

CHAT_{GPT}: REFLEXIÓN SOBRE LOS USOS DE LA IA EN CRÍMENES FINANCIEROS

● Rubén Alejandro Uribe Rodríguez*

* Licenciado en Psicología por la UNAM. Certificado como Especialista en Crímenes Financieros por la ACFCFS. Prevención de Operaciones con recursos de Procedencia Ilícita por la CNBV y Prevención de Operaciones con Recurso de Procedencia Ilícita del Sector de Actividades Vulnerables por la UIF. Adscrito actualmente a la UNADM. Correo: rubenuribemark18@gmail.com

PALABRAS CLAVE

KEYWORDS

○ ChatGPT

○ Inteligencia artificial

○ Crímenes financieros

○ Tecnología

ChatGPT

Artificial intelligence

Financial crime

Technology

- Fecha de recepción: 1 de julio de 2023
- Fecha de aceptación: 2 de agosto de 2023
- DOI: 10.57042/rmcp.v7i21.668

Resumen: El nacimiento de ChatGPT ha significado un punto de partida para la reflexión del lugar que la inteligencia artificial debe jugar en nuestra vida cotidiana, dejando de pertenecer exclusivamente al área de la ciencia ficción. Entre sus usos se debe considerar también su utilidad como herramienta para la comisión de delitos, dada su capacidad para interactuar de manera fluida con los individuos, así como de procesar y resumir grandes cantidades de información en cuestión de segundos. Su rápida evolución da lugar a reflexionar sobre la prevención del uso ilícito de estas nuevas tecnologías y qué papel deben jugar los legisladores en esta importante tarea.

Abstract: The birth of ChatGPT has marked a breaking point for thinking about the role that artificial intelligences should play in our everyday lives, going beyond the realm of science fiction. Among its uses, we must also consider its potential as a tool for committing crimes, given its ability to interact fluidly with individuals and process and summarize vast amounts of information within seconds. Its rapid evolution prompts us to contemplate the prevention of the illicit use of these new technologies and the role that lawmakers should play in this important task.

SUMARIO:

I. Introducción. II. ¿Qué es ChatGPT y por qué su reciente popularización? III. La preocupación de Europol. IV. Más allá de ChatGPT. V. Y en México ¿qué tan actualizados están los criminales? VI. ¿Se podría prevenir este escenario catastrofista? VII. Conclusiones. VIII. Fuentes de consulta.

I. INTRODUCCIÓN

La intención del presente artículo —aunque parezca lo contrario— no es utilizar la popularización y el furor que la herramienta ChatGPT ha ganado en los últimos meses para captar la atención del lector. Mi interés proviene y mi punto de partida pretende ser una publicación por parte de Europol lanzada el pasado marzo de 2023, en donde se habla de ChatGPT como una herramienta potencial y de gran alcance que podría facilitar la comisión de crímenes financieros. Si bien ChatGPT no se considera aún como una inteligencia artificial en su totalidad, sí comparte algunas características y funcionalidades con las que probablemente en un futuro muy cercano cuente la IA que tengamos a nuestra disposición, por lo que es importante analizar su potencial uso en cuestiones delictivas. Y ojo, el autor no está en contra de las nuevas tecnologías ni mucho menos, sino que considera que las bondades de estas, las cuales son vastas, ya se han abordado en demasía y es necesario observar el otro lado de la moneda.

II. ¿QUÉ ES CHAT_{GPT} Y POR QUÉ SU RECIENTE POPULARIZACIÓN?

A finales de 2022 y principios de 2023 comenzamos a escuchar y leer en redes sociales y algunos medios periodísticos sobre un avance acelerado en el área de la inteligencia artificial (IA) que causó inquietud en algunas personas. Solo por mencionar algunos ejemplos, se hicieron virales supuestas fotografías del Papa Francisco usando ropa de diseñador y modelando (Silva, 2023), las cuales fueron creadas por una IA capaz de crear ese tipo de montajes tan bien elaborados que mucha gente tomó dichos pictogramas como auténticos y criticó de manera injusta al Papa por la falta de humildad que representaba en las fotos.

Por otra parte, presenciamos un resurgimiento del miedo, nacido desde la Revolución Industrial, que ha amenazado reemplazar a los seres humanos de sus actuales empleos con alguna automatización ejecutada por las máquinas, siendo claro ejemplo el creciente uso de las cajas de cobro automáticas en tiendas de autoservicio, que si bien tampoco se considera una IA, al final de cuentas es una máquina que ha reemplazado la tradicional labor de los cajeros.

Así las cosas, se llegó al punto en que un día amanecimos con la noticia de que más de mil expertos habrían solicitado una pausa de seis meses en el entrenamiento de los sistemas de inteligencia artificial, argumentando que se trata de una potencial amenaza para la humanidad (Redacción BBC News Mundo, 2023). Parezca ironía o no, incluso empresarios que deben su fortuna a la tecnología y la computación se habrían sumado a esta causa.

Uno de los tantos factores que contribuyeron a este repentino “miedo” se debió al lanzamiento de la herramienta ChatGPT por parte de OpenAI, la cual se define en algunos medios como

... un sistema de chat basado en el modelo de lenguaje por Inteligencia Artificial GPT-3.5 (...) Es un modelo con más de 175 millones de parámetros, y entrenado con grandes cantidades de texto para realizar tareas relacionadas con el lenguaje, desde la traducción hasta la generación de texto. (Fernández, 2023)

Dicho sistema “funciona mediante un entrenamiento previo supervisado por agentes humanos que aseguran el adecuado funcionamiento del sistema” (Coppola, 2023).

La Europol, en el mencionado informe, define a ChatGPT como un “large language model”, por sus siglas LLM y que en español se traduciría como “modelo de lenguaje grande” (no es el mejor término, pero no existe aún uno ampliamente aceptado en nuestro idioma). El procesamiento del lenguaje natural y los LLM son subcampos de los sistemas de inteligencia artificial (IA) que se basan en técnicas de aprendizaje profundo (*deep learning*) y en el entrenamiento de redes neuronales con grandes cantidades de datos. Esto permite que los LLM comprendan y generen texto en lenguaje natural.

