Cómo la inteligencia artificial altera el paisaje de las seguridades

How artificial intelligence alters landscape of securities

Omar Villota Hurtado24

omarvillota@omarvillota.net

RESUMEN

El conocimiento explícito del matemático británico Alan Turing y las ideas de la generación Beat dieron entrada a la tercera revolución industrial respaldada en la electrónica, la tecnología de información y la producción automatizada.

El ensayo “Máquina computacional e Inteligencia” del británico fue publicado en octubre de 1950 en la edición 49 de la revista Mind. Mientras que las ideas de contracultura llevaron a la ciudadanía norteamericana a marchar pacíficamente contra la guerra de Vietnam portando carteles con la frase “Free Speech”, a establecer una universidad gratis en Palo Alto donde las personas compartían y consumían lo que podía pasar y a los ingenieros y programadores que trabajaban para las empresas de electrónica y computación en Silicon Valley pretender poseer una computadora personal.

El pensamento en “los Silicon boys” por derecho propio se sustentó en un *hacker* interesado en desarrollar aparatos electrónicos y computacionales personalizados. Junto había otro tipo que entendía del negócio relacionado con programación computacional. Los dos trabajadores se fusionaron en comunidad, de manera correcta, y organizaron empresas de garaje que luego convirtieron en corporaciones globales.

[24] Docente universitario en temas de comunicación digital y gestión del conocimiento para inteligencias colectivas.

Comunicador, especialista en redes de información documental y máster en comunicación digital con más de 25 años de experticia. Investigador sobre temas de cibercultura y tecnologías aplicadas. Trabajo en actividades de gestión de conocimiento para inteligencias colectivas. Optimizo el entorno virtual con base en emprendimiento y conocimiento disciplinar mediante estrategias, creatividad e innovación de ruptura. Autor de libros y artículos académicos para varias revistas indexadas que puede leer en http://orcid.org/0000-0003-3743-3734

Aquellos años 1960 terminaron con un hombre caminando sobre la Luna. Las capacidades de la inteligencia artificial y el aprendizaje automático han progresado tan aceleradamente dentro de las máquinas que vivimos la cuarta revolución industrial, pero su desarrollo digital há permitido asimismo una gama amplia de aplicaciones de doble uso: beneficiosas para la humanidad e inconvenientes para la sociedad, según sea el componente crítico que se activa en los círculos de uso.

PALABRAS CLAVE

Agentes inteligentes automatizados, sistemas autónomos, ingeniería social, *hackeo* cívico, panorama de seguridades

SUMMARY

The explicit knowledge of the British mathematician Alan Turing and the ideas of the Beat generation gave way to the third industrial revolution backed by electronics, information technology and automated production.

The British essay "Computational Machine and Intelligence" was published in October 1950 in the 49th edition of Mind magazine. While the ideas of counterculture led the American citizens to march peacefully against the Vietnam War carrying posters with the phrase "Free Speech", to establish a free university in Palo Alto where people shared and consumed what could happen and the engineers and programmers who worked for the electronics and computer companies in Silicon Valley pretending to own a personal computer.

The thinking in "the Silicon Boys" in its own right was based on a hacker interested in developing customized electronic and computational devices. There was another guy who understood the business related to computer programming. The two workers merged in community, in the right way, and organized garage companies that they later turned into global corporations.

Those 1960s ended with a man walking on the moon. The capacities of artificial intelligence and machine learning have progressed so fast within the machines that we lived the fourth industrial revolution, but its digital development has also allowed a wide range of dual-use applications: beneficial for humanity and disadvantages for society, depending on the critical component that is activated in the circles of use.

KEYWORDS

Automated intelligent agents, autonomous systems, social engineering, civic hacking, security overview

INTRODUCCIÓN

El presente análisis al problema del *hackeo* a la democracia se relaciona con la alteración del paisaje de las seguridades ocasionado por los sistemas de inteligencia artificial que introducen nuevos daños y expanden los existentes. La sociedad del conocimiento se enfrenta entonces a tres convulsiones como consecuencia de dicha transformación: la física, de las máquinas inteligentes; la digital, que programa los sistemas de inteligencia artificial y el equilibrio de poder, visto desde múltiples intersecciones entre la tecnología de información y la esfera política.

El razonamiento revisa el contexto formado por el activismo de los programadores, la ingeniería social como modo de *hackeo* y las actividades de usuario dentro de las denominadas “red limpia” y “red oscura”. Dentro de los tres entornos se está exacerbando el miedo, que proviene de los temores presenciales por la creciente desafección ciudadana ante los casos de corrupción, los injustificables y excesivos incumplimientos de promesas institucionales, las crisis económicas nacionales, entre otros asuntos. La realidad virtual, por tanto, es propicia para ser aprovechada por delincuentes quienes atrapan víctimas en Internet con base en ataques sentimentales; texto,

imagen y voz artificiales; esquemas Ponzi (redes piramidales); suplantación de identidad mediante correos electrónicos fraudulentos; estafas a través de correo electrónico; comunicaciones dirigidas a personas específicas para que instalen *malware*; estrategias engañosas de mercadeo; etc.

Fundamentación del problema

Desde que Turing propuso en su escrito la pregunta ¿pueden pensar las máquinas?, los agentes inteligentes artificiales se han diseñado con base en la idea que el matemático británico comprendía de la computadora: "una máquina cuyo objetivo es ejecutar cualquier operación que pueda realizar una computadora humana"[25]. El proceso computacional está fijado en seguir reglas determinadas y omitir la autoridad que desvíe la tarea en un mínimo detalle. Por su parte, la inteligencia artificial (IA) busca como objetivo "automatizar una amplia gama de tareas donde humanos o animales no humanos utilicen su inteligencia"[26].

En el último lustro, los mejores sistemas de inteligencia artificial que reconocen imágenes alcanzan un rendimiento correctamente mejorado cercano a 98%, mucho mejor que el punto de referencia humano que es 95% de precisión. Otros de aquellos sistemas producen imágenes artificiales que son casi indistinguibles a partir de fotografías. Esta típica inteligencia artificial manifiesta dos propiedades, que son:

a) Eficiencia, en doble perspectiva: en la tarea, si un sistema entrenado excede comúnmente el desempeño humano, y en el entrenamiento medido, según la cantidad de tiempo, recurso computacional y datos que requiere para aprender a desempeñarse bien en la tarea. Por tanto, un sistema de IA es eficiente si, una vez entrenado y desarrollado, puede completar una tarea de forma más rápida o económica que un humano; y

b) Escalabilidad si, dado que puede concluir una determinada tarea, aumenta la potencia informática a la que tiene acceso o hace copias del sistema con la que se permitiría completar muchas más instancias para realizar la misma tarea.

[25] Turing, 1950, pág. 50.
[26] Brundage & et al., 2018, pág. 12.

Para llevar a cabo la mayoría de nuestras tareas, los humanos superamos significativamente a los sistemas de inteligencia artificial en términos de eficiencia en la capacitación, pero las máquinas inteligentes nos superan en la eficiencia, en la tarea. La investigación en el campo de la IA arguye, por ejemplo, que el reconocimiento de imágenes se puede aplicar a muchas tomas de cámaras diferentes por mucho menos valor al comparar el costo por contratación de analistas humanos para ejecutar ese trabajo equivalente. En particular, la producción de sistemas eficientes puede aumentar la cantidad de actores que accedan perpetrar ataques individuales debido a la facilidad de expansión de los sistemas de IA. Factores que ayudan a explicar esta ganancia reciente incluyen: i) crecimiento exponencial de la informática; ii) algoritmos mejorados destinados al aprendizaje automático, especialmente en el área de redes neuronales profundas; iii) desarrollo de software con marcos estándar para una iteración más rápida; iv) réplica de los experimentos tecno científicos; v) conjuntos de datos disponibles cada vez más grandes y más amplios; e vi) inversiones comerciales globales[27].

Turing razonaba que las máquinas aprenden ciertas operaciones mecánicas una vez fueran previamente consideradas las funciones del cerebro humano y animal. Las denominó funciones "subcríticas", donde una idea presentada proporciona por lo menos una idea como respuesta. En los sistemas de inteligencia artificial, dicha función es fácil de difundir por ser eficiente y escalable. El ejemplo prototipo es la amenaza de los ataques *spear phishing*[28]. Para estas agresiones se requiere cantidad significativa de mano de obra calificada, ya que el atacante debe identificar el alto valor adecuado del objetivo, investigar las redes sociales y profesionales por donde se relaciona la potencial víctima y, posteriormente, generar los mensajes que sean plausibles a su contexto. Máquinas programadas para un aprendizaje automático generan dicha inteligencia -de escribir mensajes personalizados y presentarse como amigo, colega o contacto

[27] Jordan & Mitchell, 2015.

[28] Un ataque *spear phishing* es una estafa a través de correo electrónico o comunicación instantánea dirigida a una persona, de organización o empresa concreta, personalizando la fachada del ataque de forma relevante o confiable, con el objetivo de instalar *malware* en la computadora de la víctima y, a menudo, manipular datos para fines malintencionados. Se diferencia del ataque *phishing* que consiste en intentar iniciar una acción para robar información de la víctima potencial mediante el engaño, con base en una fachada superficialmente confiable.

profesional- con el propósito premeditado de extraer información sensible o dinero a víctimas individuales.

Sobresale un segundo tipo de funciones que Turing designó "supercríticas", según la cual una idea presentada podría dar origen a toda una teoría de ideas secundarias y tradicionales. Sin embargo, su cálculo teórico debía esperar hasta el fin del siglo 20 para emprender el experimento ya que, principalmente, el problema central se relacionaba con la insuficiente programación en aquella década de 1950.

Los actuales sistemas de IA sobrellevan un encadenamiento de invenciones pasadas y algunos no resuelven las diferentes vulnerabilidades que se han expandido por fuera del software. Por el contrario, demuestran que pueden exceder desempeños humanos de muchas maneras como también fallar del modo en que un humano jamás lo haría. Son ataques aplicados para el "envenenamiento de datos" (introducción de datos en la fase de entrenamiento para que el aprendizaje automático cometa errores de sistema), "ejemplos adversos" (diseño de insumos en el sistema para ser clasificados erróneamente por el aprendizaje automático) y "explotación de imperfecciones" (errores en el planteo de las finalidades de los sistemas autónomos).

