



# **INSTITUTO DE CIBERDEFENSA DE LAS FUERZAS ARMADAS**

## **DIPLOMATURA UNIVERSITARIA EN GERENCIAMIENTO DE LA CIBERDEFENSA**

TRABAJO FINAL INTEGRADOR

### **GUERRA COGNITIVA: EL USO DE LA INTELIGENCIA ARTIFICIAL PARA MANIPULAR LA OPINIÓN PÚBLICA Y DESESTABILIZAR AL ADVERSARIO**

#### **Integrantes del Equipo Nro 4:**

ERIC JOSUE DORADO  
ALEJANDRO ARIEL FERREYRA  
MAXIMILIANO ESTRADA

#### **Títulos Profesionales / de grado**

Licenciado en Administración

Licenciado en Administración

Técnico Superior en Electrónica Aeronáutica

09 de noviembre de 2024

## Resumen

Se evidencia un vertiginoso desarrollo de la inteligencia artificial en los últimos años empleando el ciberespacio como plataforma base. Coincidente con ello la guerra cognitiva se presenta con importantes particularidades. La influencia de la inteligencia artificial en este tipo de conflicto representa una nueva dimensión en el control de información y manipulación de percepción. De allí que surgió la inquietud por investigar el vínculo entre este tipo de conflicto y la inteligencia artificial

El interés central estuvo dado por analizar la posibilidad de identificar la aplicación de esta nueva inteligencia en los conflictos actuales y de esta manera comprender, de manera general, el funcionamiento. Todo esto con especial atención en la manipulación de la opinión pública.

Se usó un análisis bibliográfico del material brindado por la cátedra como así datos de fuentes abiertas, principalmente páginas de internet para los conceptos teóricos. Luego una investigación general de los hechos históricos seleccionados. Esto permitió la vinculación de lo teórico con lo concreto.

En primer orden se entendió el funcionamiento general de la inteligencia artificial y seguidamente de qué manera los actores (estatales o no) hacen uso para la manipulación.

A través del estudio de casos reciente se identificó la manipulación de la opinión pública causada por la intervención de algunos actores mediante el uso de las redes sociales, la inteligencia artificial en la manipulación de creencias y en la creación de narrativas específicas.

Se concluyó que la guerra por la mente de las personas, de alta complejidad, tiene una arista importante como el funcionamiento de la inteligencia artificial al servicio de esta. Se apreció que estos efectos pueden fragmentar y polarizar a una sociedad, desestabilizando al adversario, afectando principalmente la cohesión social de un país.

**Palabras Claves:** guerra cognitiva – inteligencia artificial – opinión pública

## Índice General

Justificación / Fundamentos / Aportes.....	1
Planteamiento del problema.....	1
Solución propuesta.....	1
Objetivos.....	2
Marco teórico.....	2
Metodología.....	6
Capítulo 1: La Inteligencia Artificial y la Opinión Pública .....	7
Sección 1: Tipos de Inteligencia Artificial .....	7
Sección 2: La opinión pública .....	10
Conclusión parcial .....	12
Capítulo 2: La Inteligencia Artificial y la Guerra Cognitiva .....	14
Sección 1: La Guerra Cognitiva .....	14
Sección 2: Casos de Estudio .....	15
Sección 3: Desestabilización del adversario .....	21
Conclusión parcial .....	24
Conclusiones Finales.....	25
Bibliografía .....	26

## **Justificación / Fundamentación / Aportes**

Una característica importante de la guerra es que, al igual que la sociedad, también ha evolucionado en todo sentido, y lo seguirá haciendo. Desde la prehistoria hasta nuestros días, ha sufrido variaciones considerables desde de una guerra de palos y piedras, flechas y escudos, hasta conflictos de botones de altísima tecnología, de misiles y de destrucción nuclear masiva.

Según sus características y los propósitos de sus participantes, las guerras pueden catalogarse de diferentes nodos: convencional, civil, religiosa, subsidiaria, híbrida, y psicológica, entre otros. Durante el desarrollo del presente trabajo nos enfocaremos en este último concepto en particular, deteniéndonos en lo concerniente a la guerra cognitiva.

El ciberespacio ofrece una guerra sin fricción ni dolor en términos generales. Aquí la guerra cognitiva encuentra un terreno propicio para su desarrollo. Asimismo, puede considerarse como no convencional e híbrida ya que puede utilizar herramientas cibernéticas para alterar los procesos cognitivos del oponente, explotar sus sesgos mentales, el pensamiento irreflexivo y provocar distorsiones en su pensamiento. De esta manera sería posible influir en la toma de decisiones y las acciones, con efectos negativos, tanto a nivel individual como colectivo.

La ingeniería social que se puede considerar que es una mezcla de arte y ciencia con la finalidad de manipular y/o influenciar a una o un grupo de personas juega un rol preponderante en la guerra cognitiva.

De allí se observa la necesidad de investigar sobre hechos concretos pasados el uso y aplicación plausible de la inteligencia artificial (IA) con intención de alterar la opinión pública y determinar la posibilidad de desestabilizar a un oponente.

## **Planteamiento del Problema**

### **Formulación del Problema**

¿En qué grado influye el uso de la IA en la opinión pública y cómo afecta la estabilidad del oponente?

### **Solución Propuesta**

Debido al creciente y vertiginoso desarrollo de la IA y sus múltiples usos se observa la necesidad de realizar un análisis en hechos históricos y presentes. Esto permitiría apreciar convenientemente los impactos positivos y negativos sobre la opinión pública. Advirtiendo que esta es un pilar fundamental para el equilibrio dentro de

cualquier nación políticamente organizada, su debida valoración es esencial para considerarla como una posible arma estratégica.

Bajo esta perspectiva el uso global de redes sociales y de la inteligencia artificial se vio cómo influye en la opinión pública, la cual es profunda y multifacética. Se investigó de que forma la IA contribuye a la capacidad de moldear creencias, reforzar narrativas específicas y manipular emociones, logrando que parte de la población tomen decisiones basadas en información sesgada, parcial o fabricada.

Este tipo de manipulación puede fragmentar a una sociedad, polarizarla y debilitar la cohesión social, lo que tiene implicaciones directas para la estabilidad de cualquier nación

## **Objetivos**

### **Principal o general**

Reconocer posibles usos de la inteligencia artificial orientados a manipular la opinión pública y desestabilizar al adversario.

### **Particulares o Intermedios**

#### ***Objetivo Particular Nro 1***

Describir el funcionamiento de la inteligencia artificial y su influencia sobre la opinión pública.

#### ***Objetivo Particular Nro 2***

Analizar el grado de desestabilización estratégica que logran los diferentes actores estatales y no estatales al aplicar la inteligencia artificial, y su impacto en la ética.