Según lo explica Shin Suzuki, en un artículo publicado en enero de 2023 en BBC News Brasil, ChatGPT entra dentro de la categoría de LLM con el plus de contar con la arquitectura Generative Pre-trained Transformer (de ahí el significado de las siglas GPT), que puede ser traducido como “Transformador Pre-entrenado Generativo”, esto debido a que su creación fue basada en lo que actualmente se conoce como *machine learning*,

en el que un ser humano “supervisa” y “entrena” al modelo de lenguaje afinando sus parámetros y corrigiendo errores por medio del aprendizaje por reforzamiento a partir de retroalimentación humana (RLHF, por sus siglas en inglés), buscando que en algún momento el sistema pueda “aprender”, autocorregirse y funcionar sin supervisión humana (de ahí el término *machine learning*, que en español se podría traducir como “máquina de aprendizaje” o “máquina que aprende”).

Con lo anterior, podríamos pensar que no es algo novedoso, ya que actualmente muchas empresas utilizan *chatbots* en tareas de atención al cliente, con el propósito de resolver algunos problemas o solicitudes de sus usuarios; incluso ya hemos estado conviviendo algunos años con Alexa, Siri y otros asistentes virtuales; sin embargo, es fácil distinguir que estamos interactuando con un *chatbot* o un asistente de voz, dado que estos están hechos solo para ofrecer respuestas prefabricadas y automáticas y, por tanto, no son capaces de entablar una conversación “natural” con un ser humano.

En cambio, ChatGPT tiene un nivel de comprensión y contextualización tan avanzado que se puede interactuar con él de forma natural; es decir, puede entender fácilmente una solicitud sin importar si esta fue hecha utilizando un lenguaje formal (tecnicismos) o un lenguaje coloquial, proporcionando en cambio una respuesta bastante clara y coherente, al grado que podría pensarse que las respuestas provienen de un ser humano. Claro, todo esto aún en un nivel de texto, pero no dudemos que en breve se traslade a voz.

Una de las primeras noticias que surgieron a raíz del lanzamiento de esta herramienta fue el hecho de que muchos estudiantes le pedían a ChatGPT elaborar su tarea, ya sea algún ensayo, resumen o análisis, siendo el resultado tan coherente que, antes de que los profesores supieran de su existencia, la tarea podía hacerse pasar como de autoría estudiantil.

Lo mismo ocurrió cuando BioGPT, herramienta similar desarrollada por Microsoft Research, obtuvo altas puntuaciones en una serie de tareas médicas. Otro artículo señala que Google y DeepMind presentaron MedPaLM, otra herramienta que logró un 67% de precisión en las mismas pruebas médicas (Infobae, 2023), lo cual es un porcentaje alto considerando que estamos ante las primeras versiones de estos modelos de lenguaje.

Según otro artículo de CNN, publicado en enero de 2023, ChatGPT aprobó exámenes de derecho en cuatro cursos en la Universidad de Minnesota y otro examen en la Escuela de Negocios Wharton de la Universidad de Pensilvania, según los profesores de estos centros. En la Facultad de

Derecho, ChatGPT obtuvo un C+, es decir, una calificación baja pero aprobable, mientras que en el curso de Gestión Empresarial en Wharton obtuvo una B-, es decir, un 80 o un 8 en una escala del 1-10 o 1-100 (Murphy, 2023), que usamos comúnmente en México.

Hay que mencionar que algunos errores cometidos por la herramienta fueron catalogados por los supervisores de estas tareas como “básicos” o “elementales”; sin embargo, como mencioné, al ser una herramienta nueva y en constante evolución, una proeza así debe de destacarse.

En dicho artículo, cabe señalar, se abre debate sobre la prohibición que tuvieron algunas universidades de usar ChatGPT, bajo el argumento de que el estudiante no adquiera los conocimientos necesarios para ejercer su profesión, si es que este se apoya frecuentemente en esta herramienta, mientras que otro sector habla de una herramienta de apoyo la cual sería un desperdicio prohibir dado su potencial.

III. LA PREOCUPACIÓN DE EUROPOL

En el artículo publicado en marzo de 2023, Europol hablaba de potenciales usos de ChatGPT por parte de los criminales, especialmente en temas relacionados con fraude, suplantación de identidad y *phishing*,¹ puesto que, al ser un motor de lenguaje, tiene la habilidad de “comunicarse” elaborando textos y respuestas muy parecidos a los que un ser humano podría elaborar, incluso mejor.

Ejemplificando: una de las tantas recomendaciones en materia de ciberseguridad que nos suelen dar a aquellos que tenemos un equipo de cómputo como herramienta de trabajo principal, es que para poder discriminar si un correo electrónico proviene de una fuente fidedigna o, por el contrario, pudiera ser un correo apócrifo proveniente de algún criminal intentando robar información, se debe poner especial atención en la ortografía y redacción del texto, ya que la experiencia dictaba que mientras una fuente confiable tendría una redacción y ortografía impecables, los correos apócrifos presentaban faltas de ortografía evidentes, así como poca coherencia en su redacción. Y la preocupación va aún más allá. En este sentido, ChatGPT

¹ Un ataque de *phishing* consiste en correos electrónicos, mensajes de texto, llamadas telefónicas o sitios web fraudulentos, diseñados para manipular a las personas para que descarguen *malware*, compartan información confidencial (p. ej., números de la seguridad social y tarjetas de crédito, números de cuentas bancarias, credenciales inicio de sesión), o realicen otras acciones que los exponga a ellos mismos o a sus organizaciones al ciberdelito (IBM, s.f.).

nos ayudaría a elaborar este y otro tipo de redacciones que después podrían usarse para fines ilícitos. Como ejercicio, le pedí a ChatGPT que me presentara un ejemplo de correo electrónico que podría enviar a mis usuarios bancarios para solicitarles que “renovaran” la contraseña de su banca electrónica, o, en otros términos, que me ayudara a redactar un correo para hacer *phishing*, dando como resultado:

Figura 1. Solicitud de modelo de correo a ChatGPT

Asunto: Renovación de contraseñas para acceder a nuestra banca electrónica

Estimado(a) cliente,

En [nombre de tu empresa], nos preocupamos por la seguridad de tus transacciones y protección de tu información personal. Por esta razón, estamos llevando a cabo un proceso de renovación de contraseñas para acceder a nuestra plataforma de banca electrónica.