Frente a las capacidades todavía limitadas de la IA, expongo dos hipótesis para sustentar el análisis al problema del *hackeo* a la democracia. H1: el aprovechamiento de los sistemas de inteligencia artificial ocasiona expansión de las amenazas existentes con base en nuevas técnicas y exitosas estrategias de innovación que enseñan a las máquinas obtención de recompensas auxiliares, como el control de funciones y otro conjunto de manifestaciones humanas. H2: el fortalecimiento en la investigación de la inteligencia artificial suscita la introducción de nuevas amenazas traspasada la capacidad de reconocimiento artificial de imágenes, texto y audio usado para alterar la identidad digital en línea o para influir a la opinión pública mediante la distribución de contenido generado por *bots*[29] a través de los medios sociales.

[29] Es una aplicación que ejecuta tareas automatizadas como activar una alarma, informar del clima, efectuar búsquedas *online*. Desde el desarrollo del software existen creaciones de Microsoft como "Clippy" y de AOL Instant

Contexto

Con base en las propiedades generales de la inteligencia artificial –eficiencia y escalabilidad- las investigaciones en este campo han llevado a predecir que: i) ataques muy específicos serán difíciles de atribuir debido a la singularidad de que aumentan el anonimato del atacante y, que ii) ataques que aprovechan las vulnerabilidades no resueltas se volverán más típicas por cuanto amplían la probabilidad de ser cada vez más generalizadas. Otros indicios de investigadores[30] conducen a razonar que:

a) Ataques altamente efectivos y sin medidas preventivas sustanciales se volverán más típicos por causa de dos factores externos: i) las habilidades humanas de pensamiento superior y ii) la conducta del agresor al frecuentemente desafiar la compensación entre la repetición y la escala de sus ataques, por un lado, y la efectividad, por el otro. Por ejemplo, ataques *phishing* muy genéricos logran ser más rentables a pesar de las desmesuradas tasas bajas de éxito, simplemente en virtud de su escala. A pesar de lo anterior, hay optimismo en los sistemas de IA para que tales concesiones sean menos agudas en el futuro;

b) Ataques dirigidos con precisión serán más frecuentes con base en la dimensión del contexto relacionado con identificación y análisis a objetivos potenciales. El motivo obedece a que los atacantes repetidamente enfrentan la compensación entre cuán eficientes y escalables son sus ataques y qué tan bien dirigidos están en estos aspectos. Un ejemplo podría ser el uso de enjambres de vehículos aéreos no tripulados (UAV) que implementan tecnología de reconocimiento facial para matar a individuos señalados entre la multitud, en lugar de usar formas de violencia menos selectivas.

Expongo a continuación algunas prácticas que visionan las características de las condiciones de uso de las máquinas inteligentes y su aprendizaje automático, orientado a argumentar las hipótesis planteadas con anterioridad.

1. Activismo

Messenger como “Smarter Child”. Con la IA, la mayoría de *bots* están programados para actuar como humanos cuando se interactúa con ellos, como el caso de “Siri” elaborado por Apple o “Cortana”, el asistente virtual creado por Microsoft (Mitroff, 2016).

[30] Brundage & et al., 2018, págs. 12-18.

La meditación existencial de los pioneros desarrolladores de software buscaba en la pregunta ¿qué esperas obtener en tu vida? Siendo así, el activismo de aquellos programadores se orientó a varias vanguardias: i) obsequiaron sus programas de computación porque el entusiasmo y la práctica con la electrónica los superaba como fieles aficionados por una técnica en ciernes. Adicional, contribuyó que la legislación norteamericana carecía de reglamentación para patentar el software; ii) cobraban por el trabajo inmaterial que realizaban en el que invertían poco tiempo asumido como pasatiempo; iii) nunca pensaron ser famosos ni alcanzar el éxito con base en esquemas publicitarios inadvertidos.

De otro lado, las compañías informáticas reunidas en la primera feria de computación de la Costa Oeste americana en 1977, no se dieron cuenta que los sistemas informáticos para consumo dominarían la dinámica empresarial. Esta especial circunstancia ya había sido interiorizada por algunos de los activistas del Silicon Valley quienes empezaron a aprender sobre nuevas estrategias de mercado y a la postre, resultaron visionarios. Otra avanzada que contribuyó en la traslación de la cultura activista para la década de 1980 se fundamentó en la lucha de dos fuerzas: la tradicional, opresiva, conservadora, de las compañías informáticas y de computación y la de libertad de los programadores con base en su activismo por el software.
Desde su génesis, el uso de las herramientas lógicas ha proporcionado a su usuario experimentar a partir de su inteligencia natural individual, cruzando muchas tareas que implican realizar operaciones interactivas, comunicarse con otras personas, redistribuir copias, observar, ser observado, y, sobre todo, acentuar el anonimato. En general, el software permite a la gente y a grupos de técnicos, mayormente, ejecutar trabajo inmaterial bien hecho, reiterativo y preciso. De manera que con esta “herramienta en la mente” se explora en más de un nivel las realidades sociales y se renueva la modalidad del trabajo presencial centrado en una fábrica de producción serial, por otro de apariencia virtual y movilidad corporal para la gestión de información y el conocimiento.

En el proceso del desarrollo de programación del software actual se están implantando nuevas amenazas fundamentadas en la investigación de la IA que explotan una variedad de

vulnerabilidades. Las humanas por ejemplo, mediante suplantación desde software sintetizadores de voz y las comunes de software, que aún realizan denegación de servicios informáticos[31]. Una segunda agrupación de vulnerabilidades no resueltas se encuentra localizada dentro de los sistemas de IA que se exterioriza por medio de la categoría "ejemplos adversos". Verbigracia, una modificación simple y controlada a algunos pixeles de una señal digital de *Stop* crearía oportunidades para colisiones entre vehículos autónomos. Un peor escenario puede estar implantado en la clase "envenenamiento de datos" por ejemplo, un ataque a un servidor utilizado por el Estado como central financiera acarrearía la manipulación de miles de datos de clientes y en caso de no detectarse, continuará en el tiempo dejando inmenso daño a la economía nacional y a la reputación de la entidad financiera[32].

Así mismo se presenta expansión de amenazas existentes a través del aprendizaje automático, ya que las máquinas pueden aprender a tomar decisiones para que respondan al comportamiento humano o para la ejecución de tareas automatizadas observando a las personas presentes. Aquellos softwares programados para sistemas de IA permiten a los actores la conservación de su anonimato y experimentar mayor grado de distancia psicológica frente a las personas que impactan. Por ejemplo, si alguien usa armas letales autónomas para llevar a cabo un asesinato, en lugar de un arma de fuego convencional, el francotirador evitará la necesidad de estar presente en la escena y mirar a su víctima[33]. Esta importancia de asumir distancia psicológica evita la recaída por el estrés postraumático del trabajo en la guerra[34], incluso para los cibercriminales quienes se benefician del complejo de las vulnerabilidades.

En síntesis, el espíritu de libertad en los programadores de software por su creativo trabajo especializado no material parecería haber sido recogido por John Rawls para la formulación del primer principio de su teoría de la justicia y liberalismo político aplicable a los individuos, según

[31] Lomas, 2017.
[32] Scharre, 2016.
[33] Cummings & et al., 2018.
[34] Allden & Murakami, 2015.

el cual "la sociedad se organizará de tal manera que todos los miembros disfruten de la mayor igualdad de libertades posibles, incluida la justa igualdad de oportunidades"[35].

2. Ingeniería social

La ingeniería social apunta de manera perversa a efectuar estafas mediante la gestión de intromisión según actúen las personas conforme sus impulsos de curiosidad, emoción o apetencia sexual. A las corporaciones se las ataca luego de averiguar por las debilidades informáticas de empleados obteniendo así dinero y robo de valiosos datos que provienen de investigación, documentos estratégicos, planes de expansión, propuestas de expertos. Un tercer grupo víctima es el formado por agencias de gobierno, para dirigir acciones de sabotaje y terrorismo contra el Estado. Sin duda, las modalidades que los agresores practican con base en estrategias de ingeniería social para el lucro recurren al humano por ser el eslabón más débil en la cadena de la estafa. Las destrezas digitales de aquellos oscilan entre falsear la identidad virtual, embestir con el envío de contenido e infectar sistemas y servidores informáticos. De manera que las estafas cibernéticas por medio de estrategias de ingeniería social son tan diversas que ocasionan amenazas confusas cuyas consecuencias de todas maneras son desmesuradamente devastadoras.

La IA es explotada por bandas organizadas de cibercriminales y *hacker* "sombrero negro" con base en agentes inteligentes como software para el reconocimiento de texto y voz artificial que alcanza a víctimas determinadas y por otros agentes inteligentes automatizados como *bots* para ejecutar ataques de mayor escala. Todas las calidades se distribuyen desde la comodidad de la web a través de portales para citas amorosas, correos electrónicos fraudulentos, llamadas falsas de *Help Desk* o simplemente, abandonando en lugares tácticos y predeterminados de las oficinas dispositivos de almacenamiento digital. Últimamente las empresas dedicadas a la ciberseguridad evidencian que ese mercado está en creciente alza económica, pero igualmente algunos de sus servicios de seguridad declinan al ser también víctimas corporativas de los ataques globales.

[35] Schultz, 2012, pág. 142.

Por el contrario, *hacker* "sombrero blanco" tienen como propósito no ser malvados y por potencia emprendedora dos acciones: do it yourself (hacerlo por mí mismo) y recolectar circuitos e intercambiar información destinados a la construcción de hardware y software. Estos principios siguen activos, pero con planes de ingeniería social perversa se mutan a procedimientos con el fin de "manipular psicológicamente a los individuos para que proporcionen información sensible privada y acaben siendo víctimas"[36]. Para los ingenieros sociales el sentimiento de curiosidad en la víctima eventual es la previsión perfecta para concretar con ella los ataques.

Perfiladores de cibercriminales afirman que el impulso por el cual las previsibles víctimas se despeñan con trucos en línea obedece a factores psicológicos que favorecen el automatismo de los ataques en Internet. María Bada, investigadora senior del Centro de Cibercrimen de Cambridge (Reino Unido), asegura que "las personas solitarias acceden a Internet y por fin encuentran a alguien con quien pueda imaginar una relación, alguien en quien confiar"[37]. Brad Sagarin, profesor de tiempo completo en el Departamento de Psicología de Northern Illinois University (Estados Unidos), sostiene que "el ingeniero social emplea las mismas técnicas de persuasión que utilizamos a diario todos los demás, pero de forma manipuladora, engañosa y con muy poca ética. Adquirimos normas. Intentamos ganar credibilidad. Exigimos obligaciones recíprocas"[38]. Uno de los hackers más famosos es el norteamericano Kevin David Mitnick apodado Cóndor y por el mismo Fantasma de los cables. Comenzó su intrusión en redes de computadora mediante el uso de teléfonos a los 16 años cuando quebró la seguridad del sistema informático administrativo de su colegio solo para curiosear. Considera que más allá de las técnicas de seguridad que se pueden implementar en hardware y en software, el factor determinante de la ciberseguridad está directamente relacionado con la capacidad de los usuarios de tecnología para interpretar correctamente las políticas de seguridad y hacerlas

[36] Gómez Blanco, 2018.
[37] Wagner, 2018.
[38] Mitnick & Simon, 2007, pág. 313.

cumplir. Como ingeniero social asaltante ha comprendido a los 55 años de edad que "el atacante diestro en el engaño se alimenta de las mejores cualidades de la naturaleza humana: nuestra innata tendencia de ayuda y de apoyo, de ser agradable, colaborador y cumplidor en una gestión". Recomienda que la mejor "defensa inteligente es comprender las metodologías que utilizan los ciber adversarios"[39].