## **Marco Teórico**

Para introducirse en la temática se considera algunos conceptos básicos que luego serán útiles como base de análisis.

La IA es un campo de la ciencia que incluye muchísimas disciplinas, desde la informática, el análisis de datos y estadísticas, la matemática, la ingeniería de hardware y software, la lingüística, la neurociencia e incluso la filosofía y la psicología, y se la define como una inteligencia simulada por algoritmos o máquinas (COECKELBERGH, p. 64).

Las aplicaciones para el uso de la IA son inmensas, desde robots que pueden circular solos (...), sistemas de ciberseguridad (...) para reconocer un ataque, asistentes virtuales que pueden "entender" lo que alguien está diciendo y actuar con esa información, reconocimiento de las emociones (...) e incluso la creación de videos y sonidos totalmente artificiales muy difíciles de distinguir de los reales (MUÑOZ ITURRIETA, p. 145).

Entre otras aplicaciones de la IA tenemos el transporte, el marketing, la medicina, las finanzas, los seguros, los distintos ámbitos de la ciencia, la educación, el entretenimiento, e el arte, la agricultura, la manufactura, la ciberseguridad, la industria militar, e incluso la composición de textos (COECKELBERGH, p. 103).

Como se observa la IA es compleja y de gran amplitud por lo que es advierte como un primer elemento considerado en el marco de la guerra cognitiva.

El concepto de Guerra Cognitiva (GC) ha emergido como una extensión de las formas ya conocidas de conflicto, abarcando la manipulación de percepciones, la desinformación y el control del proceso de toma de decisiones del adversario. Se la puede considerar como el arte del uso de tecnologías para alterar la cognición de blancos humanos, la mayor parte de las veces sin conocimiento ni percepción, ni consentimiento del objetivo. Podría ser la base disruptiva final de la guerra asimétrica, donde la GC explotaría los prejuicios, sesgos y automatismos mentales que provocan distorsión de la realidad, alteración y hasta la inhibición de acciones con desastrosas consecuencias a nivel individual y colectivo

Esta forma de guerra se centra en afectar la mente del enemigo, debilitando su capacidad para interpretar la realidad y tomar decisiones acertadas. Con la introducción de la IA, la GC ha encontrado nuevas herramientas para influir en los comportamientos colectivos, amplificando el impacto de la información dirigida y la manipulación estratégica. En este marco, el estudio del rol de la IA en la influencia de la opinión pública y desestabilización del enemigo requiere especial atención.

En este contexto se trata de dar al presente trabajo una orientación, definiendo el concepto de guerra cognitiva y analizando de que forma la IA comenzó la transformación de las estrategias en los conflictos contemporáneos.

Como se observó en la cátedra, la guerra cognitiva tiene como objetivo principal no solo alcanzar la superioridad física o tecnológica, sino principalmente lograr la "superioridad mental". Esta forma de conflicto se enfoca en influir y manipular la percepción de la realidad, tanto de los decisores como en las poblaciones civiles.

Según estudios de defensa y seguridad (Pamment, 2021), esta guerra opera en el dominio de la información, considerado como el sexto. Se utilizan canales de comunicación como ser las redes sociales, propaganda, campañas psicológicas, desinformación, y noticias falsas para provocar cambios en el comportamiento del adversario o enemigo.

Desde otro punto de vista la GC se insertaría con NEUROARMAS en la Híbrida, la cual será una forma de guerra que no se limite al producto del enfrentamiento entre dos fuerzas oponentes diferentes (regulares contra insurgentes, guerrilleros, fanáticos religiosos); sino que además enfrenten culturas que privilegien de manera diferente el

valor de la vida y el tiempo para obtener resultados en relación al gasto de recursos materiales, cantidad de bajas civiles y militares aceptables por parte de sus connacionales. Son fuerzas que se oponen en forma no cooperativa, sin equilibrio de su poderío militar ni misma matriz estratégica ni operativa.

A diferencia de otros tipos de conflicto, la guerra cognitiva se desarrolla principalmente en el espacio virtual e informacional. Las fuerzas en oposición ya no necesitan utilizar violencia directa para obtener ventajas estratégicas, sino que se concentran en erosionar la confianza, confundir al adversario y reducir su capacidad de comando y control.

Para destacar como ejemplo de lo referido, es preciso observar a una potencia como la República Popular China, quien lidera claramente la doctrina de la GC. Este actor ha demostrado comprender cabalmente su concepto y en ese sentido la aplica. Los autores, Qiao Liang y Wang Xiangsui, de la “Guerra Sin Restricciones” expresan en su prefacio

” reconocemos que los nuevos principios de la guerra ya no son “el uso de la fuerza armada para someter al enemigo a la voluntad de uno”, sino que están utilizando todos los medios, incluso a las fuerzas armadas o no armadas, fuerzas militares o no militares y medios letales y no letales para obligar al enemigo a aceptar los propios intereses”

En ella los militares chinos analizaron el empleo de toda forma de lucha no convencional y sin límites éticos, ante un enemigo superior en fuerza, tecnología e influencia política y económica. El uso de la *Big Data*, la proliferación de redes sociales, la captura de patrones por IA, permiten la anticipación de escenarios por *wargaming* a las que hoy se suma el uso de tácticas de “desgaste cognitivo” mediante técnicas de interferencia y contra interferencia, combinando ingeniería de sistemas y a las ciencias sociales configurando PSYOPS de alto impacto.

El auge de la IA ha potenciado estas dinámicas al hacer posible la automatización y personalización de campañas de influencia masiva en el quinto dominio. Este nuevo elemento introdujo otras capacidades en la guerra cognitiva, permitiendo una manipulación más precisa de las percepciones del adversario. Por ejemplo, las redes neuronales facilitan crear campañas de desinformación dirigidas a segmentos específicos de la población, mientras que los algoritmos de aprendizaje automático identifican patrones de comportamiento y optimizan mensajes para causar el mayor impacto. Por lo tanto, es necesario el estudio de la conformación y desarrollo de estos componentes.

El conocido uso de *Bots* en redes sociales o el análisis automatizado de las emociones, las técnicas de generación de contenidos por IA, como los "*Deepfakes*",

permiten crear videos e imágenes falsos altamente convincentes, socavando la confianza en las instituciones y generando caos social. Estas tecnologías son integradas a las tácticas tradicionales de manipulación, alcanzando que la frontera entre lo real y lo ficticio se conviertan cada vez más indivisibles. Por este motivo es por lo que se estudió la relación entre ellos.

