

## Las fotos manipuladas de Rosalía y el deep fake porno: ¿cuáles son los límites de la inteligencia artificial?

Por: **FABIANA SOLANO. 31/05/2023**

JC Reyes, rapero y artista, compartió una foto de Rosalia desnuda creada con inteligencia artificial. El consentimiento y la violencia digital, en el eje de la problemática.

La **Inteligencia Artificial** (IA) ha llegado para quedarse, y no necesariamente para facilitarnos la vida. En los últimos meses ha crecido exponencialmente su uso en diferentes ámbitos, hemos conocido las habilidades creativas del ChatGPT o la minuciosidad con la que algunos programas generan fotos casi perfectas de personas famosas en situaciones extraordinarias. Lo demostraron los ejemplos recientes de las imágenes del expresidente de Estados Unidos Donald Trump siendo detenido o las fotos del Papa Francisco con su campera de trapero blanca. Pero el registro público sobre estas tecnologías puso un freno a la exacerbación y se corrió de la mirada del asombro para dar lugar a la advertencia sobre el uso indiscriminado y la manipulación de la inteligencia artificial.

La semana pasada la víctima fue la artista catalana **Rosalía**. **En las redes sociales se difundió una imagen suya con el torso desnudo como si fuera real**. Quien la compartió fue el rapero Juan Manuel Cortés Reyes, conocido como JC Reyes, quien además se jactó de hacerlo con la frase “lo mejor que vas a ver hoy” para sus más de 230 mil seguidores. La fotografía se trata de una **manipulación sin consentimiento de un contenido que ella había compartido previamente en sus redes**.

La autora de hits como “Despechá” y “Con Altura” hizo su descargo en su cuenta de Twitter luego de la sexualización y definió al episodio como “violencia”: “Ir a buscar clout(influencia) faltando el respeto y sexualizando a alguien es un tipo de violencia y da asco, pero hacerlo por 4 plays es lo que da más pena”, aseguró en su cuenta de Twitter. Y agregó: “El cuerpo de una mujer no es propiedad pública, no es una mercancía para tu estrategia de marketing. Esas fotos estaban editadas y creaste una falsa narrativa alrededor cuando ni te conozco. Existe algo llamado

consentimiento y todos a los que os pareció gracioso o plausible espero de corazón que un día aprendáis que venís de una mujer, que las mujeres somos sagradas y que se nos ha de respetar. Bye”.

Ir a buscar clout faltando el respeto y sexualizando a alguien es un tipo d violencia y da asco pero hacerlo por 4 plays de + lo q da es pena

— R O S A L Í A (@rosalia) [May 23, 2023](#)

Algo similar le ocurrió a Johanna Villalobos una influencer y periodista de Costa Rica que denunció la viralización de una falsa foto suya desnuda creada a través de inteligencia artificial: “Usaron mi foto de la playa y con Inteligencia Artificial “me quitaron la ropa”.. estoy harta de ver cómo nos utilizan para su entretenimiento, la foto está en Grupos de WhatsApp, redes sociales, NO SOY YO y lo más lamentable de todo es que no puedo hacer nada al respecto”.

Usaron mi foto de la playa y con Inteligencia Artificial “me quitaron la ropa”.. estoy harta de ver cómo nos utilizan para su entretenimiento, la foto esta en Grupos de WhatsApp, redes sociales, NO SOY YO y lo más lamentable de todo es que no puedo hacer nada al respecto. [pic.twitter.com/Er78wXkTb](https://pic.twitter.com/Er78wXkTb)

— Johanna Villalobos ? (@johavillam) [May 23, 2023](#)

Estos casos han trascendido ya que se trata de personas públicas con fuerte llegada en las audiencias y con una voz amplificada que permite que sus denuncias se legitimen rápidamente. No obstante, por el crecimiento de la inteligencia artificial y el acceso rápido y sencillo a herramientas de manipulación, existen miles de casos de Deepfake, es decir la técnica de síntesis de imágenes humanas. Las imágenes sexualizadas hiperrealistas producidas sin consentimiento alimentan el mercado de la pornografía y la pedofilia, crean contenido para adultos, reproducen la violencia digital de género en la dark web y son muy difíciles de identificar.

**El desafío del sector pareciera ser entonces, en el corto y mediano plazo, la determinación de límites técnicos y legislativos que pongan un freno a la**

inteligencia artificial. No obstante el debate a dar es también político y profundamente sociocultural.