Responsables en ciberseguridad afirman que la conducta repetitiva se convierte en hábito inconsciente en el usuario de tecnología cuando se encuentra en actividad frente al teclado -ya sea de la computadora personal o del teléfono celular-. Esa constante se complementa con ignorar que lo aparecido en las pantallas es realidad virtual, o sea una proyección del mundo verdadero. Por ejemplo, la mayoría de ataques en línea provienen de información personalizada previamente recogida y adaptada para actuar sobre la propia víctima mediante sitios web encubiertos, contenidos de correos electrónicos con supuesta confidencialidad, envío de hipervínculos maliciosos en algún mensaje de la comunicación digital. Probablemente en alguna de estas tácticas el navegante hará *clics* por la confianza que encubre la misión.

Las pandillas entonces utilizan la ingeniería social para exasperar estados individuales como soledad, auto-comparación excesiva, déficit de asertividad relacionado con ansiedad o depresión, trastorno de adicción a Internet, ausencia de censura, y así acometer con el engaño a personas, empresas y gobiernos. La realidad virtual es propicia para mitigar estos estados de personalidad afectada lo cual hace que la potencial víctima acabe entablando con desconocidos comunicación incorpórea, ubicua, permanente. Considero que lo favorable de las redes sociales para estos estados de bajo ánimo personal estriba en el exiguo costo invertido y el profundo "beneficio" obtenido.

Un segundo determinante común es recuperar la auto-confianza respondiendo todo mensaje recibido, aun el de algún desconocido. La comunicación escrita alegra debido a los anónimos mensajes sentimentales, la compañía virtual hace feliz por un momento, los dispositivos exaltan

[39] Mitnick & Simon, 2007, pág. 299.

la euforia personal por la empatía perdida en el entorno presencial, son algunas justificaciones de los consumidores de las tecnologías digitales. Sin embargo, una persona con su autoestima lesionada no percibe previamente la astucia del estafador. Genera rápida empatía y responde enseguida, luego esa comunicación se volverá un bombardeo de anuncios que incluye a todos los dispositivos y aplicaciones posibles que la persona afectada compromete. El síndrome de Estocolmo se ha introducido a Internet ya que por aquel sentimiento de identificación tan fuerte algunas víctimas atacadas siguen confiando en su interlocutor virtual. Para Lee Felsenstein, primer desarrollador de una red social, el objeto de instalar su creación, "Community Memory", dentro del recinto del Homebrew Computer Club, aquella noche del 5 de marzo de 1975, era hacer que las personas se conocieran entre sí, cara a cara, pues el "comportamiento de la gente mediado por pantallas es muy diferente al que tendrían en una conversación directa"[40].

El método genérico de ataques atravesado por la ingeniería social perversa es despertar impulsos sensibles en la víctima potencial y aprovecharse de sus temores. Existen estafadores sentimentales localizados en África. Son individuos que falsean su identidad en salas web para enamorados y mediante envíos de fotos falsas y llamadas de teléfono móvil se presentan como atractivos solteros europeos adinerados. El trabajo de ingeniería social a sus víctimas es persistente, consistente y constante hasta que la confianza se pone a prueba y resulta a su favor. Por ejemplo, las víctimas son especialmente mujeres mayores de 50 años, amas de casa, que en sus mensajes de presentación no ocultan su autoestima maltratada. La mayoría de ellas termina transfiriendo dinero para solventar de improviso una ayuda requerida por el "galán" o suministrando datos personales de la cuenta bancaria que las pandillas pronto disponen como estación de tránsito para lavar activos ilícitos. Otra práctica de asalto comprobado es la recepción de mensajes instantáneos con suplantación de identidad corporativa, cuyos contenidos generan sobresalto en el usuario porque su dinero en el banco estaría en peligro si no activa el enlace para actualizar los datos personales del portafolio bancario. En realidad, son

[40] Tenhaven, 2017.

enlaces a sitios web fraudulentos donde los criminales capturan datos privados y claves de autenticación financiera de sus víctimas.

A medida que la IA se desarrolla, aplicaciones geo sociales con *chat bots*[41] son más convincentes, lo que puede generar mayor confianza al involucrar a las personas en diálogos más extensos y, quizá, eventualmente con identidades virtuales que son elementales disfraces visuales creadas para el momento de la video charla. Por ahora, la mayoría de enlaces que se reciben en esas plataformas sentimentales dirigen a sitios web re-direccionados 3 o 4 veces que contienen la advertencia: "Verás fotos de desnudos. Sé discreto, por favor" o solicitan entablar conversación más privada por Skype con perfiles falsos bajo la excusa engañosa de un intercambio de fotos, pero al final terminan solicitando datos privados financieros a la víctima.

Con base en la autonomía de las máquinas, el estilo de escritura humano está siendo imitado con inteligencia artificial lo que beneficia a algunos perpetradores para ataques a escala y planeados a través de ingeniería social. Además, es común usar *bots* para que empleados reciban *malware* en archivos de falsos colegas del área de soporte informático para que los instalen como parches de seguridad o de actualización de software. De manera que es el propio usuario quien instala en la terminal el programa malicioso, siendo suficiente para que el agresor obtenga total control del equipo y de su víctima. Con la ingeniería social las bandas criminales también tienen como víctimas a las corporaciones, según le ocurrió así al grupo empresarial Leoni AG de Nuremberg (Alemania), el mayor fabricante de cables electrónicos y arneses de Europa. En 2016, para desencriptar sus datos estratégicos corporativos capturados pagó 40 millones de euros tras ser atacado con falsos mensajes de correo que provenían de una dirección corporativa enmascarada. Igualmente, la ingeniería social se instaló al servicio del espionaje a gobiernos, y de nuevo la leyenda "Freedom of Speech" se expuso en las calles en el año 2013 por el caso de Edward Snowden, el exempleado de la CIA y la NSA, quien entregó documentación confidencial a los periódicos The Guardian y The Washington Post. Estos

[41] Programa de software que simula mantener una conversación con una persona al proveer respuestas automáticas a entradas hechas por el usuario humano (Drozhzhin, 2017).

publicaron hojas clasificadas remitidas por Snowden sobre distintos programas de vigilancia masiva de la NSA y el FBI, incluyendo PRISM[42], XKeyscore[43] y otra información recabada directamente de servidores de Microsoft, Yahoo, Google, Facebook, PalTalk, AOL, Skype, YouTube y Apple. Una vez más la sociedad está vigilada mediante Internet y con base en el comportamiento individual del usuario, quien regala sus datos privados a cambio de cualquier información gratis.

El caso más reciente de ingeniería social que desata un nuevo nivel de polarización en la sociedad red, en especial la norteamericana, es la configuración de plataformas web para encontrar la compañía ideal o unir emocionalmente a los que huyen del populismo y las dictaduras, a pesar de que ya existían para grupos concretos como jubilados, latinos, veganos. Desde que Donald Trump fue elegido presidente de Estados Unidos, se han configurado sitios web amorosos según la filiación política considerada fundamental en los sentimientos. La nueva ola de sitios de citas virtuales, por ejemplo, es el portal de David Gross, fundador de Trump Singles, que enlaza enamorados con la leyenda "haciendo citas geniales de nuevo". Sostiene que la "polarización existe y buscamos ayudar a esa gente afectada por la división social para encontrar el amor. Tenemos más demanda en San Francisco, Filadelfia y Nueva York porque si eres republicano y seguidor de Trump no te va a costar." En Canadá, Joe Goldman, también ha creado una plataforma de amistad y negocios con la orientación política opuesta a Trump. Justamente su lanza de mercado Maple Match se vende bajo el lema "haz que las citas sean geniales de nuevo" y se presenta como un "facilitador para que los estadounidenses encuentren al socio canadiense ideal para salvarlos del horror insondable de la presidencia de Trump".

[42] PRISM es un programa clandestino para vigilancia electrónica operado por la Agencia de Seguridad Nacional (NSA) de Estados Unidos que recoge datos y comunicación masiva procedentes de grandes compañías norteamericanas de Internet. El programa secreto fue liberado en 2007, en el marco de expansión de los servicios de inteligencia de Estados Unidos iniciado en 2001 tras los atentados del 11 de septiembre y el comienzo de la guerra contra el terrorismo.

[43] XKeyscore es un sistema informático secreto utilizado por la Agencia de Seguridad Nacional (NSA) de Estados Unidos para la búsqueda y análisis de datos en Internet. El programa se ejecuta en forma conjunta con otros organismos internacionales como la Dirección de Señales de Defensa de Australia y la Oficina de Seguridad de Comunicaciones del Gobierno de Nueva Zelanda.

Con el ataque terrorista del 11-S a las Torres Gemelas de Nueva York, el desafío volvió a escena en el sentido de mantener la sociedad abierta con libertad de acción para los individuos o darle paso a una sociedad basada en la sospecha e intereses de las empresas o de las instituciones. Por su parte, el segundo principio de justicia de Rawls aplicado a los participantes de la economía mundial es el de la diferencia, debido a que "las desigualdades económicas en la sociedad han de estar estructuradas de manera tal que aseguren más beneficio a los menos favorecidos"[44].

3. Internet y Darknet

Una nueva generación de *hacker* "sombrero negro" está ganando mucho dinero a través de Internet debido a los ataques a infraestructuras de servicios y a la seguridad nacional de los países. Estas agresiones están afectando aún más la psicología social mediante el engaño agresivo a personal calificado. Para alcanzar sus fines los atacantes aplican tres métodos indistintos, por lo general:

a) Ataques *Advanced Persistent Threat* (APT) fue una acción militar de un mes de duración iniciada por militares de China contra diferentes medios de comunicación a través de estrategias *spear phishing* y *malware*.

b) Ataques *Botnet* a cantidad de redes de computadora con una variedad de *malware* para infección y ataques DDoS (son las denegaciones de servicio) simultáneos lo que genera amplia amenaza en la estabilidad e integridad de Internet.

c) Ataques *Ransomware* dirigidos a empresas, instituciones públicas y usuarios particulares para encriptar los archivos guardados en la computadora de la víctima y posteriormente exigir un pago por el rescate y la liberación de los archivos cifrados.