Haciendo un paralelismo histórico, durante la Segunda Guerra Mundial, en la Operación Fortaleza, se puede destacar un relevante ejemplo de GC llevado a cabo por los aliados. Esta operación fue diseñada para engañar a los alemanes sobre el lugar del desembarco aliado en Europa, haciéndoles creer que el ataque principal ocurriría en Pas-de-Calais y no en las playas de Normandía.

En este caso se utilizaron técnicas de desinformación, como la creación de falsos ejércitos mediante radio falsificación y la difusión de mensajes que sugerían movimientos ficticios. El éxito de esta se basó en manipular la percepción del enemigo provocando que los alemanes mantuvieran tropas en la zona equivocada. Si bien esta operación se llevó a cabo antes del desarrollo de la IA, es un ejemplo de cómo la manipulación de información puede ser clave para la victoria en un conflicto.

En la actualidad, la IA facilitaría este tipo de engaños a través del análisis predictivo de decisiones del adversario y la generación de falsos escenarios en tiempo real. Retrocediendo algunos años, durante las elecciones presidenciales del 2016, en Estados Unidos, el mundo fue testigo del pleno empleo de las redes sociales y algoritmos de segmentación para influir en los votantes y sembrar discordia social. A través de campañas masivas de desinformación, se promovieron mensajes para polarizar y explotar las divisiones ideológicas y sociales existentes. Creando perfiles de votantes y personalización de mensajes políticos. Este caso ilustra cómo la IA puede influir en la percepción pública y manipular las decisiones de votantes, provocando efectos a largo plazo en la estabilidad política de una nación. En un contexto bélico, estas tácticas podrían ser utilizadas para socavar la moral del enemigo o desorganizar sus esfuerzos estratégicos.

La GC representa un nuevo paradigma en el que la manipulación de la información y la influencia sobre el comportamiento del adversario se convierten en las armas principales. La IA ha transformado estas dinámicas, permitiendo una automatización y personalización sin precedentes en las campañas de desinformación. Como sería posible observar en los ejemplos históricos de la Operación Fortaleza y la interferencia electoral de 2016, la guerra cognitiva ha evolucionado a lo largo del tiempo, integrando cada vez más tecnología para alcanzar sus objetivos. En la actualidad, la IA plantea oportunidades o desafíos, ya que puede ser utilizada tanto para defender como para socavar la estabilidad de una sociedad. Por ello, el estudio de su impacto en la guerra

cognitiva es fundamental para comprender los conflictos contemporáneos y anticipar los desafíos futuros.

### **Metodología**

Se empleó un método exploratorio y cualitativo para analizar hechos de conflictos a la luz de conceptos vertidos por autores referentes previamente seleccionados por el grupo de investigación y el temario dado en la cátedra. Para ello se realizaron distintos análisis y descripciones durante el desarrollo de cada capítulo a fin de obtener conclusiones parciales surgidas de cada uno en orden a una respuesta para el objetivo principal.

Se aborda el uso de la inteligencia artificial para generar caos, fomentar divisiones internas y erosionar la confianza en las instituciones del adversario. Se examinan casos específicos y las implicaciones en la desestabilización del oponente.

## Capítulo 1

### La Inteligencia Artificial en la opinión pública

#### Sección 1

#### La Inteligencia Artificial

##### Primer tipo de IA utilizadas - Simbólicas

El algoritmo emplea sentencias condicionales para las reglas de decisión.

- Si ocurre esto (condición).
- Entonces pasa esto (resultados).

Un ejemplo de esto podría ser lo que se están utilizando para diagnosticar afecciones en la vista. La IA ha demostrado que es posible detectar problemas oculares empleando escaso tiempo en comparación con lo que le demandaría a un profesional de la salud.

##### Segundo tipo de IA utilizadas – Redes Neuronales

Son métodos de aprendizaje automático útiles para realizar predicciones analíticas en oportunidad de crear modelos que analicen relaciones complejas. Se recomiendan para determinar relaciones no lineales en conjuntos de datos, especialmente cuando se carece de fórmula matemática conocida para el análisis de los datos. Conforme a la idea de redes interconectadas basadas en unidades simple.

##### ***Algoritmos de aprendizaje automático (Machine Learning)***

Son herramientas que analizan y predicen comportamientos basadas en inteligencia artificial. Estas pueden procesar grandes cantidades de datos para identificar patrones y anticipar acciones tanto de actores estatales como no estatales. La IA se utiliza para mejorar la toma de decisiones, detectar amenazas potenciales y optimizar operaciones militares mediante algoritmos de aprendizaje automático. Estas capacidades permiten responder rápidamente a escenarios complejos.

El proceso militar para medir la eficacia se basa en una arquitectura vertical de tipo *go-no-go*. Esto significa que una medida solo existe si cada uno de los *indicadores* de esa también existe. Del mismo modo, un indicador solo existe si todos los efectos de ese indicador también están presentes. Es un proceso binario de todo o nada que puede aplicarse a la IA tan fácilmente como a la actividad militar convencional.

Es un tipo de IA que recibe datos nuevos constantemente y de allí la maquina “aprende” generando nuevos modelos para poder realizar predicciones cada vez más precisas o analizar voces y datos.

En resumen, realiza un proceso estadístico para reconocer patrones en los datos y realizar predicciones futuras.

### *Análisis predictivo*

El sitio Amazon Web Services [AWS] plantea que:

El análisis predictivo consiste en estudiar los datos históricos y actuales para hacer predicciones sobre el futuro. Usar una mezcla de técnicas matemáticas, estadísticas y de machine learning avanzadas para analizar los datos y así determinar y extrapolar las tendencias ocultas.

Muchas empresas y organizaciones usan el análisis predictivo para guiar decisiones futuras. Por ejemplo, los analistas de marketing usan el análisis predictivo para determinar futuras ventas de sus productos, las estaciones meteorológicas lo usan para predecir el tiempo y los corredores de bolsa lo hacen para maximizar las ganancias de la compra y venta, Detección de fraude, Predicción financiera (pp.1-2)

### *Procesamiento del lenguaje natural (NLP)*

Es una tecnología de machine learning que brinda a las computadoras la capacidad de interpretar, manipular y comprender el lenguaje humano. Hoy en día, las organizaciones tienen grandes volúmenes de datos de voz y texto de varios canales de comunicación, como correos electrónicos, mensajes de texto, fuentes de noticias en redes sociales, vídeo, audio, etc. Utilizan software de NLP para procesar de forma automática estos datos, analizan la intención o el sentimiento del mensaje y responden en tiempo real a la comunicación humana.

