¿Qué son los deepfakes y cómo las organizaciones pueden combatirlos?

A la par del crecimiento de la transformación digital en las organizaciones, también han aumentado las diversas formas de fraude electrónico o cibercrímenes. Existe uno en particular que viene aumentado en estos últimos tiempos: el fraude a través de los **deepfakes**.

Esta técnica que, si bien era usada como una técnica simple de face swapping o cambio de rostros, se ha vuelto más problemática en los últimos años al causar desinformación y alterar la percepción de la confianza que se pueda tener sobre la información.

A partir de esto, las personas se han vuelto más susceptibles a la información que se les presenta tanto en las redes como en medios tradicionales, sin tener la certeza de lo que puede ser real o no.

"A pesar de lo futurista que suena, es posible que ya hayamos interactuado con esta tecnología en nuestras redes sociales. De acuerdo con la firma de seguridad Sensity, al mes de julio de 2019 el 95% del uso del deepfake estaba destinado al entretenimiento. Desde conocidos videos deepfake de personajes como Mark Zuckerberg, Cofundador de la red social Facebook, describiendo cómo una organización le habría mostrado el potencial de manipular a la sociedad para así obtener gratuitamente la información de las personas y sus seres

cercanos, finalizando con la frase "mientras más te expresas, más nos perteneces"; hasta Vladimir Putin, presidente de Rusia, hablando acerca de la manipulación de las campañas electorales de Estados Unidos en el año 2016", comentó **Henry Matta**, Gerente Senior de Riesgos de Integridad de **EY Perú**, firma de servicios de auditoría, impuestos, estrategia, transacciones y consultoría.

A continuación, **EY Perú** señala a detalle qué son los deepfakes, las medidas que las organizaciones pueden tomar para combatir este nuevo tipo de fraude electrónico, y las diferentes modalidades en las que este tipo de cibernético puede ocurrir:

Deepfake: Esta técnica hace uso de tecnologías disruptivas como la inteligencia artificial para extraer los patrones de datos contenidos en un conjunto de imágenes o audios relacionados a una persona.

El siguiente paso involucra la aplicación de técnicas de aprendizaje como Machine Learning para superponer de forma convincente esos datos extraídos sobre otra persona. Finalmente, el objetivo es lograr una personificación que puede llegar a igualar no solo la apariencia y la voz, sino también los gestos y vocalización.

Campo de uso

Usualmente, esta técnica ha sido utilizada en celebridades o

políticos para crear narrativas falsas o impactar en la imagen de una determinada persona.

Sin embargo, los deepfakes han incursionado recientemente en el mundo laboral; por ejemplo, en agosto de 2019, el Wall Street Journal publicó un artículo en el cual se describe cómo una empresa (cuyo nombre no fue revelado) domiciliada en **Reino Unido** fue víctima de un ciberataque que utilizó deepfake para suplantar la voz del CEO corporativo.

De acuerdo con el ejecutivo británico, la voz del CEO corporativo habría emulado también su acento alemán y habría solicitado realizar una transferencia urgente a la cuenta bancaria de un proveedor húngaro; ocasionando así una pérdida de US$ 243,000.

Medidas preventivas

Entrenamiento: Debido a que el problema generado por el deepfake se basa en la percepción de confianza sobre la información, es importante entrenar de forma efectiva a todo el personal con el objetivo que puedan evaluar con escepticismo las solicitudes poco usuales y recurrir a mecanismos seguros de comunicación para confirmar la información o instrucciones recibidas.

Análisis forense: La aplicación de técnicas avanzadas de cómputo forense permiten evaluar la integridad de un mensaje, a través del análisis de la metadata (atributos tecnológicos)

del mensaje, así como la identificación de patrones en el contenido del mismo (naturalidad del movimiento, secuencia de cuadros, consistencia de la iluminación y sombras, entre otros).

Blockchain: Esta tecnología permite insertar datos de validación dentro de la información corporativa, proporcionando así un nivel de seguridad e integridad ante la manipulación o falsificación de contenido. Con apoyo de esta tecnología, el material oficial y verídico en una organización tendría una validación exitosa en la base de datos Blockchain; mientras que el material editado tendría una validación fallida.

"El deepfake es un ejemplo actual de cómo los ciberataques se vuelven cada vez más sofisticados, aplicando inclusive tecnologías disruptivas como la inteligencia artificial a los tradicionales esquemas de phishing e ingeniería social. Este tipo de fraude aprovecha las características de la modalidad actual de trabajo remoto y tiene el potencial de propagarse rápidamente. A futuro, el deepfake podría generar un importante problema de confianza en las organizaciones y en el público en general, por lo que resulta clave poder trabajar pronto en soluciones que permitan a las organizaciones mitigar el potencial impacto al que podrían estar expuestas", finaliza el experto de **EY Perú**.