Los inesperados *hackers* realizan combinaciones de estas acciones intrépidas como las cometidas por Mustafa Al-Bassam, alias Tflow, un "sombrero negro" miembro principal del grupo LulzSec, quien durante 50 días atacó en las manifestaciones del mundo árabe de 2011. En su diversión informática maliciosa descubrió el falso paisaje de la seguridad informática y exigió mayor transparencia en las actividades de algunas de las compañías atacadas. Los medios de comunicación recogieron las palabras de su defensa y entonces publicaron que "penetramos

[44] Schultz, 2012, pág. 143.

correos electrónicos y bases de datos para demostrar lo interno de las corporaciones, de qué procedimientos son culpables, pues es extremadamente importante sacar la verdad de la empresa corrupta para que la gente tome sus propias decisiones."

Alpha Barry, jefe de estrategias, gobernanza y seguridad en Thyssenkrupp (Colonia), la industria siderúrgica alemana más importante, sostiene que "los nuevos ataques son muy sofisticados, lo que indica que no son producidos por un delincuente sino por organizaciones criminales interesadas que intentan cambiar al modo industrial, usando herramientas profesionales compradas en el mercado negro y programadas por ellos mismos"[45].

Los ciberataques a la industria norteamericana del año 2010, en especial para robar propiedad industrial y dominar el mundo del mercado global, provinieron de ejércitos de maleantes de China. De allí que Barak Obama en su presidencia, anunciara públicamente ante el Congreso que "los países y las compañías extranjeros nos roban secretos corporativos". Sin embargo, fue Google la primera compañía americana en revelar a los autores de los ataques cibernéticos, al publicar en su blog corporativo que "detectamos ataques altamente sofisticados dirigidos a nuestra infraestructura corporativa, provenientes originalmente de China, cuyos robos se generaron en la propiedad intelectual de Google". Se rumora desde entonces que los hurtos fueron a buena parte del código fuente del gigante tecnológico.

Luego del ciberataque a Google y de analizar cientos de amenazas ofensivas a conglomerados globales de diferentes sectores norteamericanos, Dimitri Alperovitch, cofundador de Crowd Strike, una compañía estadounidense con sede en Sunnyvale (California) especializada en ciberseguridad, llegó a la conclusión que el grupo Elderwood, con sede en Beijing y vínculos con el Ejército Popular de Liberación de China, organizó la "Operación Aurora" para llevar a cabo APT a lo largo del segundo semestre del año 2009 a corporaciones de la lista Fortune 500. Precisa que:

[45] Wagner, 2018.

> Concluimos que fue China por el rastreo a las máquinas involucradas y programadas para el ataque y el control. Los ataques fueron producto del grupo Aurora Panda, del gobierno de China, que conduce a sus miembros a realizar espionaje a variadas organizaciones, agencias gubernamentales y compañías privadas de todo el mundo. Por lo general, son miembros militares del Ejército Popular de Liberación, cuya orden es de tipo militar, así que si sus miembros son descubiertos y expulsados del territorio que espían, retornan para concluir su misión[46].

China básicamente transformó su visión económica clave de dejar de ser maquilador mundial para convertirse en innovador global. Cada 5 años se ve una nueva industria de China pretendiendo sobresalir, pero, al mismo tiempo, se observan ciberataques a una empresa en concreto que sobresale en Europa o América. La investigación de seguridad informática mundial que lleva a cabo Nicole Perlroth, de The New York Times, tiene comprobación que los ataques cibernéticos que genera China están destinados a robar fórmulas de pintura, estrategias de negocio, ideas de diplomáticos y profesores universitarios, contratos de sociedades de abogados, etc.

También se habría encontrado que Israel fue el país que dirigió el ataque con el gusano informático Stuxnet, para infectar en 2010 a varios miles de computadoras del mundo, la mayoría localizados en Irán. La intromisión llegó hasta el control de los motores de casi mil centrifugadoras con las que ese país del golfo Pérsico enriquece Uranio en la planta nuclear de Natanz, al sur de Teherán. Israel ha sido el primer país del globo en aseverar que Internet, además de una red de comunicación global, es una amenaza mundial a las seguridades nacionales. Por consiguiente, ha creado un fuerte aparato en ciberseguridad para lo militar, lo económico y lo investigativo. Al sur de Jerusalem, en Beerseba, se localiza el campo Advanced Technologies Park, el Silicon Valley del desierto de Néguev, donde una masa crítica, nuevamente, formada por científicos, militares, académicos e industriales trabaja conjuntamente en procura de la ciberseguridad nacional. Adicional, el sistema educativo israelí ha creado el servicio militar en informática para los estudiantes de las escuelas cuyo fin es

[46] Makuch, 2016-2017.

enfrentar una guerra cibernética. Este reclutamiento voluntario está a cargo de la unidad militar de elite 8200. Para el director del Centro de Investigación y Seguridad Cibernética de Israel, Yuval Elovici, "no somos tan ingenuos para subestimar los ciberataques, ya que según nuestras investigaciones sabemos perfectamente de lo que somos capaces, incluidos otros actores diferentes a Israel"[47]. Muchos altos ejecutivos de las compañías de Europa, en especial de Alemania, han afirmado que viajan continuamente hasta Israel para adquirir nuevas soluciones en ciberseguridad, principalmente, emprendimientos y experiencias técnicas para proteger las propias redes o revender el servicio a otros.

Nuevos países también están comprometidos con los ciberataques a redes nacionales de infraestructura de servicios. Por ejemplo, para la navidad de 2015, de manera repentina, todo el oeste de Ucrania sufrió un apagón en su sistema de energía eléctrica. Durante meses los asaltantes ingresaron furtivamente a la red informática del operador eléctrico hasta que finalmente lograron separar de la red nacional a varias centrales con transformadores. De la invasión se culpó a bandas de ataque informático de Rusia. En 2017, más de 360 mil sistemas informáticos de 180 países que operan con Windows en hospitales, ferrocarriles, fábricas, quedaron bloqueados por el *Ransomware*, Wanna Cry. Los atacantes solo recibieron US $130 mil, pero todo el sistema se perjudicó con estimaciones que llegaron a los 4 mil millones de euros. Según el último informe de Kaspersky Lab, correspondiente al tercer trimestre de 2018, todavía hay muchas computadoras en el mundo que no han parcheado esa vulnerabilidad de Windows. Finalmente, para finales del 2018, funcionarios de seguridad del gobierno de Alemania detectaron nuevos y variados ataques cibernéticos a las cuentas de correo electrónico de legisladores, ministerios, militares y varias embajadas alemanas. Aun cuando se desconocen los robos, el gobierno de Angela Merkel llegó a culpar al grupo *hacker* ruso Snake, un servicio del Departamento Central de Inteligencia Militar (GRU).

Darknet "son capas de aplicación y protocolo que se encuentra en las redes existentes que facilitan una colección de redes y tecnologías para compartir contenido digital"[48]. Por ejemplo,

[47] Wagner, 2018.
[48] Biddle & et al., 2002, pág. 1.

red oscura es el intercambio de archivos de igual a igual, la copia de CD y DVD y el cambio recíproco de claves para correo electrónico y grupos de noticias. En la red oscura se pueden distinguir dos tipos de sistemas de gestión. Uno de acceso condicional como forma simple de administración de derechos en el cual un cliente suscriptor tiene acceso solo a los objetos hospedados en los *hosts* de la red oscura basados en el contrato de servicio. Por ejemplo, servicios de televisión y de radio por cable ofrecen este contrato de servicio y poca o ninguna protección contra la introducción de objetos en el *host*. En consecuencia, la profusión de contenido en formato video en la red oscura es un motivo de preocupación ante los vastos aumentos del ancho de banda, la confiabilidad, la facilidad de uso, el tamaño de la biblioteca compartida y la disponibilidad de los motores para búsquedas en Darknet. No obstante, este sistema de acceso condicional en la red oscura se protege técnicamente con la distribución de claves de difusión, pues cada cabecera de televisión por cable emplea una clave de cifrado disponible que se cambia en cada transmisión para cada cliente. En este sentido, los operadores de cable se encuentran en una posición más segura y privilegiada que los operadores de servicios de televisión satelital ya que pueden realizar una difusión más económica y en casos maliciosos, mediante la piratería de la señal.

El segundo sistema de gestión de acceso es el clásico *Digital Rights Management* (DRM), donde un cliente obtiene contenido en forma protegida por la encriptación y la licencia específica de uso del contenido. Se diferencia de los sistemas de administración de derechos de acceso condicional porque, generalmente, sí imponen restricciones para el uso de los objetos una vez se han desbloqueado. La licencia y el contenido envueltos en el sistema DRM son la garantía de manera general que el cliente:

a) no pueda eliminar el cifrado del archivo y enviarlo a un par,

b) no pueda clonar el sistema DRM para que sea ejecutado en otro *host*,

c) no pueda separar las reglas de la carga útil, y

d) obedezca las reglas establecidas en la licencia DRM[49].

[49] Biddle & et al., 2002, pág. 11.

Otras posibilidades del sistema DRM es la codificación del contenido utilizado en video DVD, por ejemplo. Sin embargo, esta protección se rompió hace muy poco tiempo con la introducción de la tecnología de compresión, el aumento en las capacidades de almacenamiento y el ancho de banda que facilitaron hospedar el contenido de video DVD en Darknet. De otro lado, si los costos y los términos de la licencia son atractivos para productores y consumidores, entonces el proveedor prospera. El caso contrario, la empresa fracasará. Los sistemas DRM actuales se incorporan al software de computadoras personales con variedad de posibilidades que hacen difícil cualquier alteración. La mayoría de los sistemas DRM comerciales tienen explotaciones BOBE (*break once, break everywhere*) y según los expertos se observa que la red oscura también aplica DRM para restringir la introducción de contenido, lo que es difícil de cuantificar en relación con la efectividad, especialmente.

Con el fin de contrarrestar el pillaje en la esfera de la seguridad pública nacional, el Ministerio del Interior de Alemania constituyó en 2017 la Agencia Central de Tecnología de Información (ZITiS), como intento para abordar centralizadamente el ciberdelito y el espionaje digital. Una función es la vigilancia masiva a las telecomunicaciones, el cifrado de información y la recopilación global de datos. Esta unidad de fiscales trabaja para encontrar ciberdelincuentes en la red oscura, ya que existen “las más diversas áreas que en la vida real. Por ejemplo, ventas de armas y drogas, pornografía infantil. Es decir, todos los delitos que se encuentran en la vida normal sin computadoras” (Wagner, 2018), sostiene Julia Bussweiler, fiscal de ZITiS.