Las tecnologías de NLP permiten por ejemplo que los *bots* de chat y voz imiten con buena calidad la conversación entre clientes (personas reales) y la empresa. Las organizaciones utilizan los *chatbots* para optimizar la capacidad y la calidad del servicio al cliente manteniendo bajos costo en simultáneo.

Algunos ejemplos actuales:

- Historias clínicas.
- Clasificación de currículos.
- Detección de nombres de personas, entidades, lugares, etc.
- Clasificación de incidencias.
- Búsqueda en internet por voz.
- Traducción automática de comentarios en foros.

### ***Aprendizaje profundo (Deep Learning)***

En este tipo se presentan varios niveles. Es decir, utilizan multiplicidad de capas para el procesamiento de grandes volúmenes de datos. En la medida que va traspasando los estadios o capas se alcanza un mejor refinamiento y optimización en la interpretación de datos.

Al utilizar datos sin estructura fija es posible su uso en un amplio rango de sectores.

Un ejemplo es el que se utiliza para el reconocimiento facial y de voz; traducción y generación de textos, etc.

### **Tercer tipo de IA utilizadas – Estrecha o debil (ANI)**

Está desarrollada para realizar acciones con alta especificidad y eficiencia dentro de un área estrecha.

Ejemplos de ello son los motores de recomendación o clasificación de imágenes. Una aplicación distintiva es ChatGPT que se orienta a comprender y generar texto en función al procesamiento de un lenguaje natural. Pero también se sirve del aprendizaje profundo y las redes neuronales artificiales.

Ejemplo de la especificidad puede considerarse que una IA para reconocimiento facial no tendría la capacidad de traducir un texto o procesar lenguaje natural.

### **Cuarto tipo de IA utilizadas – General (AGI)**

A diferencia de la anterior lo hace en múltiples áreas intentando replicar a un ser humano. Tendría la capacidad de sentir, pensar y actuar como este.

### **Quinto tipo de IA utilizadas – Super (ASI)**

Es aquella que debería tener la habilidad de funcionar mejor que la inteligencia humana. Al igual que la AGI al momento se considera que nunca podrán ser alcanzadas completamente.

### **Automatización de la desinformación**

En cuanto a la automatización de la información, Ana Siu (2018) afirmó que:

Si bien los *bots* son una de las principales tendencias en la comunicación digital, muchos usuarios no advierten, posiblemente por desconocimiento, su poder para influir en la opinión pública mediante la difusión de desinformación y propaganda. De acuerdo con Giovanni Luca Ciampaglia, investigador de la Universidad de Indiana, los *bots* son un sencillo programa informático que maneja una o varias cuentas de redes sociales para llevar a cabo una acción determinada a través de ellas. Aunque muchos de estos programas se utilizan para marketing, el uso malicioso de *bots* supone un nuevo desafío para los periodistas (pp.1-2)

Estos elementos son fundamentales para automatizar la desinformación mediante la propagación de noticias falsas, y en ocasiones son grupos de personas que también se complementan con “personas” digitales, logrando inundar los medios de comunicación y las redes sociales. Es común encontrar personas de corta edad que

emplean su tiempo en este tipo de actividades que, claramente, carecen de un nivel mínimo de ética y moral, cuya finalidad es la de implantar una idea mediante la difusión de información falsa, logrando así alterar la percepción de la realidad y la verdad. Las actividades de los *bots*, físicos o virtuales, podrían tener gran repercusión sobre la población de un país en actos relevantes como ser la elección de sus gobernantes

### **Algoritmos**

Como bien se ha mencionado los algoritmos son una base fundamental para el funcionamiento y desarrollo de la IA. Es posible considerarlos como el motor de los diferentes vehículos que se emplean para el transporte de los datos e información.

El autor Muñoz Iturrieta (2023) plantea que los algoritmos son fórmulas matemáticas diseñadas para llevar a cabo procesos específicos y que se aplican en el funcionamiento de la tecnología (p.127). Actúan como herramientas para recopilar datos del ciberespacio, con un enfoque particular en la información proveniente de redes sociales y plataformas que facilitan la interacción entre personas.

Son también modelos dinámicos, es decir, con cada nueva información se van adaptando y perfeccionando, por lo que cuanto más información relevante se provea, mejor será el algoritmo. Aunque hay que notar que “mejor” no es equivalente a “bueno”. Un algoritmo puede funcionar tremendamente bien y ser completamente inmoral. (MUÑOZ ITURRIETA, 2023, p. 127).

## **Sección 2**

### **La opinión pública**

La opinión pública está conceptualizada como la manifestación del sentir de la mayoría de la población de una sociedad sobre un tema de interés general. Además, juega un rol trascendental en los actores políticos. Esto es debido a que su legitimidad reside en gran medida en la voluntad popular.

Cotidianamente se tiende a confundir la opinión pública con la opinión publicada, es decir, la opinión de una minoría que alcanza algún medio de difusión y posteriormente es amplificada por este, y que no necesariamente es coincidente con las apreciaciones e ideas de la mayoría. Es cierto que se presenta en la sociedad una amplia cantidad de ciudadanos que prefieren callar y, en caso de opinar, solo hacerlo de manera "políticamente correcta", posibilitando la proliferación de este tipo de contenido. Allí radica un punto pivote que posibilitará el desarrollo de esta sección de contenidos. El desafío de discernir entre la verdad de la opinión pública y lo enunciado

en medios de comunicación y redes sociales requiere estudio, preparación y capacidad de abstracción tal que disminuya la probabilidad de ser engañados.

### **Características de la opinión pública.**

- Consiste en la sumatoria de las posturas de la sociedad frente a un tema de interés, especialmente cuando existe una postura mayoritaria o dominante.
- Es cambiante y susceptible de ser influenciada por los eventos ocurridos y las informaciones expresadas en los medios de comunicación.
- Se considera un elemento de peso en el debate político, aunque no siempre es tomada en cuenta por parte de los gobernantes y los actores políticos. Por ejemplo, en una dictadura la opinión pública es ignorada y castigada.
- Puede expresarse a través de medios institucionales (como el voto), acciones de fuerza (huelgas, manifestaciones) o instancias colectivas (asambleas, publicaciones).
- Tiene consecuencias directas en el ejercicio de la política y la conducción de la nación.

(CONCEPTO, 2024)

### **Manipulación de la opinión pública con IA**

#### ***Deepfakes y videos manipulados***

Los *deepfakes* o "falsedades profundas" son archivos de vídeo, imagen o voz manipulados mediante un software de IA de modo que aparenten ser originales, auténticos y reales. Aunque los *deepfakes* existen desde finales de 1990, cobraron interés en 2017, cuando un usuario de Reddit publicó material pornográfico falso con los rostros de varias actrices famosas.