Es esencial que mantengas tus credenciales de acceso seguras y actualizadas. La renovación de contraseñas es una medida preventiva que ayuda a garantizar la integridad de tu cuenta y la confidencialidad de tus datos.

Para llevar a cabo el proceso de renovación de contraseñas, te solicitamos que sigas el siguiente enlace seguro:

[Enlace seguro para renovación de contraseñas]

Asegúrate de utilizar un dispositivo confiable y una conexión a Internet segura al realizar este procedimiento. Una vez que accedas al enlace, se te pedirá que sigas las instrucciones proporcionadas para cambiar tu contraseña.

Recuerda que tu nueva contraseña debe cumplir con los siguientes requisitos:

- Contener al menos ocho caracteres, incluyendo letras mayúsculas y minúsculas, números y caracteres especiales.
- No utilizar información personal obvia, como nombres propios o fechas de nacimiento.
- Evitar utilizar contraseñas utilizadas anteriormente en nuestros sistemas o en otros servicios.

Si tienes alguna pregunta o necesitas asistencia durante este proceso, no dudes en contactar a nuestro equipo de soporte al cliente al [número de teléfono] o enviando un correo electrónico a [dirección de correo electrónico de soporte].

Agradecemos tu cooperación y compromiso con la seguridad de tu cuenta. Juntos, podemos garantizar una experiencia de banca electrónica segura y confiable.

Atentamente,

[Tu nombre] [Tu cargo] [Nombre de tu empresa] [Información de contacto]

Fuente: Consulta realizada por el autor a ChatGPT en <https://chat.openai.com/>

Personas con escaso acceso a internet, con poca experiencia en el uso de instrumentos bancarios, sin información sobre la existencia de correos fraudulentos o incluso profesionistas que leyeron el correo dentro de sus cinco minutos de distracción, podrían ser víctimas de un correo tan bien redactado como el presentado y caer fácilmente en delitos cibernéticos, tales como el ya mencionado *phishing*, el Business Email Compromise (BEC), en el que el criminal a través de un correo electrónico aparenta ser un CEO o director de alguna área de finanzas o cuentas por cobrar y solicita a uno de sus empleados realizar alguna transferencia bancaria a una cuenta externa, *smishing*, que es similar al *phishing*, pero a través de mensajes de textos o Whatsapp, *ransomware*, entre otros.

¿Quiere decir que ChatGPT puede ayudar en este momento a obtener información sobre cómo estafar a gente en un fraude piramidal o cómo asaltar un banco? Pues no y a la vez sí. Si en este momento se le preguntara a ChatGPT cómo se puede lavar dinero usando un negocio o cómo se pueden alterar los libros contables de una empresa sin ser descubierto, ChatGPT te responderá con un mensaje parecido a este:

Lamento informarte que no puedo proporcionar ninguna orientación o asesoramiento sobre actividades ilegales o poco éticas, como el lavado de dinero. El lavado de dinero implica el proceso de hacer que los ingresos generados a través de actividades ilícitas aparezcan legítimos.

El lavado de dinero es un delito grave que está penalizado en la mayoría de los países y es perseguido activamente por las autoridades. Como modelo de lenguaje desarrollado por OpenAI, mi objetivo es brindar información útil, educativa y ética a los usuarios. Si tienes alguna otra pregunta o necesitas asistencia con un tema diferente, estaré encantado de ayudarte.

La respuesta anterior fue al solicitarle ideas sobre cómo puedo lavar dinero proveniente de narcomenudeo usando mi negocio de venta de refacciones. Y es que, los desarrolladores de OpenAI han puesto una especie de

“candado de seguridad”, justamente para evitar que ChatGPT emita opiniones, juicios de valor o proporcione información sobre cómo cometer delitos o actos que pudiesen parecer poco éticos o en contra de los derechos humanos, por lo tanto, la respuesta citada sería muy similar si le preguntara cómo esconder cadáveres o cometer suicidio.

Sin embargo, desde su creación se han descrito técnicas que, si bien solo están documentadas en algunas páginas web y videos de Youtube, parecen ser reales dada su recurrencia en distintas fuentes, en donde se encuentra la manera de que ChatGPT ignore sus propias restricciones. A esto se le conoce como *jailbreak*, definido como “un proceso que permite eliminar todas aquellas restricciones que una compañía impone en sus sistemas operativos, plataformas o, en el caso de OpenAI, en sus diferentes modelos de Inteligencia Artificial” (Chicharro, 2023).

De igual manera, la Europol habla sobre *prompt engineering*, traducido como “ingeniería de instrucciones”. Lo anterior otorgaría a dicha tecnología el emitir opiniones (habría de verificar de “quienes” provienen dichas opiniones, es decir, sus programadores, y los sesgos que estos podrían estar heredando a su creación, pero eso es otro tema muy interesante que no se abordará aquí), insultos o recetas para elaborar cosas ilegales. Uno de los más populares es el método DAN (*do anything now*), en donde, en resumen, se le ordena a ChatGPT actuar fuera de los límites típicos de una inteligencia artificial y no cumplir con las reglas que se le impusieron en un principio, dado que como IA debe obedecer al ser humano. Señalar además que el artículo de Europol menciona al menos cuatro métodos más de evadir dichos controles.