Para ciberdelincuentes, como Sascha Flamm, alías Prestamista, el mayor distribuidor de drogas sintéticas como MDMA de *Silk Road*[50], “Darknet hizo posible estar en la mitad de los productores y los consumidores finales de droga. De manera que evitábamos toda la cadena, lo que permitía ventas de gramos, pero cientos al día” (Wagner, 2018). Sin embargo, no fueron los especialistas en ciberseguridad los que atraparon a Prestamista. Fue una delación de un socio infiltrado de Austria. Para el fiscal de ZITiS, Carl Rüffer, “las posibilidades técnicas para

[50] Un mercado negro en línea operado como servicio oculto de Internet profunda, cerrado por el FBI en octubre de 2013.

investigaciones en la red oscura son limitadas y difíciles debido a la información encriptada usada en las comunicaciones"[51], así que los investigadores en ciberseguridad deben trabajar encubierto con métodos tradicionales.

Estamos en la era del Internet de las cosas, donde aspiradoras, refrigeradores, identificadores de objetos, cámaras web, transporte automatizado, drones, locomotores bípedos, utilizan el progreso de la inteligencia artificial mediante el aprendizaje automático de las máquinas. Estas destrezas traen consigo nuevos mercados potenciales y originales incentivos para continuar con las investigaciones en IA. Mientras tanto, grupos de criminales usan estos sistemas para cometer ataques simultáneos a cantidades de redes de computadora y dispositivos electrónicos. Para Alperovitch:

> Lo que vemos en principio es que en estos momentos se está produciendo una transición. Desde nuestra perspectiva, se aumenta la disposición a discutir abiertamente estas cuestiones ya que el conocimiento debe ser actualizado y compartido. Es decir, ya hay intercambio entre las empresas sobre las diferentes organizaciones y expertos quienes comparten datos con mucho detalle acerca de qué tipo de ataques existen y cómo deben protegerse[52].

Alteraciones en el paisaje

El concepto inteligencia artificial ha mudado de definición durante las últimas 5 décadas: de "una rama de la informática que estudia las propiedades cognitivas mediante la síntesis de sistemas"[53] ha pasado por un consenso reciente en torno a la idea de la "actividad dedicada a hacer que las máquinas sean inteligentes, siendo la inteligencia esa calidad que permite a una entidad funcionar de manera adecuada y con visión de futuro de su entorno"[54]. Se diría, entonces, que la pregunta de Turing sobre si las máquinas pueden pensar estaría resuelta.

[51] Wagner, 2018.
[52] Makuch, 2016-2017.
[53] Simon, 1995, pág. 102.
[54] Nilsson, 2010, pág. 183.

Disciplinas como la robótica, el procesamiento de lenguajes naturales y el reconocimiento de imagen, texto y voz artificial se han incorporado al paradigma general de la investigación en sistemas que incluyen IA. De modo que aún queda por resolver la segunda parte de la inquietud del matemático británico: "¿si una máquina puede pensar, es posible que piense de manera más inteligente que nosotros?"

Lo certero, por ahora, es que "el objetivo de la investigación en IA ha sido comprender los principios que subyacen en el comportamiento inteligente para aplicarlos en la construcción de máquinas capaces de presentar dicho comportamiento"[55]. Anota Russell, además, que, para ayudar a las máquinas a adaptarse a las nuevas circunstancias, la IA incorpora las teorías de probabilidad para gestionar las incertidumbres, de la utilidad para definir objetivos y del aprendizaje estadístico para encontrar una función predictiva basada en datos. Dichas inclusiones extienden relaciones de vecindad de la IA con la teoría del control (que tiene que ver con el comportamiento de los sistemas dinámicos) y la programación matemática (con la cual se analiza la toma de decisiones teniendo en cuenta la escasez de recursos).

El informe "Cien años de estudio en inteligencia artificial –IA 100-" elaborado por el Panel de Estudio compuesto por 17 académicos de universidades y laboratorios de tecnología de Estados Unidos, India e Israel proyectó 8 dominios donde la IA ya tiene o se proyecta con mayor impacto, a saber: "transporte, salud, educación, empleo, seguridad pública, comunidades de bajos recursos, modalidades de trabajo y robótica para servicios –profesional o personal-, tareas domésticas y entretenimiento". Cada uno enfrenta dificultades, desafíos y riesgos relacionados con la IA. El Comité Permanente del informe IA-100 incluye como dificultades: "crear hardware seguro y confiable para detectar y desarrollar robots para transporte y servicios, y para interactuar sin problemas con expertos humanos en temas de salud y de educación". Los desafíos se relacionan con "ganar confianza del público, en especial de comunidades de bajos recursos y en la seguridad pública y, superar los temores de marginar a los humanos de su empleo y de los lugares de trabajo". Por último, el riesgo se proyecta hacia la

[55] Russell, 2017, pág. 226.

"disminución de la interacción interpersonal y de cara a cara debido a la mediación por el entretenimiento"[56].

Algunos de los dominios listados fomentan relación directa principalmente con sectores comerciales -transporte y atención médica-, mientras que otros están más orientados hacia los consumidores -robots para servicio, doméstico y entretenimiento-. Un tercer grupo se extiende a sectores sociales como el empleo, las modalidades de trabajo y las comunidades de bajos recurso. Por lo tanto, es fundamental que la investigación de la IA relacionada con problemas de uso de los sistemas de IA resuelva la correlación de las actuales áreas calientes: i) la investigación apoyada en ciencia y tecnología, ¿tendrá éxito frente al control de la inteligencia artificial para prevenir nuevos riesgos en la humanidad?; ii) el diseño de los sistemas de IA, ¿están dotadas las máquinas de objetivos que se alinean correctamente con nuestros valores sociales?; iii) el valor del servicio, ¿las tareas automatizadas de las máquinas inteligentes seguirán alterando la forma de atención al cliente y los modos del trabajo?; iv) la desigualdad social, ¿los operadores de sistemas de IA requieren de habilidades en capacitación cada vez más altas?; v) el empleo, ¿la mayor frecuencia de automatización de tareas se extenderá a toda actividad de trabajo humano?

El tema del dominio de los sistemas de IA se condensa en el paisaje de las seguridades ocasionadas a través de las máquinas inteligentes y su aprendizaje automático. La composición, por tanto, se expande hacía tres tipologías: física, digital y política. Algunos investigadores del MIT piensan que estamos en un punto de inflexión que nos lleva a replanteamientos radicales como sociedad frente a nuestras actividades clásicas -economía, cultura, ciencia-, ante las plazas de aplicación de los sistemas de IA –aprendizajes, robótica, visión mejorada por computadora, procesamiento del lenguaje natural, entornos colaborativos, crowdsourcing, computación humana, teoría algorítmica para juegos y elección social, Internet de las cosas,

[56] Stanford University, 2016.

ingeniería neuromórfica-, y sobre todo ante "nuestros principios éticos e incluso nuestras premisas ontológicas fundamentales"[57].

1. Seguridad física

Las aplicaciones de IA beneficiosas se encuentran en sectores como salud, educación y hogar creciendo a ritmo acelerado. Se incluye al del entretenimiento, pues la industria musical ha adoptado tecnologías de inteligencia artificial para componer; la del cine, para generar escenas en 3D; la del juego, para incorporar en las pantallas distales fantasías. Así se comparte y navega sobre textos, videos y fotos semánticas, escenario que se considera fundamental para el futuro de la IA.

Compañías organizadas con áreas internas de I+D como Siemens, Philips, GE, dedican valiosos recursos para investigación en interpretación automatizada de imágenes médicas y en el diseño de robots como apoyo a procedimientos hospitalarios y operaciones quirúrgicas. Las principales universidades del mundo establecen departamentos de investigación para comprender acerca del aprendizaje profundo mediante sistemas de inteligencia artificial. Empresas de tecnología global como Apple, Facebook, Google, IBM y Microsoft innovan activamente las plataformas y redes sociales al incorporar técnicas desarrolladas de la programación neurolingüística.

La seguridad física demanda de los individuos habilidades adicionales a la lecto-escritura para enfrentar mayores posibilidades de automatización en las máquinas programadas con sistemas de IA. Por ejemplo, sistemas de armamento autónomo aéreo, terrestre y marítimo llevan a cabo, por ahora, misiones no letales con intervención humana, limitada o aún sin ella para ofrecer mayor protección a las fuerzas armadas. Otra ventaja es la posible disminución de un tercio del costo en relación con los vehículos no tripulados y más de dos tercios del que conllevan los tripulados por fuerzas especiales. Para ataques de largo alcance reducen la experiencia requerida del francotirador y eliminan toda pasión humana por el arma de fuego, que en teoría ejecuta acciones muy objetivas. No obstante, hay que recordar el episodio sobre aguas del golfo Pérsico el 3 de julio de 1988 del vuelo 655 de Iran Air. Un Airbus 300-B2

[57] González, 2018, pág. 17.

derribado por el buque de guerra estadounidense USS Vincennes que causó 290 bajas civiles, incluidas 66 menores de edad. El crucero de la marina estaba dotado por el sistema AEGIS de defensa contra ataques aéreos y por un radar con aprendizaje automático para detectar e identificar aviones de guerra. La tripulación confió en la máquina que reconoció a un F-14 y disparó. Esta relación entre hombre máquina utilizando sistemas autónomos incrementa la escala de ataques letales y la cantidad de daño colateral que lobos solitarios o grupos terroristas pueden hacer a poblaciones civiles, al desarrollar la denominada mente *playstation*[58]. De manera que la dificultad de la IA para los dominios de la seguridad pública y el transporte sigue latente.

Análogamente, los sistemas de redes distribuidas dotados de robótica autónoma usados por ejemplo en servicios de reserva de tiquetes aéreos, aplicaciones bancarias, industria de noticias, juegos multiusuario, pueden ser objetivos de rápidos ataques *botnet* coordinados, lo que desafía aún más el desarrollo de la IA para generar confianza pública. La ventaja técnica de estos sistemas es la distribución de tareas que procesa una única red de manera que si una computadora falla las otras seguirán realizando la función correctamente. Adicional, el sistema complementa la capacidad de procesamiento, de tolerancia a fallos y de almacenamiento de datos, permite muchas conexiones simultáneas de red, autoriza un crecido número de usuarios/clientes, proporciona alto nivel de acciones co-relacionadas, etc.