## Los límites de las IA: peligros, advertencias y desafíos

**Farid Murzone**, es líder de tecnología e inteligencia artificial en ÜMA Salud, y explica en principio que el impacto de la inteligencia artificial es asimilable a lo que fue la revolución industrial o a la aparición de Internet: “Está impactando mucho en todo lo que es generación de contenidos de distintos tipos de texto, artículos, papers académicos, ideas, chats, guiones, etcétera, y en imágenes para también generar fotos, marcas, logos, etcétera, y en vídeo. Inclusive ya se están haciendo algunas animaciones 100% con IA”. Desde una mirada positiva, **Murzone destaca su incorporación en la vida cotidiana y en la mayoría de las profesiones “como soporte, no como sustituto, porque todavía quedan un montón de cosas por pulir y por perfeccionar y hay un montón de cosas del intelecto humano en cuanto a creatividad que todavía la inteligencia artificial no hace”**.

Sin embargo, advierte que su uso plantea una serie de preguntas y desafíos. “Cada vez que irrumpe una nueva tecnología, obviamente también aparecen nuevos problemas que hay que legislar, que hay que cuidar. Como pasó con internet, por ejemplo, la necesidad de moderar que un menor de edad no esté chateando con un adulto que le esté pidiendo información personal y fotos. Más tarde apareció también el ciberacoso, sobre todo a adolescentes. **Con la inteligencia artificial el punto central es que se puedan generar contenidos de forma muy fácil, tanto de texto como imágenes y eventualmente vídeos**”.

**Para el especialista enfrentar el problema implica dos acciones paralelas: pensar una regulación sobre la generación de contenidos y la persecución del uso inadecuado de la generación de esos contenidos; y por otro lado generar también nuevas herramientas para poder identificar esos contenidos falsos:**

“Ya es muy difícil identificar fake news y sobre todo evitar que penetren en la sociedad. Encima ahora la inteligencia artificial te permite generar automáticamente y masivamente ese tipo de contenidos hiper realistas, tanto las imágenes falsas o creadas artificialmente como los artículos generados artificialmente”.

## **IA y los sesgos sociales: el dilema del huevo y la gallina**

Cuando se habla del avance de la inteligencia artificial muchas veces se lo suele relatar como un suceso involuntario, como un hecho de la naturaleza por fuera del radio de movimiento humano. Por el contrario, lo primero a aclarar es que la IA se trata de una tecnología creada y desarrollada por seres humanos con fuertes sesgos sociales y puesta en funcionamiento en medio de dinámicas socioculturales que producen sentido sobre las cosas.

“Hasta el día de hoy la inteligencia artificial no tiene capacidad de generar contenidos por sí sola. Todos los contenidos, las imágenes y los textos ocurren por instrucción de un humano – explica Farid Murzone – lo mismo que con internet, lo mismo que con la tecnología nuclear, siempre está la responsabilidad humana y los controles que podamos hacer para poder utilizarla adecuadamente”. **La inteligencia artificial nunca es neutral ya que reproduce los sesgos, prejuicios, miradas y estigmas de quienes la codifican, primero, y de quienes la utilizan, en segundo lugar.**

“El gran peligro inicial es que toda la inteligencia artificial está generada con datos de la vida real. Ya pasó con los modelos generativos de texto: el año pasado hubo mucha polémica en torno a GPT-3, un modelo muy revolucionario que generaba los primeros corpus de textos bastante creíbles, que parecían escritos por una persona. Sin embargo los contenidos que generaban, en la mayoría de los casos, eran racistas – relata el líder de tecnología e inteligencia artificial- Si vos le preguntabas a GPT-3 o le pedías un texto sobre una persona negra ingresando en un negocio comercial, posiblemente te redactaba un artículo sobre un delito. Si le preguntabas

sobre un musulmán, posiblemente te escribía un artículo sobre terrorismo. Tienen todos estos sesgos porque los modelos están generados en base a textos de personas que parten de miradas patriarcales, violentas, xenófobas y homofóbicas”.

Gabriela Ramos, subdirectora General de Ciencias Sociales y Humanas de la Unesco, hace un tiempo en una entrevista explicó: **“No es que las tecnologías discriminen, es que han sido entrenadas para discriminar.** Y han sido entrenadas para discriminar porque no tenemos conciencia de cómo construirlas de una manera inclusiva (...) Tú le preguntas al ChatGPT descríbeme a una mujer y te la describe en la casa, y a los hombres te los describen exitosos llegando a la Luna”. Otra herramienta de IA, por ejemplo, diseñada para Amazon con el fin de buscar personas para un puesto laboral, no percibía a las mujeres como candidatas exitosas.