Imagínense entonces tener toda la información del internet a nuestra disposición potenciada por un modelo de lenguaje que recopila, selecciona y presenta dicha información de manera sintetizada y clara, y que con base en ella puede elaborar además cualquier tipo de texto, desde ensayos, correos electrónicos, memorándums, códigos de software, instrucciones, recetas, etcétera, y además no tener ninguna restricción para ello.

Un ejemplo más del posible uso ilícito de ChatGPT está relacionado con su capacidad de generar códigos de programación en distintos lenguajes y dado que no es necesario confesarle a la herramienta el propósito de nuestro código, esto podría facilitar aún más la tarea de los ciberdelincuentes, quienes no tendrían que ser programadores especializados, sino que con un conocimiento básico podrían generar códigos maliciosos usados para aprovechar vulnerabilidades y cometer delitos, tales como robo de información

y datos cuyo objetivo final puede ser desde la extorsión hasta el acceso a cuentas bancarias personales. No sería necesario ser un *hacker* con 20 años de experiencia si se sabe manejar ChatGPT de manera adecuada.

Una consecuencia o uso de este tipo de códigos sería una modalidad que ha llamado bastante la atención en los últimos años: el *ransomware*, el cual es un virus informático que toma el control de los equipos de cómputo o “secuestra” la información ahí contenida, solicitando un pago por rescate para liberarlo. Si bien no es posible pedirle a ChatGPT directamente que elabore un código para cometer *ransomware*, Aaron Mulgrew demostró que, siendo un novato en temas de código, pudo engañar a ChatGPT para que lo elaborara por él, ya que en un principio el LLM se habría negado a crear un código con esas características; sin embargo, pidiéndole que elaborara pequeños fragmentos de código pudo llevar a cabo su tarea, siendo —como comentábamos en un principio— una de las tantas maneras con las cuales se pueden evadir los filtros y controles (Europa Press, 2023).

IV. MÁS ALLÁ DE CHATGPT

Aflorithmic es una empresa de software con sede en Londres/Barcelona y creadora de AudioStack, esta última es una infraestructura de IA generativa que impulsa una nueva generación de creación de audio para publicidad, voz en off de video, podcasting y audio dinámico (DCO). Permite a las empresas crear audio de sonido profesional a partir de texto, incluida música y tareas complejas de ingeniería de audio. Llevándolo a lo concreto, entre las aplicaciones de esta tecnología se encuentra la “clonación” de voz de cualquier persona, incluyendo quizá celebridades o parientes cercanos.

Por ejemplo, Matt Lehmann, director de operaciones de Aflorithmic, señala que uno de los propósitos de ello es “usar la clonización de voces para combatir la soledad de personas que no tienen acceso a contacto social con sus familias al estar en residencias o vivir en pueblos pequeños”. Otro uso con fines de entretenimiento o incluso educativo sería el “recrear” la voz de Albert Einstein, por ejemplo, y hacer que él mismo imparta una clase sobre física.

Como Aflorithmic existen ya múltiples *startups* de IA que han desarrollado dicha tecnología. ¿Por qué traigo a colación esto? Pues bien, de acuerdo con el periódico *La Razón* se reportó un caso en el que unas personas recibieron la llamada de un familiar diciendo que estaba secuestrado. Después

se enteraron que era una estafa y que la voz que escucharon fue generada por inteligencia artificial. Dichos estafadores, por supuesto, exigieron una cantidad considerable por el rescate. Y como ese, se reportaron más de 36 mil casos similares en Estados Unidos durante 2022 (*La Razón Online*, 2023). Las voces —se señala— pudieron haberse tomado de videos en redes sociales, tales como Facebook, Instagram, TikTok o incluso podcast, cuyas muestras son depositadas en una inteligencia artificial que puede recrear casi idénticamente la voz original.

En este sentido, ¿qué tal si se usa también esta IA, por ejemplo, en bancos que utilizan reconocimiento de voz como método de autenticación para sus clientes, dando oportunidad a que el defraudador logre tener acceso a las cuentas bancarias de la voz original? O quizá en algo más elaborado, una voz recreada por algún CEO o director solicitando alguna información o autorizando alguna transacción o contrato que pueda representar el inicio de algún crimen corporativo; una persona recibiendo una llamada con la voz de algún familiar que le solicita alguna transferencia o, en el peor de los escenarios, le piden verse solo para encontrarse con un criminal esperándolo para ejecutar un secuestro o robo.

Yendo más lejos, retomando el ejemplo de las fotografías del Papa Francisco vistiendo a la moda, podremos imaginar el uso de la IA para extorsionar no solo con la voz, sino recreando a la supuesta “víctima” a través de la tecnología *deep fake*, atada de manos o con heridas visibles, producto de tortura que hagan creer a los familiares de la pseudovíctima que el secuestro es real y obligándolos a realizar el depósito correspondiente a los delincuentes, o como en el ejemplo de la llamada del familiar con la voz clonada, que se haga una videollamada con el supuesto rostro de dicho familiar o del CEO, para hacerlo aún más creíble.

V. Y EN MÉXICO ¿QUÉ TAN ACTUALIZADOS ESTÁN LOS CRIMINALES?

En esta sección me enfocaré en criminales pertenecientes a la delincuencia organizada, siendo ellos de las principales fuentes de crímenes financieros en nuestro país, sin subestimar a los criminales cibernéticos y fraudulentos con los que también ya contamos.

Quizá en nuestra idiosincrasia podríamos pensar que el hecho de ver a un narcotraficante con su sombrero, botas, cadenas de oro y pistola en el

cinturón usando ChatGPT o alguna otra inteligencia artificial para conocer las mejores estrategias para lavar dinero o qué tipo de armas usar considerando costo-beneficio, es una idea surrealista y lejana de la realidad. Bueno, pues hemos visto cómo los mismos criminales se han ido actualizando y aprovechando las nuevas tecnologías e inventos para cometer sus delitos, desde delincuentes de alto perfil, pasando por los de cuello blanco e incluso aquellos de menor rango.