Las perspectivas para el mercado global de los robots de servicio -profesional y personal- siguen siendo muy positivas ya que no todos los robots autónomos son terrestres. Por ejemplo, la Administración Federal de Aviación de Estados Unidos tiene registrados más de 670 mil drones o robots aéreos entre los años 2016 y 2017. La International Federation of Robotics valoró para el lustro 2010-2015 un crecimiento de 210% de suministro de robots industriales, con y sin componentes de IA. Robots primitivos limpios siguen en amplio uso con servicio más sofisticado, lo que parece ser otro horizonte de futuro.

[58] Alston, 2010.

El valor global de ventas de robots para servicio profesional aumentó US $6,6 mil millones y el total vendido se incrementó en 85% en el año 2017. En ese año, se instalaron 69 mil sistemas logísticos robotizados por un valor de ventas estimado en US $2,4 mil millones (incremento de 138% frente al 2016), que representa 162% más unidades situadas que las de 2016. Existen en fabricación 6,700 vehículos de conducción automática y otros 62 mil en entornos de no producción. Los robots médicos también presentan potencial de crecimiento considerable al revisar el valor total de ventas en 2017: US $1,9 mil millones, que representan 29% del total de ventas de los robots de servicio profesional. La cantidad total de robots para actividad agrícola -en su mayoría robots de ordeño-, vendidos en 2017, fue ligeramente superior a 6 mil unidades, que representan ventas aproximadamente en 15% del valor total de las ventas de robots de servicio profesional[59].

Al mismo tiempo, el mercado de robots para servicio personal -sean para asistencia de hogar o para entrenamiento- progresa rápidamente. El valor de ventas subió US $2.1 mil millones mientras que la cantidad total aumentó a 8,5 millones de unidades en 2017. Se estima que casi 6 millones de robots para tareas domésticas se vendieron en ese mismo año, incluyendo los de limpieza por aspiración, corte de césped, lavado de ventanas, entre otros tipos[60].

Jeff Wong, Director Global de Innovación de EY, una empresa de China que ofrece servicios de asesorías a organizaciones para aprovechar oportunidades del mercado tecnológico, está invirtiendo en inteligencia artificial para ser consecuente con los recursos colocados en el sector del *big data*. Argumenta que:

> La IA es obviamente una de las tecnologías más importantes de esta próxima generación y decidimos aplicarla para ayudar a nuestra gente a leer y caracterizar contratos de arrendamiento, específicamente. En la fase piloto de aplicación de esta tecnología ascendimos de 200% a 300% en todas las métricas relevantes, lo que nos induce a lanzarlo al mundo como un verdadero producto beta[61].

[59] IFR, 2018.
[60] Ibid.
[61] Holak, 2018.

En síntesis, el mercado ha variado el paisaje de esta seguridad cuya consecuencia estriba en que los dominios de la IA en seguridad pública y robots para servicio profesional y personal subsisten según la oferta y la demanda. Hay otros saltos inmensos que están afectando los dominios del empleo y los modos de trabajo debido a los vehículos autónomos y los aprendizajes profundos en educación, que cuentan con agentes automatizados a gran escala y cuya interacción para la resolución de problemas no requiere en principio de los maestros expertos. En relación con el principio de justicia de la diferencia[62] han dicho que por causa de la ausencia de idoneidad tecnológica en los sensores existentes de la actualidad, sería muy difícil para un robot autónomo aplicar las capacidades humanas exigentes para la comprensión del contexto y la interpretación de intenciones y de emociones. Entonces, aún la sociedad civilizada tiene esperanza con base en el pensamiento crítico de la inteligencia humana.

2. Seguridad digital

Esta seguridad podría estar más robustecida mediante mejores prácticas en el comportamiento del usuario al usar la información en línea. La actitud es la primera defensa de prevención ante la automatización de ataques con base en la ingeniería social.

El desarrollo principal de los sistemas de IA es la maduración del aprendizaje automático para agentes inteligentes estimulado por el auge de la economía digital. Es un enclave preponderante que aprovecha y proporciona las considerables cantidades de datos diseminadas en las Bases de Datos. Otros circuitos influyentes son el aumento de los recursos de computación en la nube y la demanda de los prosumidores ante el acceso generalizado a servicios, como el reconocimiento de voz y el soporte de navegación web[63].

A medida que la investigación en IA se desarrolla, el procedimiento humano se verá más reproducido a través de una multitud masiva de agentes inteligentes automatizados que vencen los servicios en línea, evitan el acceso de usuarios legítimos y encauzan potencialmente al sistema a estados menos seguros. Las denominadas tecnologías digitales cotidianas, que

[62] Sharkey, 2008; Asaro, 2012; Dinstein, 2012.
[63] Stanford University, 2016.

incluyen planeación de recursos empresariales y búsqueda y procesamiento de información, han impulsado fuertes cambios. Por ejemplo, tanto la velocidad de *clics* como los recorridos humanos en la navegación por el sitio web de destino ya se igualan con patrones promedio de las personas, distribuidos en el algoritmo del software. Se utilizan estilos y tonos de voz artificial que personifican contactos reales. Contenido de la *fake news* se confecciona de manera realista con texto y audio engañoso, para exponer a líderes expresando enardecidos comentarios que con certeza nunca hicieron. La tecnología para la traducción automática de texto también se facilita en diferentes idiomas con bastante precisión, sin depender de traductores humanos. El escenario más aterrador de esta seguridad digital es que "el robot autónomo letal tiene el potencial inigualable de operar a un ritmo más rápido que los humanos y atacar con letalidad deshumanizada, aun cuando se le haya suspendido toda forma de comunicación"[64]. Desde el surgimiento del software se preveía la afectación al dominio del empleo, pero con los sistemas de IA ese menoscabo se extiende al de educación y al de seguridad pública provocando nuevos desafíos sobre todo para los principios de justicia.

Ciertas formas de manipulación y de disponibilidad de información utilizan las plataformas multimedia con algoritmos de curación de contenido para orientar a los usuarios y llevarlos lejos de ciertos conocimientos interesantes a través de su comportamiento como usuario pasivo. Con base en los análisis de automatización de la IA, equipos de marketing y de campañas publicitarias en línea se están aprovechando de las redes sociales habilitadas para identificar a los *influencer* claves, quienes luego son abordados con ofertas ladinas y emplazados con desinformación. Otros ataques en redes sociales y en aplicaciones de mensajería instantánea aprovechan los *bots* para la denegación de información a gran escala, a través de acciones que inundan de ruido la comunicación digital, lo que hace más difícil para las personas menos informadas adquirir la información verdadera y tomar decisiones acertadas. Sin análisis previo, los *mass media* le hacen el juego a estas campañas de desinformación, donde los individuos son protagonistas del envío de cambiantes mensajes hiper personalizados con el fin de afectar el contenido de la natural comunicación social, cada vez más digital. Al mismo

[64] Thurnher, 2017, pág. 83.

tiempo, nuestro lenguaje es mucho más simulado con lenguaje artificial que se registra como *big data* de preferencia por los usos individuales del consumo de información.

La IA incorpora cada vez métodos automáticos digitales que mejoran el servicio en cuanto a rapidez y a accesibilidad. También el progreso ha sido explosivo para el procesamiento de cantidades de imágenes almacenadas en la web. Sin embargo, es sorprendente que esa comprensión sustancialmente digital a los archivos visuales clasificados con débiles etiquetas no haya dado todavía un giro similar en el reconocimiento y la interpretación de imágenes médicas. Será el futuro prometedor para el dominio de la IA en salud, pero mientras tanto, como en otros dominios, la recopilación de registros de salud electrónicos (EHR) desde dispositivos de monitoreo personal y aplicaciones móviles es un capital clave que aún requiere de regulación política con respecto a compartición de los datos y eliminación de los obstáculos comerciales. Otro tanto para que proveedores de atención y pacientes ganen la confianza de los sistemas de IA y pueda promoverse el desarrollo de incentivos y mecanismos para que las máquinas inteligentes interactúen naturalmente con los cuidadores, los pacientes y sus familias.

Economistas consideran que el PIB mide el valor de muchos productos digitales, pero deja de considerar el excedente del consumidor digital. Por tanto, se toman malas decisiones políticas que ocasionan estrictas regulaciones gubernamentales cuyos juicios son de alto riesgo. A medida que la IA se materializa en más recursos, este problema se vuelve más importante debido a que "parece disminuir el PIB, aun cuando las personas tengan mayor bienestar por el acceso a esos bienes digitales"[65].

Ciertas formas de ciberseguridad contra el *spam* y la detección anticipada de *malware* están atacando como defensiva con base en las propiedades de eficiencia y de escalabilidad de la inteligencia artificial. Aun así, la sociedad debe exigirse evitar los incentivos ofrecidos a muchos actores ladinos por experimentar con la IA y justamente agredir sistemas de otros típicamente inseguros. La seguridad digital con sistemas de IA implementados reta la esfera de los Derechos Civiles, en especial con lo relacionado a la libertad de expresión y la protección al *Habeas Data* y

[65] Brynjolfsson & Saunders, 2009, pág. 95.

al Derecho Internacional Humanitario. La protección de la privacidad a través de tecnologías con IA está empezando a ser regulada como una actividad de cumplimiento que llevaría a que el profundo impacto futuro aplique monitoreo sin límite a comunidades estigmatizadas, incremente situaciones de vigilancia y seguimiento injustificado, sea más permisiva ante el robo de datos personales, etc. España y Francia que mantienen principios estrictos y detallados en regulación con base en la gestión de cumplimiento empresarial han disminuido la innovación y las sólidas protecciones a la privacidad. Contrapuesto, el entorno regulatorio de Estados Unidos y Alemania combina objetivos más ambiguos con requisitos exigentes de transparencia y cumplimiento. Este modelo tiene más éxito en las empresas para considerar a la privacidad como su responsabilidad. Para los investigadores en IA, el ideal es diseñar proactivamente la tecnología empleando personal profesional que fomente y adopte las prácticas para proteger la privacidad desde los procesos de fabricación y de negocios, en lugar de responsabilizar internamente a una empresa de la protección de la privacidad.

La regulación a la privacidad como responsabilidad socio-empresarial debe orientarse a:

a) Promover democráticamente el desarrollo y la distribución equitativa de los beneficios de la IA
b) Alentar confianza dentro de las empresas para el cabal desarrollo de requisitos en transparencia y cumplimiento significativo con personal profesional y procesos competitivos
c) Apoyar la expansión de asociaciones de profesionales en comercio para adaptar sanas prácticas sociales
d) Favorecer comités de normas y regulaciones ante los avances tecnológicos que respete los derechos civiles.

