

'Deepfakes' y reputación corporativa: ¿Cómo afrontar una potencial crisis por inteligencia artificial?

Ciudad de México, 23 de abril de 2024. – Con el creciente desarrollo, avance y uso de inteligencia artificial (IA) en los últimos años, también ha ocurrido un nuevo fenómeno que es usado por manos humanas con el objetivo de crear secuencias falsas pero realista, que se expande para crear campañas de desinformación: las *deepfakes*.

De acuerdo con la [Fundación del Español Urgente](#) (Fundéu RAE), organización asesorada por la Real Academia Española, la palabra *deepfake* es un anglicismo que alude a los sistemas informáticos que permiten, mediante técnicas de inteligencia artificial, desarrollar videos manipulados extremadamente realistas.

A través de una *deepfake* se puede modificar la apariencia física e incluso la voz de las personas, y se utilizan comúnmente para distorsionar mensajes o manipular la opinión pública, principalmente de actores políticos.

De acuerdo con Carlos Maya, SVP Client Services y [another](#), agencia de comunicación estratégica con la mayor oferta en América Latina, a medida que la inteligencia artificial avanza, también crece el desafío de detectar, y verificar, este tipo de técnicas con el uso de IA o aplicaciones que se encuentran disponibles tanto en celulares como en computadoras, es decir, cualquier persona puede tener acceso a este tipo de tecnologías.

“Aunque actualmente las ‘deepfakes’ han sido usadas en distintos contextos donde aparecen actores políticos, celebridades, actores de doblaje o creadores de contenido, la realidad es que cualquier persona puede ser víctima de este tipo de situaciones, incluyendo voceros de empresas o marcas”, indicó Carlos Maya.

En situaciones donde un vocero se ve afectado por un *deepfake*, la respuesta corporativa puede ser relativamente clara: se desmiente públicamente y se toman medidas para corregir la información errónea, subraya Carlos Maya. Sin embargo, cuando una *deepfake* involucra a un líder de opinión (KOL, por sus siglas en inglés) hablando mal de una empresa, la gravedad de la crisis aumenta exponencialmente. Este tipo de incidentes no solo son más difíciles de contener, sino que también pueden tener un impacto inmediato y duradero en la percepción pública de la marca.

Aunque la verificación de fuentes sigue siendo fundamental, especialmente en el mundo digital, la gestión de la reputación se vuelve aún más crucial en estos escenarios. La creencia errónea en *deepfakes* que difaman a terceros puede llevar a consecuencias graves para la reputación de una empresa, incluso si el contenido resulta ser falso o humorístico. Por lo tanto, las empresas deben estar preparadas para enfrentar estas crisis de manera rápida, efectiva y transparente, manteniendo siempre la integridad de su marca como prioridad principal.

“Hay personas que sí creen en las ‘deepfakes’ cuando se trata de una tercera persona o un líder de opinión, donde su imagen y voz son alterados a través de inteligencia artificial. Aquí es fundamental que las empresas cuenten con un manual de crisis actualizado para este tipo de situaciones, y es fundamental contar con un equipo especializado para activar las respuestas”, dijo Maya.

Según el experto, el uso de inteligencia artificial llegó para quedarse, sin embargo, hay todo tipo de voces que están en contra de las *deepfakes*. Por ejemplo, en febrero más de 400 expertos firmaron una carta en contra de la proliferación de estas, para exigir a los gobiernos del mundo medidas urgentes.

El documento titulado [Disrupting the Deepfake Supply Chain](#), sostiene que los contenidos con *deepfakes* “son una amenaza creciente para la sociedad y que los gobiernos deben imponer obligaciones en toda la cadena de suministro para detener su proliferación”.

Además, la proliferación de videos falsos y engañosos creados con inteligencia artificial creció un 550% entre 2019 y 2023, según el informe, [State of Deepfakes 2023 de Home Security Heroes](#). Además, la organización indicó que detectaron cerca de 95,820 videos deepfake en línea, tan solo el año pasado.

Maya advierte que el robo de identidad es una amenaza latente para todo tipo de personas, y esto incluye a las empresas. Por lo tanto, los voceros y expertos en marketing deben contar con un equipo de crisis, comunicación estratégica y herramientas tecnológicas con IA para tener un plan de respaldo si enfrentan este tipo de problemas.

En este sentido, el experto de *another* comparte una estrategia de crisis para afrontar una deepfake que afecte a cualquier vocero o líder detrás de las marcas o empresas.

Actuar con rapidez y determinación

Ante la detección de una deepfake que involucre a un vocero de la empresa o un KOL aludiendo a la empresa, es crucial actuar de manera inmediata. La velocidad de respuesta es esencial para limitar el alcance del video manipulado y mitigar el daño potencial a la reputación corporativa.

Verificación de la autenticidad

Es fundamental verificar la autenticidad del video en cuestión. Esto puede implicar el trabajo en colaboración con expertos en tecnología forense digital y herramientas de verificación de contenido para determinar si el video es genuino o una manipulación.

Comunicación transparente

Una vez confirmada la existencia de una *deepfake*, la empresa debe comunicarse de manera transparente con su audiencia. Es importante informar de manera clara y concisa sobre la situación, explicando que se trata de un video manipulado y desmintiendo cualquier afirmación falsa que se presente en el mismo.

Estrategia de Gestión

Una agencia de comunicación estratégica puede desempeñar un papel fundamental en desarrollar una estrategia diseñada para gestionar la crisis eficazmente. Esto puede implicar la creación de mensajes claros y coherentes, la preparación de portavoces para entrevistas y declaraciones públicas, y el monitoreo constante de la situación en los medios y las redes sociales, así como comunicación y relacionamiento con *stakeholders* clave.

Educación y concientización

Además de abordar la crisis actual, la empresa puede tomar medidas proactivas para educar a su audiencia sobre el problema de las *deepfakes* y promover la concientización sobre la importancia de verificar la autenticidad del contenido en línea, sin importar la fuente.