El término "*Deepfakes*" combina la palabra "*fake*" (falso) y "*deep*", proveniente de "*deep learning*" (aprendizaje profundo, que es un tipo automático de la IA). La capacidad de parecer tan reales viene dada por la posibilidad técnica de modelado que tienen los programas informáticos. Se evidencia un intento para asemejarse lo más posible al funcionamiento de las redes neuronales y del cerebro humano, facilitando que nuestros sesgos cognitivos y esquemas mentales interfieran en la interpretación.

En ocasiones se han puesto palabras en boca de una personalidad importante con el objetivo de afectar un proceso de elecciones políticas. Situaciones así se pudieron observar en las campañas del presidente de Ucrania, Zelenski, Lula da Silva o Jair Bolsonaro en Brasil durante el año 2022. También se denunciaron hechos de este tipo durante la campaña del expresidente estadounidense Donald Trump.

Los materiales *Deepfakes* pueden modificar los rasgos del rostro de una persona (o su voz) y reemplazarla por otra, generando otro archivo muy similar al verdadero. Por

ello, algunos expertos consideran que los *Deepfakes* son una evolución de las *fake news* intencionada y que busca, claramente, la manipulación de las masas gracias a la difusión de estos archivos en Internet y en las redes sociales.

### ***Usos posibles de Deepfakes y sus riesgos***

- Propagar noticias falsas.
- Generar desinformación.
- Desacreditar a una persona u organización.
- Cometer delitos relacionados con el honor, la imagen o el fraude.
- Influir en las votaciones, toma de decisiones y/o pensamiento grupal.
- Llevar a cabo venganzas.
- Manipular los mercados financieros.
- Desestabilizar las relaciones internacionales

Uno de los contenidos más famosos que orientó la atención de las personas hacia los *deepfakes* se considera que fue un video del expresidente de los Estados Unidos Barak Obama en abril del año 2018. Allí se combinó la imagen simulada de un expresidente virtual con la voz de un actor de comedia real, Jordan Peele. Este último lo hizo, en sus palabras, para concientizar sobre los peligros emergentes de esta tecnología.

La disponibilidad de los grandes modelos de lenguajes (LLM) y otras IA del tipo generativas han posibilitado el desarrollo de aplicaciones gratuitas o de escaso costo. Además, ofrecen la posibilidad de uso con pocos conocimientos por lo que facilita el acceso a una gran cantidad de personas.

Algunos de los riesgos más relevantes pueden ser la desinformación, fraude en la internet o los engaños entre otros. Esto a su vez deben competir con desarrollo tecnológico de verificación en línea.

Por lo expuesto anteriormente se puede concluir que detener o dificultar el desarrollo de estas aplicaciones tendría poco o casi nulo efecto. Es importante destacar que los denominados “contenidos sintéticos” que se producen están en una constante fluctuación entre la ética y la confianza. Es por esto por lo que lograr confianza en la sociedad es algo que estará sujeto al modo en que se usen estas herramientas. Difícilmente se pueda reducir los riesgos de una manera simple. Queda claro que es un gran desafío que se tiene por delante.

### **Conclusión parcial**

Existen variadas y posibles clasificaciones de los tipos de IA, pero lo importante es entender que las técnicas que las componen son transversales y se combinan. Ha

quedado claro el papel motriz que juegan los algoritmos traccionando en las distintas tecnologías que conforma cualquier IA.

La información es posible automatizarla mediante las aplicaciones de IA y en el ámbito de la opinión pública juega un rol preponderante. Se evidencia la tensión latente entre los *deepfakes*, la ética y la confianza para lograr la supremacía en el dominio de la opinión pública. Esta última se tergiversa fácilmente por la alta velocidad de desarrollo y el acceso libre a las aplicaciones.

Es necesario, en orden a las características de la opinión pública y los riesgos presente en el uso de los *deepfakes*, considerar como primordial la formación y capacitación en la temática para evitar caer en el engaño y manipulación.

La IA constituye una herramienta esencial y medular cuando un actor deliberadamente posee intenciones de manipular la opinión pública exacerbando o atenuando ideas. Esto es facilitado, en gran parte, por el desconocimiento general del funcionamiento de la IA a través de sus algoritmos.

Se considera que los métodos de Deep fakes son quienes tienen un mayor impacto sobre el sentir de la mayoría debido a su dificultad para la correcta identificación. Se entiende que esto puede suceder porque los algoritmos diseñados para atrapar la atención de las personas han modificado su conducta. Entonces, por lo general, se omite verificar la veracidad de lo que transita por el ciberespacio.

## Capítulo 2

### La Inteligencia Artificial y la Guerra Cognitiva

#### Sección 1

##### La Guerra Cognitiva

En una primera aproximación, se puede considerar como aquella que se realiza psicológicamente contra un pueblo, por parte de fuerzas locales o internacionales. Kaleman Levitín (2024) la define como “término asociado a los medios por los cuales un estado o un grupo de influencia manipula los mecanismos cerebrales de cognición de un enemigo a nivel militar y civil a fines de debilitarlos, penetrarlo, incluirlo, dominarlo o destruirlo”.

En la Guerra Cognitiva el cerebro es sometido a un bombardeo intenso de información hostigando y presionando a la población civil para afectarla económica y socialmente. Esto a través de una campaña mediática con informaciones falsas o tergiversadas (*fake news*), buscando desesperar a la sociedad y generar escenarios que permitan condiciones de ventaja militar para los involucrados. Mediante el empleo de los medios de comunicación tradicionales e internet, la sociedad es sometida a un aluvión de diferentes enfoques informativos. Es posible utilizar grandes cadenas informativas, cibernautas, grupos de personas comunes, colegas, amigos y familiares, etc. En el terreno militar la información es cada vez más amplificadas, variada, profunda y muchas veces hasta contradictoria.

##### Inteligencia Artificial en la Guerra Cognitiva

###### ***Generación de noticias falsas (fake news)***

A partir de la emergencia de Internet y de nuevas tecnologías de comunicación e información, las Fake News han proliferado alrededor del mundo. Las redes sociales ofrecen la posibilidad a los usuarios de ser productores y consumidores de contenidos al mismo tiempo, además de facilitar la difusión de información falsa. De esta manera, se produce un círculo vicioso en el que una noticia falsa se replica miles de veces en tal solo unos segundos.

Estas noticias son historias inventadas, sin ética, que condicionan la formación de la opinión pública y promueven falsedades que contribuyen a su deformación. Este es el resultado de la excesiva oferta de información que se puede encontrar a través Internet y las redes sociales, así como de la fragmentación de las audiencias, que a su vez provoca la división de la esfera pública.

