsa** o no tiene fundamento en datos reales. Por ejemplo, una IA podría **inventar citas, fechas o datos que no existen**.

La IA genera **respuestas basadas en patrones probabilísticos** en lugar de en conocimiento verificable, lo que puede llevar a la fabricación de datos erróneos.

A medida que han ido evolucionando los modelos e incorporado herramientas de verificación, se ha reducido el factor de alucinación, pero por su propia lógica de funcionamiento, **existe la posibilidad de que un modelo de IA genere datos erróneos**.

Para evitar este problema, **es crucial verificar toda la información generada por IA** y complementarla con investigación y supervisión humana.



Es decir, los editores, en tanto humanos responsables en la publicación de libros, son los responsables de:

- No **reproducir y expandir los sesgos** sociales, mayoritariamente establecidos, que las IA tienden a desarrollar.
- No **publicar y difundir datos falsos o** erróneos que la IA pueda generar.
- No intervenir las obras con una **ética, moral y valores propios de la IA que entren en conflicto** con su visión del mundo.

La IA es una herramienta muy potente, pero somos nosotros los responsables de cara a los lectores.

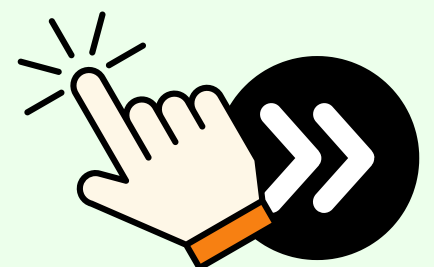


10. ¿Entonces? ¿Qué contenido se puede trabajar con IA?

En definitiva, **cada editorial es libre de definir cuál es el alcance de su política**, y en dónde considera idóneo trabajar un contenido con IA.

Por ejemplo, la creación de resúmenes, el análisis comercial o el diseño de imágenes como piezas de comunicación son procesos ideales, y que la mayoría de las empresas ya realiza cotidianamente.

En el caso de emplear IA en el tratamiento de los textos de una obra literaria (por ejemplo, para la traducción) o el diseño de portadas, ya será la propia editorial la que deba establecer los límites, y definir incluso qué herramientas serán las autorizadas.



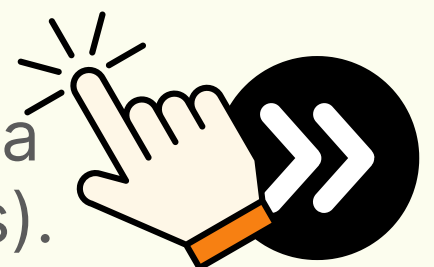
11. ¿Cuáles son las regulaciones actuales sobre IA que afectan a la industria editorial y qué implican?

Regulaciones como la **Ley de IA de la Unión Europea** están diseñadas para garantizar la **transparencia**, la **seguridad** y la **responsabilidad** en el uso de inteligencia artificial.

Esta normativa exige, por ejemplo, informar al usuario cuando se utiliza contenido generado por IA, garantizar la privacidad de los datos y prevenir el uso de modelos que puedan generar resultados discriminatorios.

Actualmente es una de las **pocas regulaciones específicas que existen en el mundo**.

Indirectamente, aplican otras normativas como los **tratados internacionales sobre derechos de autor** o el **RGPD** (que protege la privacidad de la información de las personas).



Una guía para definir la política editorial

Debemos por tanto establecer una política que incluya todas estas definiciones:

- ¿Qué **leyes** o **normativas** regulan el uso de las herramientas dentro del equipo de trabajo? ¿Qué dicen estas leyes y cómo podemos cumplirlas?
- ¿Qué nivel de **transparencia** se establecerá con el equipo interno y de cara a los lectores en el uso de la IA?
- ¿En qué **tareas** específicas se considera idóneo el uso de la IA y en cuáles estará prohibido?
- ¿Sobre qué **contenidos** se autoriza a utilizar IA y en cuáles no?
- ¿Qué **herramientas** se aconsejan emplear y de qué forma?