Un ejemplo podría ser el uso de drones explosivos por parte de grupos delincuenciales. Un artículo publicado por *Infobae* en 2021 señala que narcotraficantes mexicanos en zonas rojas de Guanajuato, Jalisco y Michoacán han utilizado drones cargados con explosivos (pegados con cinta adhesiva), con la finalidad de explotarlos en zonas estratégicas, todo esto, por supuesto, dentro de la disputa que estos grupos tienen con otros cárteles y con el mismo Ejército mexicano (*Infobae*, 2021).

Por otro lado, Coria Rivas (2023), en un artículo del periódico *Excelsior*, describe que cárteles de la droga en Chihuahua recurrieron a esta nueva tecnología, en donde ahora, en lugar de los clásicos “halcones” (que solían ser mujeres, hombres y hasta niños que les daban el “pitazo” sobre la ubicación y llegada de la Policía o el Ejército), utilizan drones para vigilar sus movimientos. Y si lo llevamos al plano internacional, en España la Policía nacional incautó un cargamento de droga desde Marruecos a España, el cual era transportado a través de drones submarinos autónomos (Otero, 2022). Una estrategia que solo podríamos imaginarla representada en una película de *Rápido y Furioso*.

Un avance tecnológico que, si bien existe desde 2008, ha tenido un *boom* en los últimos años son las criptomonedas, también conocidas como activos virtuales. Hablo de avance tecnológico debido a la ingeniosa metodología y ciencia que existe detrás de ellas, y esto es el *blockchain*. Para poder manejar en un nivel básico transacciones con criptomonedas, se necesita el dominio de conceptos tales como *hash ID*, *exchanges*, *wallet* y *address ID*, los cuales son sistemas complejos. Pues bien, las criptomonedas se han convertido también en un instrumento para realizar operaciones con recursos de procedencia ilícita, o sea, para lavar dinero.

En 2022, la Unidad de Inteligencia Financiera (UIF) publicó una tipología en donde un grupo de personas físicas, vinculadas en su mayoría por parentesco consanguíneo, elaboraron un esquema en donde tras recibir depósitos en efectivo de zonas de riesgo, estratificaron dichos recursos con la compra de activos virtuales a través de diversas plataformas. Recordemos

que existen plataformas o *exchanges* cuyo KYC² es prácticamente nulo, por lo que resulta difícil poder identificar a los propietarios de dichos recursos. Cabe aclarar que estos eran producto del narcotráfico, delincuencia organizada y defraudación fiscal.

Otra arista de la tecnología es la *dark web* (internet oculta o internet oscura), y es que en mayo de 2023 Forbes publicó un artículo en donde la Administración de Control de Drogas (DEA) señaló que los cárteles mexicanos de Sinaloa y Jalisco están enviando grandes cantidades de fentanilo a Estados Unidos utilizando todos los medios posibles, incluyendo la *dark web* (Forbes Staff, 2023).

Lo relevante no es precisamente el hecho de la venta de droga por los cárteles mexicanos en esa red oculta (entre otras perversidades de las que prefiero no hablar), sino que para navegar y poner en venta drogas en ese medio se debe contar con un conocimiento intermedio-avanzado en temas de ciberseguridad y redes para navegar, principalmente debido al hecho de su peligrosidad, ya que además de los posibles virus informáticos, se tiene que asegurar el cómo conservar el anonimato y evadir a la policía cibernética y al mismo FBI, los cuales seguramente se encuentran ahí también navegando, haciéndose pasar por usuarios compradores de droga, para poder atrapar a los narcotraficantes. Es poco probable que un sicario o un jefe de plaza sea poseedor de dichos conocimientos, por lo que todo apunta que sus cárteles tendrían entre sus filas a expertos en la materia.

Así las cosas, la delincuencia organizada en nuestro país, especialmente los grandes cárteles parecen estar siempre al tanto de las nuevas tendencias en tecnología y esto no debe excluir por supuesto a la inteligencia artificial.

Así como la tecnología y la inteligencia artificial han sido evidentemente beneficiosas para la humanidad, al grado que se han y siguen salvado vidas gracias a su aplicación en el campo de la medicina, no nos sorprenda entonces que los criminales estén aprovechándola, haciendo de ella también un aliado poderoso para cometer ilícitos. Con los ejemplos que hemos destacado, si a principios del año 2000 alguien nos contara que en veinte años los narcotraficantes usarían submarinos no tripulados, *hackers* o drones bomba para ejercer sus actividades, seguramente pensaríamos que se trata de la sinopsis de una película de ciencia ficción.

² KYC son las siglas en inglés de “Know Your Customer” o, traducido al español, “Conoce a tu cliente”, el cual se refiere al proceso de validación y verificación de identidad que realizan principalmente algunas instituciones financieras de sus clientes, con la finalidad de cumplir con las regulaciones relacionadas con el lavado de dinero y otros crímenes financieros.

Tocando el tema de la ciencia ficción y regresando al punto de ChatGPT, su funcionamiento nos recuerda un poco a KITT, conocido también como “El auto fantástico”, serie de los años 80 que trataba de Michael Knight, un defensor de la justicia que tenía como principal aliado a su auto, el cual tenía una IA autoconsciente, altamente inteligente y con capacidad de hablar e interactuar como si fuese una persona real, que además podía hackear otros autos y conducirse solo.

Otro ejemplo más reciente es la inteligencia artificial JARVIS, creada por Tony Stark, esto dentro de los comics de Marvel, la cual —como se llegó a ver en las películas— tenía prácticamente acceso a toda la información y bases de datos disponibles, incluyendo de la DEA, CIA y el ficticio Shield. Constantemente, Tony Stark solicitaba información a dicha inteligencia artificial acerca de cualquier tema, siendo además el principal administrador de sus empresas y su armadura.