El ejemplo más grande de la influencia que pueden llegar a tener las noticias falsas es lo que ocurrió durante la campaña presidencial norteamericana de 2016. Las Fake news que abordaron los comicios en Estados Unidos se multiplicaron en los tres meses previos a las votaciones y, en particular, en Facebook éstas tuvieron más “likes”, compartir y comentarios que las informaciones verdaderas.

### **Sesgos algorítmicos**

Esto se refiere a los errores sistémicos y repetibles de un sistema informático que crean resultados acarreado consecuencias como privilegiar a un grupo arbitrario de usuarios frente a otros. Si bien las causas que provocan estos sesgos son variables, generalmente responden a algoritmos defectuosos o parciales, falta de diversidad en los datos de entrenamiento, prejuicios humanos en el diseño, entre otros.

### **Tipos de sesgo algorítmico**

Estos pueden tener un origen consciente o inconsciente, dependiendo de la intencionalidad del actor.

- **Sesgo en la muestra:** cuando existe un desequilibrio en los receptores y la muestra no es una representación fidedigna del conjunto. Por lo tanto, los datos no representarían realmente el entorno donde se espera su funcionamiento.
- **Sesgo de exclusión:** se cercena ciertas características consideradas irrelevantes por desconocimiento en la temática.
- **Sesgo del auditor:** se presenta al momento de analizar los datos realizándolo con prejuicios o punto de vista erróneo. Esto producido por la experiencia del investigador.
- **Sesgo de prejuicios:** el desarrollador se ve influenciado, inconscientemente por sus prejuicios personales. Algo común es permitir que el objetivo condicione la muestra, por ejemplo: las encuestas.
- **Sesgo de medición:** producto al emplear un instrumento equivocado en la recolección de los datos. Ejemplo: una cámara con lente defectuoso.
- **Sesgo sociocultural:** existe una influencia de los diferentes contextos sociales, económicos y culturales que rodean a los datos posibles de ser recolectados. Esto afectará la muestra que ingresará al sistema de IA.

## **Sección 2**

### **Casos de estudio**

#### **Elecciones políticas en Estados Unidos 2016**

Juan Pablo López (2023) analizó el caso de la elección presidencial estadounidense de 2016:

En enero de 2017, a pocos días de que Donald Trump asumiera como presidente electo de los Estados Unidos, las tres principales agencias de inteligencia de dicho país (Agencia Central de Inteligencia, Buró Federal de Investigaciones y la Agencia Nacional de Seguridad) publicaron un informe en el que aseguraban la injerencia del gobierno ruso en la elección presidencial de 2016 (Dirección de Inteligencia Nacional 2017). Según el documento, el mandatario ruso Vladimir Putin habría ordenado una campaña dirigida a influenciar dicho proceso electoral con dos objetivos primordiales: debilitar la fe pública en el proceso democrático estadounidense y desacreditar a Hillary Clinton (candidata por el Partido Demócrata) para dañar sus posibilidades de ser electa. Si bien consideraron que el ataque no estuvo vinculado al sistema de cómputo de votos, el informe sí establece que se realizaron ciberataques a diversas figuras del Partido Demócrata, así como filtraciones de información sensible en diversas redes sociales y sitios como *WikiLeaks* (BBC 2017). En diciembre de 2016 la CIA también emitió un comunicado responsabilizando a Rusia de ejecutar un “kompromat” (término ruso que significa “material comprometedor”, se utiliza para describir la acción de utilizar dicho material en para crear publicidad negativa, cometer chantaje o asegurar lealtad), esta vez interviniendo los correos electrónicos de algunos miembros del equipo de campaña de Clinton (p. 14).

El informe sobre la interferencia rusa en las elecciones de 2016 destaca dos métodos principales utilizados por el Kremlin. Primero, la Agencia de Investigación de Internet (IRA) llevó a cabo operaciones de desinformación en redes sociales y plataformas como *WikiLeaks* para generar discordia en la sociedad estadounidense.

Dentro de las tácticas implementadas por la IRA se encontraban:

- Publicaciones impulsadas a población afroamericana incentivando a boicotear la elección o promoviendo errores procedimentales en el ejercicio del voto,
- Fanpages, comunidades y grupos dirigidos a comunidades de extrema derecha y
- Difusión de noticias de contenido sensacionalista, conspirativo o con información errónea a votantes de todo el espectro político.

Aunque no se halló evidencia de coordinación entre el entorno de Trump y la IRA, sí se presentaron cargos criminales contra individuos y empresas rusas vinculadas a estas actividades.

La figura siguiente presenta un recorte de imágenes utilizada en el grupo “Being Patriotic”, uno de los más populares en materia de interacciones dentro de Facebook, el cual según Howard et al. (2018) fue creado por IRA y estuvo en actividad entre junio de 2015 y principios de 2017, alcanzando más de 1.5 millones de interacciones tan solo en octubre de 2016.



*Imágenes publicadas en la fanpage de Facebook "Being Patriotic".*

El segundo método incluyó operaciones de hackeo por parte de actores rusos que accedieron a correos electrónicos de personas cercanas a Hillary Clinton y al Partido Demócrata, con el fin de diseminar información sensible.

Esto llevó a la presentación de cargos contra funcionarios militares rusos por conspiración e interferencia electoral. Sin embargo, no se encontró evidencia concluyente de una conspiración coordinada entre la campaña de Trump y Rusia, a pesar de las supuestas ofertas del Kremlin.

### **Estallido social chileno en 2019. Uso de bots y redes sociales en el contexto de una crisis social**

El aumento del boleto del subterráneo efectuado el 6 de octubre de 2019 fue la Uchispa que dio comienzo al estallido social en Chile. Una ola de manifestaciones por un modelo económico más justo y derechos sociales derivó en las peores protestas en el país de los últimos tiempos. En el contexto de los graves incidentes que estaban teniendo lugar, el Ministerio del Interior anunciaba el 19 de diciembre de 2019, que había entregado al Ministerio Público un informe de 112 páginas, en que se resume el comportamiento de las redes sociales durante el estallido social.

En este, el ministro del Interior, Gonzalo Blumel, señaló que existía "información extraordinariamente sofisticada a partir del análisis de *Big data*, con tecnologías de la información, que dan cuenta de antecedentes importantes para la investigación penal". Haciendo clara referencia a las causas abiertas por el Ministerio Público ante los graves incidentes. Estos incluían el incendio de autobuses, edificios, hoteles, iglesias, saqueos a supermercados y locales comerciales. Especialmente el motivo relevante era la destrucción de 78 estaciones de metro en un espacio de tiempo de 48 horas, en

Santiago, durante los primeros días de la crisis social. En dicho informe se analizaron sesenta millones de comentarios publicados en Twitter por cinco millones de usuarios, muchos de ellos de cuentas extranjeras. Asimismo, el informe destacaba que había identificado la participación de más de 11.000 usuarios considerados bots o usuarios con actividad anómala cuyas cuentas fueron creadas después de 18 de octubre de 2019, es decir, una vez iniciadas las protestas sociales. Por este motivo, el gobierno chileno propuso una teoría conspirativa para la crisis social, organizada por grupos antisistema, radicales y apoyados por terceros países, como Venezuela o Cuba, y presentaba el informe como prueba de ello.