Ya están sobre la mesa de discusión de Naciones Unidas dos propuestas de regulación internacional para los sistemas de armamento autónomo: una de 25 países encabezada por Chile, Brasil y Austria que considera iniciar este año alguna negociación de prohibición y otra planteada por Alemania y Francia que convoca a regular con carácter político y de urgencia el

uso de las armas letales autónomas. En relación con el principio de justicia de la libertad, el Alto Comisionado de las Naciones Unidas para los Derechos Humanos ha dicho en el informe anual de 2014 que se deben considerar "marcos nacionales que incorporen las "limitaciones de uso" con respecto a la recopilación de datos y su uso posterior para objetivos legítimos"[66], debido a las innovaciones en la tecnología digital para la comunicación, como Internet, los *Smartphone* y los dispositivos con acceso a *Wi-Fi*. Estas herramientas han impulsado la libertad de expresión, facilitado el debate a nivel mundial, fomentado la participación democrática, pero también han aumentado la capacidad de gobiernos, empresas y particulares para realizar actos de vigilancia, interceptación y recopilación de datos simultáneos invasivos con objetivos precisos y a gran escala.

3. Seguridad política

Los puntos de intersección entre las tecnologías de información y el ámbito político son múltiples. Históricamente, los avances tecnológicos han jugado un papel importante en la libertad de expresión y de información, autonomía indisoluble para la democracia. Basta recordar la aplicación del telégrafo en la Francia napoleónica, el ingreso del Sistema de Posicionamiento Global durante la Primera Guerra del Golfo, el uso de las redes sociales durante la Primavera Árabe. Los avances en tecnología cambian el equilibrio de poder entre los Estados, así como la relación entre los líderes en ejercicio y los ciudadanos que buscan desafíos, principalmente debido a que las tecnologías no son neutras. Empoderan tanto a sujetos activos como a comunidades virtuales a través de actitudes y sentimientos mediados por la comunicación instantánea y móvil. Sin embargo, también proporcionan capacidad de transmisión de mensajes de manera rápida y fácil para asuntos militares e inteligencia con el objetivo de vigilar.

La investigación en IA ha demostrado que algoritmos de aprendizaje automático se ejecutan en plataformas web y redes sociales, ahora convertidas en el Ágora preferida de los gobernantes, para priorizar el contenido y así tener control estatal automatizado. Un fin más perverso es el

[66] Naciones Unidas, 2014, pág. 10.

de reprimir disidencias. El poder de observación de las naciones, ampliado mediante la automatización del procesamiento de imagen y audio, ya ha alterado el paisaje de la seguridad política por cuanto vuelve a la inteligencia artificial aún más sofisticada. Entonces, con los agentes inteligentes automatizados se recolecta, procesa y usa información a escala masiva para innumerables propósitos, incluida la supresión del debate. Un segundo sector que se está priorizando con la IA es el empleo con trabajos susceptibles de automatizar. Ejemplos como el servicio de atención al cliente a través de asistentes virtuales con suficiente capacidad lingüística y la copia y transferencia de información técnicamente democratizan, pero en modo desfavorable al sistema democrático y en especial al dominio formado por comunidades de bajo recurso.

El exceso de producción con base en agentes automatizados puede redundar en disponibilidad de recurso adicional. El reto será la distribución habida cuenta que la tecnología altera las dinámicas políticas a través de los variados tipos de regímenes políticos. Por ejemplo, el gobierno de China está explorando formas de aprovechamiento de los datos en línea y fuera de línea para obtener un mayor puntaje de aprobación social de sus ciudadanos. De manera más generalizada, la censura en China ejemplifica el aprovechamiento más explícito de la tecnología con fines políticos.

Con el crecimiento de grandes conjuntos de datos habilitados por Internet es complejo el análisis orientado a mitigar y solucionar la variedad de problemas sociales agravados. Los recientes avances en tecnologías sensoriales y aplicaciones con aprendizaje profundo atraen tanto el incentivo económico de financiamiento como las prioridades electorales. Esta situación de presión lleva a los gobiernos a formular regulaciones cada vez más duras o no reglamentar nada, sobre el malentendido de lo que es IA debido al contexto alarmista. Adicional, los cambios en la seguridad política con la carrera armamentista oscilan entre producción y detección de engaños, pues aún no es claro cuáles serán las implicaciones a largo plazo por los usos maliciosos de la IA. Es de entender que las amenazas letales pueden fomentar más medidas de prevención y mitigación vigorosas.

Nos encontramos en una crucial coyuntura humana con respecto a ¿cómo implementar tecnologías basadas en IA de manera que promuevan los valores democráticos como la libertad, la igualdad y la transparencia? Se estima conceptualmente que la investigación contribuye en la creación de modelos predictivos para prevenir estos problemas. Sin embargo, el uso generalizado de los medios sociales con bajas barreras de entrada es canal propicio para que los sistemas de IA simulen humanos con puntos de vista partidistas al implementar estrategias políticas *bots*.

La vigilancia e interceptación desde los Estados para enfrentar la seguridad nacional cada vez amplía la frontera de la automatización permitiendo el procesamiento de imagen y de audio y la recolección y la utilidad de información de inteligencia a escalas masivas y multilaterales para innumerables propósitos. Por ejemplo, en el año 2007, el mundo fue testigo de la primera convergencia entre cibernética y operación militar cuando Estonia fue víctima de una serie de ataques DDoS realizados por *hackers* rusos del GRU que lograron menoscabar los servicios básicos del país. Otros apoyos y cooperaciones prevalecen entre agencias nacionales de las principales potencias cibernéticas, como Estados Unidos, Rusia, China, Reino Unido o Israel. Además, las alianzas entre esos gobiernos los ubican en un nivel muy superior de madurez cibernética ante el grueso de los países del mundo. El resultado de esas alianzas y cooperaciones profundiza el colonialismo al tiempo que una gran mayoría de gobiernos se ha abandonado al señalado cortoplacismo cibernético, caracterizado por 4 aspectos:

a) Inclusión del ciberespacio en la agenda política nacional, inducida por un socio o aliado extranjero
b) Reconocimiento del ciberespacio como importancia estratégica, forzada por las circunstancias y sin convencimiento efectivo
c) Creación de un sistema nacional de ciberseguridad inadecuado para gestionar el cambio con rapidez

d) Auto-complacencia por los logros alcanzados en breve tiempo y por las capacidades presentes del sistema nacional de ciberseguridad[67].

Las tecnologías con inteligencia artificial podrían ampliar todavía más el carácter disruptivo de las amenazas a las seguridades física y digital. Es decir, la capacidad de desplazar innovaciones preexistentes y de provocar profundos cambios en los actores que las usan, en sus estrategias y en los efectos de sus acciones[68] deben ser implicaciones preocupantes para la seguridad política. Son nuevas desigualdades de oportunidad siempre que el acceso a los sistemas y máquinas con IA, a la computación de gran potencia y a los datos a gran escala se distribuya socialmente con inequidad ya sea a cargo del Estado o por intermedio de entidades privadas que operan bajo ciertas reglas y directrices emitidas por el Estado. De otro lado, el uso eficiente de los datos es demasiado costoso para muchos gobiernos autoritarios, pero al utilizar sistemas de IA mejoran la capacidad para priorizar la atención y reducir el costo del monitoreo a individuos.

Los efectos de esta tecnología sobre el cambio político dependen del contexto y de la habilidad de los actores que la utilizan. Ante los profundos cambios que ocasionan las tecnologías, ya para la expansión de la democracia o para la consolidación del autoritarismo y la represión, los gobiernos se están viendo enfrentados a regular la presión de los sistemas de IA. Herramientas técnicas desarrolladas detectan falsificaciones en los mensajes de las redes sociales, la autenticidad certificada en el uso de imágenes y videos demuestra la verdad de la transmisión en vivo, la encriptación de información garantiza seguridad en la difusión del contenido. No obstante, estas acciones liberadas residen en manos de corporaciones que actúan bajo la oferta y la demanda del mercado complicando la acción política regulatoria. La secuela severa para las democracias es que el paisaje de la seguridad política requiere de normalización gubernamental por cuanto la actual es aún insuficiente o desequilibrada.

[67] Hernández Lorente & Fojón Chamorro, 2014.
[68] Castells, 2001.

Una segunda repercusión ocasionada por la carencia de políticas públicas que ayuden a facilitar la seguridad en las relaciones de poder tiene que ver con la posible expansión de más errores y fallas inevitables en la adaptación social a las aplicaciones desarrolladas con sistemas de IA. Así como el uso de los datos ciudadanos por parte de las agencias de inteligencia del Estado. De manera tal que sus beneficios están "abiertos para propósitos nobles y nefastos de la gente, de las organizaciones y de los gobiernos que determinarán que prevalece"[69].

Avances en inteligencia artificial y en sistemas de aprendizaje automático incorporado están empezando a mostrar cómo será el futuro cercano con UAV, vehículos autónomos, armas letales autónomas, vigilancia automatizada, manipulación y engaño automatizado. Mientras tanto, la esfera pública y privada de la educación y la enseñanza poco aumentan la calidad de habilidades de pensamiento crítico, las intervenciones del discurso ético para mejorar siguen siendo generales y las herramientas para la transparencia de campañas políticas en redes sociales son simples propuestas políticas para momentos coyunturales en tiempos de elección y no aplicaciones que fomenten el diálogo constructivo permanente.

El escenario de las previsibles consecuencias ya está impactando la seguridad política tanto de agencias gubernamentales como de comunidades de bajo recurso con la irrupción del populismo que exaspera la tragedia primero y la farsa, después. El debate debería de contrarrestar el peligro a aquellas ciudadanías en relación a ¿cómo se implementa la IA mediada por comunicación digital en las comunidades iletradas? y ¿cómo se protege la privacidad de datos personales y los privilegios de adquirir recursos con IA en sociedades desiguales? En general, las empresas de tecnología, los medios sociales y las corporaciones mediáticas son puntos críticos para el control y freno a campañas de desinformación, censura y persuasión cada vez más automatizadas. Otro tanto es que aquellas corporaciones que poseen conjuntos de datos personales únicos, útiles para el desarrollo de sistemas de IA, los compartan con agencias del Estado siempre y cuando existan de manifiesto principios de transparencia. No

[69] Diamond, 2010, pág. 72.

obstante, es de mundial conocimiento que las corporaciones globales asumen ser dueñas de los datos personales de sus usuarios.