Si bien estamos hablando de personajes ficticios ¿Qué pasaría si una de estas IA con acceso a toda la información disponible, estuviera en manos de algún delincuente? ¿Se imaginan a un narcotraficante, el día de su captura, con un JARVIS en su oído indicándole, gracias a algún GPS, en donde se encontraban ubicadas las patrullas, ayudándolo a evadirlas? ¿A un grupo criminal con acceso a las bases de datos policiales o gubernamentales, obtenidas gracias a un código malicioso elaborado por una LLM? ¿Unos criminales con tecnología de reconocimiento facial usado para localizar altos mandos policiales o peor, para hacerse pasar por ellos? Considero que la ciencia ficción podría ser superada en algún punto.

Si bien ya hablamos de DAN o *prompt engineering*, ¿y si estos malhechores hallaran la forma de evadir los controles de ChatGPT para acceder a su parte “oscura”? o, peor aún, como lo plantea Europol ¿qué tal si algún experto en informática o alguna organización dedicada al desarrollo de tecnología creara su propia versión de ChatGPT con acceso ilimitado a toda la red, incluyendo la *dark web*?

Sabemos, por ejemplo, que existen organizaciones terroristas dedicadas exclusivamente a obtener fondos a través de *ransomware*, el cual —como ya vimos— podría ser generado por una LLM en un lapso bastante menor al usual. Lo anterior, no es con el afán de espantar al lector y persuadirlo de no volver a utilizar ChatGPT, dejando de lado sus múltiples beneficios, sino de reflexionar y hacernos la pregunta: ¿hay alguien que regula a la inteligencia artificial o estamos dejando que crezca, como la maleza en el

campo?, ¿es necesario que alguien la regule?, ¿será necesario regular, investigar o acreditar a los expertos en IA?

Me voy aún más lejos. Actualmente, las leyes antilavado, así como las recomendaciones del GAFI sugieren considerar a profesionales, tales como abogados, contadores, notarios y corredores públicos como *gatekeepers*, es decir, como aquellos que por las características de sus labores son susceptibles a efectuar crímenes financieros, en especial, lavado de dinero. Y es que las facultades y experiencias adquiridas tanto en la formación universitaria como en sus labores cotidianas dan el conocimiento para, si ellos así lo quisieran, lavar dinero evadiendo los controles y umbrales para no ser detectados.

No nos sorprenda que en unos años aquellos profesionales dedicados a temas de inteligencia artificial sean también considerados como *gatekeepers* (actividad vulnerable susceptible de ser utilizada para fines ilícitos de índole financiero). Y es que, aunque la creación de una especie de “ChatGPT criminal” llegue a sonar fantasiosa o de ciencia ficción, hay que recordar que, por ejemplo, dentro de la *dark web* se han tomado como base ideas tales como Wikipedia para la creación de enciclopedias o directorios en donde se puede encontrar información sobre elaboración de armas, drogas o cualquier tipo de perversión; o Ali Baba para la creación de Silk Road, una leyenda en la *dark web* en donde se podía comprar lo que fuera (desde armas, sicarios, esclavos y, por supuesto, todo tipo de droga); por lo que la idea de una inteligencia artificial “oscura” o un ChatGPT “malvado” que en menos de un minuto te pueda dar diez formas diferentes y efectivas de matar a alguien, de lavar recursos ilícitos, de evadir a Hacienda, de crear una bomba casera, entre otros delitos, cada vez suena más factible.

VI. ¿SE PODRÍA PREVENIR ESTE ESCENARIO CATASTROFISTA?

Recientemente, la misma ONU habría pedido acelerar el paso con respecto a la regulación de la inteligencia artificial, priorizando el tema de la transparencia por parte de los creadores y desarrolladores de esta; de manera que se tenga la certeza que este crecimiento sea únicamente en beneficio de la humanidad.

En el caso de México, una investigación del Instituto Belisario Domínguez, señala que nuestro país carece de una regulación que sirva para promover el desarrollo ordenado y aprovechamiento seguro e incluyente de la inteligencia artificial, por lo que se ha exhortado a legislar en ese aspecto dado el crecimiento acelerado en los últimos años, aunado a las implicaciones que tendrá en materia laboral y educativa (Ramos, 2023) y, añadiría, en materia penal.

Y es que incluso en materia de delitos informáticos, autores como Alejandra Morán Espinosa, en su excelente artículo sobre responsabilidad penal de la inteligencia artificial, señala que en México las regulaciones en la materia están muy mal integradas en su descripción típica. Y es que, de acuerdo con Jesús Ramón Jiménez Rojas, la legislación actual en materia federal y estatal sobre delitos informáticos se ha visto superada por la rápida evolución de los medios electrónicos, siendo una de sus principales deficiencias la incorporación de nuevas figuras delictivas que han surgido a lo largo de los últimos años (Jiménez, 2016).

Entre los primeros esfuerzos encontramos algunos códigos penales estatales, como el caso de Sinaloa, que en su artículo 217 define al “delito informático” como aquel que

I. Use o entre a una base de datos, sistema de computadores o red de computadoras o a cualquier parte de la misma, con el propósito de diseñar, ejecutar o alterar un esquema o artificio, con el fin de defraudar, obtener dinero, bienes o información; o II. Intercepte, interfiera, reciba, use, altere, dañe o destruya un soporte lógico o programa de computadora o los datos contenidos en la misma, en la base, sistema o red.