Sin embargo, el fiscal Nacional rápidamente puso en duda la seriedad de este informe al señalar que no se trataba de uso de inteligencia, sino de fuentes abiertas, y fundado en el análisis de redes sociales y de la prensa nacional y extranjera. Asimismo, expresaba no haber encontrado “ningún valor en la información que contenía”, y que aportaba pocos antecedentes que pudieran “hacer valer en juicio”.

Este caso presenta un problema complejo, derivado del uso de las redes sociales y la tecnología de IA, para convocar, organizar, promover o dirigir crisis sociales, pero también como herramienta para investigaciones de los gobiernos que, a través de análisis de *big data*, y justificados en razones de seguridad y orden público, juegan en las fronteras de la libertad de expresión de las personas y el derecho a la protesta pacífica.

### **Conflicto Rusia - Ucrania**

Podemos definir dicho conflicto entre ambos estados como del tipo híbrido por la multiplicidad de elementos intervinientes y la respectiva variedad de acciones, en las que se observan fracciones militares, organizaciones legales, ilegales e individuos solitarios.

La guerra entre Ucrania y Rusia ha introducido una dimensión compleja en el conflicto contemporáneo, donde la IA se utilizó ampliamente para moldear y manipular la opinión pública en ambos países y en la comunidad internacional. El uso estratégico de la IA ha facilitado la creación y difusión de propaganda, desinformación y narrativas polarizadas, influyendo en las percepciones públicas de manera decisiva y generando consecuencias profundas en la estabilidad social y en las políticas de ambos países.

Rusia está utilizando la IA generativa para intensificar las campañas de desinformación contra Ucrania. Esto fue enunciado por un alto funcionario ucraniano en los márgenes de una conferencia cibernética.

Al respecto, el viceministro de Relaciones Exteriores de Ucrania Anton Demokhin manifestó que se ha podido observar que la densidad de los ciberataques (de Rusia)

continúa, pero poniendo mayor acento en la desinformación. "Ahora que llegó la IA generativa, eso permite que las narrativas de desinformación se multipliquen y distribuyan en un nuevo nivel más complejo", añadió.

Estas nuevas campañas de desinformación son difíciles de detectar, dijo, con una amplia actividad falsa en las plataformas de redes sociales para aumentar su credibilidad.

En Rusia, el uso de la IA ha sido integral en las campañas de propaganda para consolidar el apoyo interno y justificar las acciones militares en Ucrania. Mediante el empleo de algoritmos de aprendizaje automático y redes neuronales, los medios rusos, controlados en su mayoría por el Estado, han diseñado y distribuido mensajes cuidadosamente adaptados para diferentes grupos de población.

Estas campañas digitales emplean tanto *bots* como cuentas automáticas para amplificar mensajes favorables al gobierno y silenciar narrativas disidentes, generando un ambiente de cohesión interna que apoya la intervención militar en Ucrania.

Funcionarios ucranianos dijeron en agosto pasado que muchos conciudadanos estaban siendo blanco de la desinformación en línea del servicio de seguridad y la agencia de inteligencia militar rusa FSB. En respuesta, Moscú, acusó a Ucrania y Occidente de desatar una sofisticada guerra informativa contra Rusia.

### ***Desinformación y narrativas controladas***

Los sistemas de IA han permitido a las agencias rusas crear y distribuir noticias falsas y videos manipulados o "deepfakes" que presentan una realidad distorsionada de la guerra. Allí las acciones rusas son justificadas como una defensa legítima ante una amenaza percibida. Esto se ve acentuado por el uso de algoritmos que refuerzan narrativas progubernamentales en plataformas digitales limitando la exposición de la población rusa a perspectivas alternativas. Las consecuencias de este tipo de manipulación en la opinión pública han sido notorias, ya que gran parte de la población percibe el conflicto bajo la óptica del nacionalismo y la defensa de la soberanía, lo que reduce la oposición interna y refuerza el apoyo a las políticas.

### ***Polarización y represión del gobierno ruso***

El gobierno ruso ha utilizado la IA para monitorear y controlar las redes sociales, identificando y censurando de manera efectiva a personas y grupos críticos. La capacidad de los sistemas de IA para rastrear patrones de comportamiento permite al gobierno responder rápidamente a cualquier intento de movilización social en contra del conflicto. Esta vigilancia digital, combinada con la represión de voces disidentes, ha generado un ambiente en el que la disidencia es limitada, lo cual ha polarizado aún más a la sociedad rusa. La percepción pública ha quedado dividida entre quienes apoyan

activamente las políticas del Kremlin y quienes, temerosos de represalias, optan por la neutralidad o el silencio.

### ***Influencia de la IA en la opinión pública en Ucrania***

Por su parte, Ucrania ha utilizado la IA de forma diferente, enfocándose en la construcción de una narrativa de resistencia y solidaridad, tanto a nivel nacional como internacional. Desde el inicio del conflicto, el gobierno ucraniano ha implementado campañas que utilizan IA para contrarrestar la propaganda rusa y fortalecer el espíritu patriótico, buscando unir a la población en una causa común y sensibilizar a la comunidad global sobre los desastres de la guerra.

La IA ha sido clave en la generación de campañas de influencia en redes sociales, donde bots y cuentas automatizadas transmiten en tiempo real información sobre el conflicto, destacando los daños civiles y los ataques contra infraestructura ucraniana. Estas campañas han generado una gran repercusión en la opinión pública mundial y han contribuido a que gobiernos extranjeros apoyen a Ucrania, ya sea mediante sanciones a Rusia o con ayuda militar y humanitaria.

El uso de la IA permite diseñar mensajes impactantes y adaptar el contenido a las tendencias de opinión, garantizando que el conflicto siga siendo un tema relevante en la agenda pública.

Dentro de Ucrania, el gobierno ha empleado la IA para monitorear y contrarrestar la desinformación, asegurándose de que la población acceda a información que favorezca el sentido de unidad y la resistencia principalmente frente al invasor. Los sistemas de IA identifican rápidamente contenido que podría erosionar la moral y, mediante campañas específicas, contrarrestan cualquier intento de generar pánico o desaliento. Esto ha contribuido a una mayor cohesión social, reforzando el compromiso de la población con la defensa de la soberanía nacional y generando una narrativa de heroísmo, fomentando la unión nacional.