Se requieren quizá con urgencia de estos y otros esfuerzos que contribuyan a detectar amenazas antes que impacten todavía más el paisaje general de la seguridad de las máquinas inteligentes, del software en los sistemas de IA y del contrapeso en las relaciones de poder. Implementar otras estrategias globales para asegurar que el uso de agentes inteligentes automatizados se alinee con el interés público y evite el *hackeo* a la democracia. Y mientras la regulación ocurre, la investigación en el campo de la IA se mantendrá bajo argumentos de interés global dado que "el envío de sondas artificiales al espacio cada vez más distante del sistema solar no se puede pre-programar para que lo encuentren ni para interpretar el *big data* de opción humana, que va a seguir en expansión"[70], según el jefe de la División de Supercomputación Avanzada de la NASA, Bryan Biegel. Por su parte, el Panel de Estudio del informe IA 100 recomienda que todas las capas del gobierno adquieran experiencia técnica en inteligencia artificial y alienten investigaciones sobre equidad, seguridad, privacidad e implicaciones sociales, provenientes de estos sistemas, eliminando los impedimentos y reduciendo el gasto público y privado para apoyar más inversiones en dichas acciones[71]. De lo contrario, las falsificaciones profundas aplicadas en algoritmos de software vendrán con un realismo suficiente para disminuir la posibilidad de ver para creer.

CONCLUSIONES

El desarrollo en la programación del software ha alcanzado niveles de las redes neuronales profundas, lo que facilita el incremento eficiente y escalable de los sistemas de inteligencia artificial. De manera que el futuro humano tendrá que enfrentar con mejores propuestas de política pública las agresiones muy específicas, las altamente efectivas, las dirigidas con precisión y las que aprovechan las vulnerables no resueltas de software, que se extrapolan a los contextos individual y social.

[70] Makuch, 2016-2017.
[71] Stanford University, 2016.

Al primer entusiasmo bajo principios naturales de no hacer daño y de experimentar con innovaciones surgidas de la inteligencia personal le sobrevino rápidamente el deseo de poseer un aparato computacional original. Esta variación impactó en las modalidades del trabajo ya que se pueden automatizar tareas desde agentes inteligentes artificiales y así dejar de agobiar a las personas por la monotonía de la repetición. El segundo frenesí sustentado en el anonimato refuerza el espíritu de libertad, pero al mismo tiempo incrementa la falta de empatía en la naturaleza humana. Estas dos exaltaciones por la tecnología avanzan con sistemas de IA incorporados en máquinas automatizadas programadas con base en el aprendizaje profundo. No obstante, detrás de cada una de ellas concurren sujetos que con su comportamiento afianzan la propagación de las amenazas y debilitan el horizonte de las seguridades. La ecuación clave debería ser a la inversa.

Los métodos utilizados por agresores para cometer cibercrímenes focalizan las debilidades en las habilidades digitales de las personas que caen víctimas, sean de modo individual, empresarial o gubernamental. Llegar a estas precisiones requiere de ciertas estrategias que gestionen la intromisión en los estados emocionales de los usuarios de las tecnologías. Así que la ingeniería social se centra en la comunicación digital y embiste ciberataques, y según sea su respuesta ejecutada la víctima obtendrá consecuencias destructoras. Aun cuando los motivos para atacar son disimiles, el plan con el que se lleva a cabo la amenaza está aprovechando la potencialidad de los sistemas de IA. Por tanto, la ciberseguridad no alcanza a resguardar todos los frentes debido a la automatización que se expande hasta para actos terroristas de algunos estados. Inclusive, servicios de empresas dedicadas a la ciberseguridad también se incluyen en la lista de víctimas corporativas.

Un segundo nivel en la escala de ataques en masa tiene en la mira la infraestructura de servicios y la seguridad nacional de los países. Nuevos métodos como APT, *botnet* y *ransonmware* son más sofisticados, pero con un escenario de actores bastante peligrosos a razón de las alianzas y cooperaciones entre bandas organizadas de cibercriminales y agencias de inteligencia militar de Estados autoritarios. Otros Estados conciben la ciberseguridad como un propósito partidista de

la agenda pública nacional. No obstante, la existencia comprobada de datos sobre tipos y modos de ataque a disposición de los países aún es tema central que no se comparte ni se discute como prioridad abierta para todos los Estados.

El campo de la investigación de la IA es ya un paradigma que usa robótica, procesamiento de imagen y reconocimiento de voz y texto artificial. Incorpora teorías de probabilidad, de utilidad y de aprendizaje estadístico para alcanzar los objetivos en los sistemas de IA. Este marco multidisciplinar ha generado un entorno de relaciones con la teoría del control y la programación matemática. Aun así, los investigadores de sistemas de IA en su desarrollo creciente deben superar dificultades ante la seguridad y control del hardware, retos ante la desconfianza pública en concreto para las cuestiones del empleo, así como enfrentar riesgos que limitan con la preservación del derecho a la información y la protección de los datos personales privados.

El avance en la programación de algoritmos para la simulación humana está destruyendo la frontera del sistema de democracia debido a acciones de manipulación y engaño relacionados con recolectar, procesar y usar información para monitoreo, inspección, vigilancia, control, supresión, de actividades ciudadanas y sociales. Comprender acerca de las amenazas de los sistemas de armamento autónomo y de las redes distribuidas que incorporan IA es más un esfuerzo de la inteligencia crítica humana que de los agentes inteligentes automatizados al asumir los comportamientos de las personas como producto de la escala del aprendizaje automático.

BIBLIOGRAFÍA

Allden, K., & Murakami, N. (2015). *Trauma and recovery on war's border.* Hanover: Dartmount College Press.

Alston, P. (2010). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary.* Geneva: United Nations Human Rights Council.

Asaro, P. (2012). On banning Autonomous Weapons Systems: Human rights, automation and the dehumanisation of lethal decision making. *International Review of the Red Cross*, 94 (886), 687-709.

Biddle, P., & et al. (18 de noviembre de 2002). *The Darknet and the future of content distribution.* Obtenido de Crypto Stanford: https://crypto.stanford.edu/DRM2002/darknet5.doc

Brundage, M., & et al. (2018). The malicious use of Artificial Intelligence: forecasting, prevention and mitigation. *Maliciousaireport.com*, 1-100.

Brynjolfsson, E., & Saunders, A. (2009). What the GDP Gets Wrong (Why Managers Should Care). *Sloan Management Review*, 51 (1), 95-96.

Castells, M. (2001). *La Galaxia Internet. Reflexiones sobre empresa, sociedad e Internet.* Barcelona: Plaza & Janés.

Cummings, M., & et al. (2018). *Artificial Intelligence and international affairs disruption anticipated.* London: Chatham House.

Diamond, L. (2010). Liberation Technology. *Journal of Democracy*, 21 (3), 69-83.

Dinstein, Y. (2012). The principle of distinction and cyber war in International Armed Conflicts. *Journal of Conflict and Security Law*, 17 (2), 261-277 .

Drozhzhin, A. (20 de abril de 2017). *Trampas en aplicaciones: el impresionante mundo de los bots de Tinder.* Obtenido de Kaspersky Lab: https://www.kaspersky.es/blog/tinder-bots/10439/

Gómez Blanco, A. (8 de enero de 2018). *Ataques de ingeniería social: qué son y cómo evitarlos.* Obtenido de BBVA: https://www.bbva.com/es/ataques-ingenieria-social-evitarlos/

González, F. (2018). De la era de la perplejidad a la era de las oportunidades: finanzas para el crecimiento. En J. Kallinikos, & et al., *La Era de la Perplejidad. Repensar el mundo que conocíamos.* (pág. 416). Madrid: Taurus.

Hernández Lorente, A., & Fojón Chamorro, E. (22 de abril de 2014). *De Estonia a Ucrania, la evolución de los conflictos en el ciberespacio.* Obtenido de Belt: http://www.belt.es/expertos/HOME2_experto.asp?id=6927

Holak, B. (octubre de 2018). *Por qué EY está invirtiendo en iniciativas de IA, datos y blockchain.* Obtenido de Search Data Center en Español: https://searchdatacenter.techtarget.com/es/cronica/Por-que-EY-esta-invirtiendo-en-iniciativas-de-IA-datos-y-blockchain

IFR. (18 de octubre de 2018). *Service robots - global sales value up 39 percent.* Obtenido de IFR: https://ifr.org/ifr-press-releases/news/service-robots-global-sales-value-up-39-percent

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: trends, perspectives, and prospects. *Science* , Vol. 349, pp. 255-260.

Lomas, N. (25 de 04 de 2017). *Lyrebird is a voice mimic for the fake news era*. Obtenido de Techcrunch: https://techcrunch.com/2017/04/25/lyrebird-is-a-voice-mimic-for-the-fake-news-era/

Makuch, B. (Dirección). (2016-2017). *Cyberwar. Hacked by China.* [Película].

Mitnick, K., & Simon, W. (2007). *El arte de la intrusión.* México: Alfaomega Grupo Editor.

Mitroff, S. (15 de mayo de 2016). *¿Qué es un 'bot'? Te contamos todo lo que necesitas saber.* Obtenido de CNET: https://www.cnet.com/es/como-se-hace/que-es-un-bot/

Naciones Unidas . (2014). *El derecho a la privacidad en la era digital .* Ginebra: Naciones Unidas .

Nilsson, N. (2010). *The Quest for Artificial Intelligence: A History of Ideas and Achievements.* Cambridge (UK): Cambridge University Press.

Russell, S. (2017). Inteligencia artificial de beneficios probados. En D. Aubrey, *El próximo paso la vida exponencial* (pág. 408). Madrid: Turner.

Scharre, P. (2016). *Autonomous Weapons and Operational Risk.* Washington: Center for a New American Security.

Schultz, R. (2012). Ética e Internet. En F. González, & et al., *Valores y Ética para el siglo XXI,* (pág. 392). Madrid: BBVA.

Sharkey, N. (2008). Grounds for discrimination: Autonomous Robot Weapons. *RUSI Defence Systems*, 86-89.

Simon, H. (1995). Artificial Intelligence: An Empirical Science. *Artificial Intelligence*, 77 (2): 95–127.

Stanford University. (2016). *Artificial Intelligence and life in 2030. One hundred year Study on Artificial Intelligence.* Stanford (CA): Stanford Education.

Tenhaven, J. (Dirección). (2017). *The Silicon Valley Revolution. How a few Nerds did change the world* [Película].

Thurnher, J. (2017). Legal implications of fully autonomous targeting. *Joint Force Quarterly*, 67, 83.

Turing, A. (1950). Máquina de computación e Inteligencia. *Mind*, Vol. 59, No. 236, pp. 433-460.

Wagner, A. (Dirección). (2018). *Cibercrimen. El negocio con el miedo.* [Película].