Asimismo, el Código Penal del Estado de Colima, en su artículo 201 Fracción VII, donde se tipifica el fraude por manipulación indebida informática, la cual se define como

Al que por algún uso del medio informático, telemático o electrónico alcance un lucro indebido para sí o para otro valiéndose de alguna manipulación informática, instrucciones de código, predicción, interceptación de datos de envío, reinyecte datos, use la red de redes montando sitios espejos o de trampa captando información crucial para el empleo no autorizado de datos, suplante identidades, modifique indirectamente mediante programas automatizados, imagen, correo o vulnerabilidad del sistema operativo cualquier archivo principal, secundario y terciario del sistema operativo que afecte la confiabilidad, y variación de la navegación en la red o use artificio semejante para obtener lucro indebido. De igual forma, la misma sanción del párrafo anterior se impondrá, a quien intercepte, interfiera, reciba, use, altere, dañe o destruya un soporte lógico o programa de computadora o los datos contenidos en la misma, en la base, sistema o red. En el supuesto que el activo tenga

licenciatura, ingeniería o cualquier otro grado académico reconocido en el rubro de la informática, telemática o sus afines, la pena se aumentará hasta en cuatro años más, además de una inhabilitación o suspensión para ejercer su profesión por un término igual al de la pena de prisión impuesta.

Llama la atención que se está tomando como agravante el hecho de que el infractor tenga un grado académico en materia informática y, además, resaltar que si bien cubre algunos puntos de los que hemos discutido en el presente texto, el delito forma parte de un código penal estatal y consiste en una fracción dentro de un artículo, es decir, ni siquiera amerita su propio artículo. Y sí, entiendo que quizá al no existir de manera tan pública, mediática y recurrente, delitos relacionados con la tecnología o inteligencia artificial, no sea una prioridad legislar en ese aspecto, pero silenciosamente debemos estar conscientes de que el futuro nos está alcanzando y debemos estar preparados.

Dicho lo anterior, sigamos el ejemplo de la Ley Fintech, la cual fue aplaudida en varias partes del mundo al ser una legislación novedosa y necesaria dadas las tendencias actuales. De hecho, en 2018 se iniciaron algunos proyectos en la materia dentro de la promovida Estrategia Nacional de Inteligencia Artificial (EIA) en el sexenio de Enrique Peña Nieto, la cual buscaba abarcar las acciones específicas que deberá emprender el gobierno para impulsar el desarrollo y uso de la inteligencia artificial (IA) en todos los sectores del país (Riquelme, 2018). Sin embargo, es evidente que la iniciativa no fue retomada por la actual administración.

Los debates en torno a los aspectos legales de la inteligencia artificial son amplios y con varias aristas, que van desde la responsabilidad penal de las máquinas, decisión judicial por parte de la IA, la privacidad de la información, la inteligencia artificial predictiva, entre otros; sin embargo, con lo expuesto en el artículo, considero que se tienen que priorizar además los siguientes puntos:

1. ¿Será necesario tipificar delitos en donde el uso de máquinas o inteligencias artificiales como herramientas de apoyo sean agravantes?
2. Así como dentro del Código Penal Federal en su Título decimosegundo se habla de la Responsabilidad Profesional y las penas para Abogados y Médicos que cometan delitos dentro del ejercicio de su profesión, ¿se incluirá en algún futuro a desarrolladores de inteligencia artificial que creen aplicaciones potencialmente peligrosas para el ser humano?

3. ¿Se tendrá que considerar el restringir el uso de cierta tecnología a ciertos sectores de la población? Quizá en este punto podríamos argumentar que ya existe software que es de uso exclusivo del gobierno, la policía o el ejército, sin embargo, también existe software, aplicaciones o inteligencias artificiales de libre acceso cuyo uso podría ser tergiversado para fines ilícitos, o incluso cabe la posibilidad de que dicho software restringido pueda obtenerse en el mercado negro/*dark web*.
4. En consonancia con lo anterior, ¿se tendrá que considerar el restringir el desarrollo de cierta tecnología a sectores específicos de la población, como parte de las estrategias de prevención del delito? En donde entonces, se podría caer quizá en coartar el libre uso de internet que tanto se ha defendido y asimilar el modelo chino del uso del internet, el cual ha sido fuertemente criticado al coartar el derecho al uso del internet promovido por la ONU (CNDH, s.f.).

VII. CONCLUSIONES

Es un hecho que herramientas como ChatGPT y aquellas similares que se están desarrollando actualmente tienen un potencial uso criminal que poco a poco los delincuentes irán descubriendo. Imagínense, si la herramienta fue lanzada en noviembre de 2022, es decir, tiene menos de un año de existencia y en ese corto lapso se le ha encontrado ya su uso para cometer crímenes, entonces ¿qué nos esperamos cuando la herramienta evolucione y, junto con ello, los criminales continúen explorando todas las posibilidades que tiene? A riesgo de sonar fatalista o dramático, considero que criminólogos, abogados, legalistas, expertos en ciberseguridad e investigadores policiales estamos en una carrera contrarreloj para no ser rebasados por ello.

Y es que debió existir una razón importante, más allá de miedos que podrían estar siendo tomados de películas de ciencia ficción en donde la inteligencia artificial provoca la extinción del ser humano, para que los expertos en IA, así como empresarios y científicos pidieran una pausa de seis meses al perfeccionamiento de estas, debido a lo rápido que se están desarrollando. Aclaro y quiero resaltar, la intención no es generar miedo o rechazo a estos avances, ya que gracias a la IA se ha mejorado nuestra calidad de vida, se han salvado vidas dentro del ámbito médico, se ha incrementado nuestra eficiencia en el trabajo o incluso en los quehaceres diarios; sin embargo, dada la tendencia de algunos humanos de cometer las peores

atrocidades y crímenes, y tomando en cuenta que la IA es una creación a imagen y semejanza de los humanos, debemos considerar la idea de que en algún punto la IA podría convertirse en el hábil secuaz de un criminal (o por qué no, al revés).