Por ello un grupo de investigadoras y profesionales del mundo académico, de la sociedad civil, del sector privado y de organismos reguladores de todo el mundo, están impulsando una nueva plataforma llamada Women4Ethical con la misión de aumentar el reconocimiento, la representación y el empoderamiento de las mujeres en la ética de la IA. “Dada la omnipresencia de los algoritmos de inteligencia artificial (IA) sesgados y el riesgo para las vidas humanas de los sistemas de IA defectuosos, existe una necesidad urgente de una discusión abierta y una acción concreta para abordar los peligros de la IA sin control”, explican en su página web.

## IA y violencia digital

“Este es un claro caso de violencia de género digital y es uno de los grandes peligros que existe hoy con las deepfakes de imágenes porque **las IA se están usando muchísimo con este sentido en la dark web, en páginas porno, etcétera. Uno de los grandes desafíos va a ser, así como hubo que generar herramientas contra la pedofilia en internet, vamos a necesitar crear herramientas para poder perseguir a las personas que generan violencia de género digital**”, identifica Murzone.

Los casos de manipulación de imágenes de identidades feminizadas para crear fotografías sexuales que se viralizan y comercian en la Deep web constituyen un delito que puede catalogarse dentro del tipo **violencia de género digital**. Si bien en Argentina todavía no existe la figura penal ni una regulación que permita controlarlo

y penar a quienes la ejercen, el **tema se está debatiendo a nivel social y existe un proyecto presentado en el Congreso de la Nación que contempla la Ley Olimpia y la Ley Belén, dos iniciativas inspiradas en una legislación recientemente aprobada en México.**

“Rosalía al ser una persona famosa que está más expuesta, seguramente, ya está más acostumbrada a sufrir este tipo de cosas y tiene las herramientas y el dinero para poder afrontar una situación como esta. Pero una persona común y corriente, supongamos vos o yo, si nos generan una campaña con deepfakes, con fotos falsas, etcétera, no tenemos absolutamente ningún recurso para protegernos. **Hoy no existe realmente legislaciones para esto, si hay algunas que contemplen la difamación y la información falsa, pero no existen todavía herramientas tecnológicas para poder limitar el impacto de todo eso**”, advierte el entrevistado.

## Modelos de IA y filtros de toxicidad

En algunos sectores se empezó a utilizar y aplicar lo que se llaman filtros de toxicidad, lo que significa que **“la persona que entrena los modelos de IA aplica intencionalmente filtros para que no se hable de determinados temas o, si habla de esos temas, lo haga con determinadas precauciones que se pueden establecer en la codificación”**. El chatGPT, por ejemplo, está codificado para que al preguntar sobre cuestiones relacionadas al color de piel de alguien o la orientación sexual, la inteligencia artificial no pueda responder.

Muchos de los modelos de generación de imágenes también tienen esos filtros: “La misma empresa que hace ChatGPT tiene un modelo que se llama DALL-E que genera imágenes con muy buena resolución y tiene filtros de toxicidad. **Si vos le pedís una foto de un político o de una persona famosa, no te lo genera.** El peligro está en que **hay muchos modelos open source de código abierto y por ende cualquier filtro de toxicidad puede ser anulado.** Es decir, una persona con capacidad técnica para utilizar ese modelo y generar los contenidos que quiera lo puede hacer sin aplicar filtros de toxicidad. Ahí nuevamente entran los desafíos regulatorios”.

## Regular al Frankenstein

La necesidad de algún tipo de regulación sobre las IA es urgente porque el fenómeno está creciendo a pasos agigantados. En este sentido, a fines de marzo se conoció que un grupo de mil científicos, intelectuales y especialistas, entre ellos el dueño de Twitter Elon Mask, **alertó en un documento público dirigido a la comunidad internacional sobre un posible “riesgo para la sociedad y la humanidad” y pidió detener la evolución de las IA por un período de 6 meses para establecer límites y controles.**

El gran problema es que los desarrollos de las inteligencias artificiales avanzan a un ritmo más acelerado que la capacidad que tienen las sociedades y los sistemas políticos de legislar sobre ellas: **“En general las legislaciones son bastante lentas, perezosas, y la inteligencia artificial y el desarrollo técnico no para.** Las empresas que están a la vanguardia del desarrollo no lo están deteniendo, por el contrario, cada vez más otras grandes empresas se están sumando a esa carrera. Esto nos plantea la necesidad urgente de poder generar mecanismos para legislar y poder castigar el uso inadecuado de estas tecnologías”.

**[LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ](#)**

Fotografía: El destape web

**Fecha de creación**

2023/05/31