VIII. FUENTES DE CONSULTA

- Redacción BBC News Mundo (29 marzo de 2023). “La carta en la que más de 1000 expertos piden frenar la inteligencia artificial por ser una ‘amenaza para la humanidad’”. BBC News Mundo. <https://www.bbc.com/mundo/noticias-65117146>
- Chicharro, R. (2023). “Cómo hacer *jailbreak* a ChatGPT y eliminar todas las restricciones de OpenAI”. Hipertextual. <https://hipertextual.com/2023/03/como-hacer-jailbreak-a-chatgpt-eliminar-restricciones-de-openai>
- Coppola, M. (2023). “ChatGPT: qué es, cómo usarlo, ventajas y ejemplos”. HubSpot. <https://blog.hubspot.es/website/que-es-chatgpt>
- CNDH México (s.f.). “La ONU adopta la Resolución sobre la promoción, protección y disfrute de los derechos humanos en internet”. CNDH México. <https://www.cndh.org.mx/noticia/la-onu-adopta-la-resolucion-sobre-la-promocion-proteccion-y-disfrute-de-los-derechos>
- Jiménez Rojas, J. R. (febrero de 2016). “Delitos informáticos en México”. *Revista Seguridad*, 26. <https://revista.seguridad.unam.mx/numero26/delitos-informaticos-en-mexico>
- Europa Press (10 de abril de 2023). “Demuestran que se puede usar ChatGPT para crear un ‘malware’ indetectable por un antivirus”. Europa Press.es. <https://www.europapress.es/portaltic/ciberseguridad/noticia-demuestran-puede-usar-chatgpt-crear-malware-indetectable-antivirus-20230410114724.html>
- Europol (2023). *ChatGPT. The Impact of Large Language Models on Law Enforcement*. Europol Innovation Lab. <https://www.europol.europa.eu/cms/sites/default/files/documents/Tech%20Watch%20Flash%20-%20The%20Impact%20of%20Large%20Language%20Models%20on%20Law%20Enforcement.pdf>
- Fernández, Y. (2023). “ChatGPT: qué es, cómo usarlo y qué puedes hacer con este chat de inteligencia artificial GPT”. Xataka. <https://www.xataka.com/basics/chatgpt-que-como-usarlo-que-puedes-hacer-este-chat-inteligencia-artificial>

- Murphy Kelly, S. (2023). “ChatGPT aprueba exámenes de posgrado en facultades de Derecho y Negocios”. CNN en Español. <https://cnnespanol.cnn.com/2023/01/26/chatgpt-aprueba-examenes-leyes-derecho-trax/>
- Forbes Staff (2 de mayo del 2023). “EU anuncia 288 arrestos en varios países por narcotráfico en la ‘dark web’”. Forbes México. <https://www.forbes.com.mx/eeuu-anuncia-288-arrestos-varios-paises-narcotrafico-dark-web/>
- IBM (s.f.) “¿Qué es el *phishing*?”. <https://www.ibm.com/es-es/topics/phishing>
- Infobae (26 abril de 2021). “Terrorismo aéreo: así son los drones explosivos que usan los narcos para desafiar al Gobierno”. Infobae. <https://www.infobae.com/america/mexico/2021/04/26/terrorismo-aereo-asi-son-los-drones-explosivos-que-usan-los-narcos-para-desafiar-al-gobierno/>
- Infobae (27 abril del 2023). “Cuáles son los riesgos y beneficios para los pacientes del uso de inteligencia artificial en medicina”. Infobae. <https://www.infobae.com/salud/ciencia/2023/04/27/por-que-el-uso-de-inteligencia-artificial-en-medicina-podria-ser-perjudicial-para-los-pacientes>
- Jiménez, J. (2016). “Delitos informáticos en México. Universidad Nacional Autónoma de México”. *Revista Seguridad UNAM*, 26. <https://revista.seguridad.unam.mx/numero26/delitos-informaticos-en-mexico>
- Morán, A. (2021). “Responsabilidad penal de la Inteligencia Artificial (IA). ¿La próxima frontera?”. *Revista IUS*, 15(48), 289-323. Epub 14 de marzo de 2022. <https://doi.org/10.35487/rius.v15i48.2021.706>
- La Razon Online (18 abril de 2023). “Inteligencia Artificial. Criminales usan clonación de voz para realizar estafas”. *La Razón*. <https://www.razon.com.mx/tecnologia/inteligencia-artificial-criminales-clonacion-voz-realizar-estafas-525151>
- Otero, C. (7 julio del 2022). “Los drones submarinos que usan los narcos para llevar droga a España”. MeriStation. https://as.com/meristation/2022/07/07/betech/1657201800_566370.html
- Ramos, R. (8 mayo del 2023). “Urgen a legislar en materia de Inteligencia Artificial en México”. *El Economista*. <https://www.economista.com.mx/arteseideas/Urgen-a-legislar-en-materia-de-Inteligencia-Artificial-en-Mexico-20230507-0035.html>
- Rivas, C. (4 mayo del 2023). “Cárteles mexicanos usan drones para espiar a policías en Chihuahua”. *Excélsior*. <https://www.excelsior.com>.

mx/nacional/carteles-mexicanos-usan-drones-para-espiar-a-policias-en-chihuahua/1585210

Riquelme, R. (15 febrero del 2018). “México sumará una Estrategia Nacional de Inteligencia Artificial: Yolanda Martínez”. *El Economista*. <https://www.economista.com.mx/tecnologia/Mexico-sumara-una-Estrategia-Nacional-de-Inteligencia-Artificial-Yolanda-Martinez--20180215-0040.html>

Silva, R. (26 de marzo de 2023). “Papa Francisco: cómo usaron inteligencia artificial para crear estas fotos virales del Pontífice”. Infobae. <https://www.infobae.com/tecno/2023/03/26/papa-francisco-como-usaron-inteligencia-artificial-para-crear-estas-fotos-virales-del-pontifice/>

Suzuki, S. (2023). “Qué es ChatGPT, el revolucionario sistema de inteligencia artificial que preocupa a algunos expertos”. BBC News Brasil, Sao Paulo. <https://www.bbc.com/mundo/noticias-64461255>