### ***Consecuencias Comunes en Ambos Países***

El uso de la IA en la manipulación de la opinión pública ha tenido repercusiones negativas en ambos países, reflejándose en una creciente polarización social y una erosión de la confianza en la información disponible.

En Rusia, la constante censura y manipulación mediática han generado un ambiente de escepticismo, donde la población no puede confiar plenamente en las fuentes oficiales de información. La censura, sumada a las limitadas perspectivas alternativas, ha generado un aislamiento de la opinión pública rusa respecto al resto del mundo.

En Ucrania, aunque la IA ha sido una herramienta clave para unificar el país, la constante exposición a imágenes de guerra y el enfoque patriótico también pueden

generar una sociedad altamente polarizada, donde el discurso de resistencia se convierte en la única narrativa aceptable. La propaganda de guerra, aun cuando se justifica como defensa, puede llevar a que las futuras generaciones crezcan con percepciones hostiles y estereotipos que limitan la posibilidad de reconciliación y diálogo en el futuro.

### **Sección 3**

#### **Desestabilización del adversario**

##### **Estrategias de desestabilización**

La Guerra Cognitiva puede considerar sus orígenes durante la Guerra Fría. Dentro de ella, las Operaciones Psicológicas (PsyOps) se centraban en el uso de la propaganda, las cuales podían ser blanca, gris y negra, según la fuente de origen: oficial, ambigua y hostil, respectivamente.

Sus agentes pueden propagar información falsa o engañosa, a través de cuentas que parecen e interactúan en forma humana, buscando controlar a individuos, poblaciones y cómo reaccionan a la información presentada. Este tipo de guerra se enfoca en alterar cómo piensa una población objetivo y cómo actúa; sus objetivos son desestabilizar e influenciar.

La desestabilización de poblaciones como primer objetivo se logra acelerando divisiones dentro de los grupos de la población, e introduciendo ideas diseñadas para enfrentar grupos unos contra otros y aumentar la polarización de su conjunto o sus líderes individuales.

El segundo objetivo de la Guerra Cognitiva es, influir en la población blanco. Esto se logra mediante la manipulación de la interpretación de un objetivo y la comprensión del mundo que les rodea. La intención final es que el grupo blanco tenga ideas afines sobre un tema, es decir, generar consenso entre una población, con suficiente poder para efectuar un cambio de paradigma. Las campañas de Guerra Cognitiva pueden lanzarse con la intención de reclutar civiles para una causa.

Intentar influir en el público no es nuevo. Políticos, generales, el mercado, líderes y otras personas influyentes, han utilizado la retórica, la propaganda y los mensajes, para manipular a la opinión pública durante años. Lo nuevo son las herramientas para hacerlo: el internet, las redes sociales y el ciclo de noticias de veinticuatro horas, que permiten flujos de información, que hacen más fácil influir en la mente humana.

##### **El rol de los actores estatales y no estatales:**

Antes de profundizar en el papel de los actores estatales y no estatales en la guerra cognitiva, se considera oportuno, en primera instancia, definirlos.

En el primer caso, se hace referencia a las autoridades gubernamentales de un país. Por el contrario, los actores no estatales son individuos u organizaciones que no forman parte del aparato estatal pero que tienen influencia política (empresarios, sindicatos, ONG, movimientos sociales, etc.).

En cuanto a la amenaza en su futuro potencial, se prevé que no serán los actores no estatales los que liderarán campañas de Guerra Cognitiva, sino los actores estatales; esto se deberá a su creciente interés en la capacidad de producir contenido, que podrá hacerse pasar por verdadero. Distintos gobiernos, agencias de inteligencia y grupos militares han empleado tácticas de desinformación y propaganda para promover sus intereses geopolíticos, desestabilizar a adversarios y minar la confianza en instituciones democráticas.

## **Ética y Desafíos Regulatorios**

### ***Dilemas éticos***

Las aplicaciones emergentes de la IA y el aprendizaje automático generan planteos éticos en el debate público. Un aspecto común de los principios generales en relación con la IA formulados y acordados por los distintos actores estatales y no estatales es la importancia que reviste el elemento humano para garantizar el cumplimiento jurídico y la aceptabilidad ética (esto es, la evaluación que permite identificar que una acción sea moralmente justificable y se encuadre en principios y valores éticos). Al respecto, se han redactado tratados y acuerdos en distintas partes del mundo, que ponen de relieve la importancia de que la IA esté en consonancia con los valores humanos.

### ***Regulación y gobernanza***

Para garantizar el respeto del derecho, las partes en un conflicto deben mantener el control humano cuando se aplican la IA y el aprendizaje automático como medios y métodos de guerra. Las normas del derecho internacional humanitario están dirigidas a los seres humanos. Son ellos quienes cumplen e implementan el derecho, y también los responsables en caso de violaciones de las normas. En particular, incumbe a los combatientes ejercer su criterio en virtud de las normas del derecho internacional humanitario que rigen la conducción de las hostilidades, y esta responsabilidad no puede transferirse a una máquina, un software o un algoritmo.

Estas normas exigen criterios específicos a la luz del contexto por parte de quienes planifican, deciden y realizan ataques, a fin de garantizar los siguientes principios:

- Distinción entre objetivos militares, que pueden ser atacados lícitamente, y personas civiles y bienes de carácter civil, que no deben ser atacados;

proporcionalidad, para garantizar que los daños colaterales no resulten excesivos en relación con la ventaja militar directa y concreta que se prevé alcanzar; y

- Precauciones en el ataque, de modo que los riesgos para las personas civiles puedan reducirse al mínimo.

Cuando se emplean sistemas de IA en ataques, ya sea como parte de sistemas de armas físicas o cibernéticas, o en sistemas de apoyo a la toma de decisiones, su concepción y uso deben permitir que los combatientes ejerzan ese criterio.

Con respecto a los sistemas de armas autónomos, los Estados Partes en la Convención sobre ciertas armas convencionales (CCA) han reconocido que “el ser humano debe mantener la responsabilidad” respecto de decisiones sobre el uso de sistemas de armas y el uso de la fuerza, y muchos Estados, organizaciones internacionales y organizaciones de la sociedad civil han enfatizado la exigencia del control humano para garantizar el cumplimiento del derecho internacional humanitario y la compatibilidad con los valores éticos.

### **Futuro de la IA en la Guerra Cognitiva**

La profesora y abogada por la Universidad de Buenos Aires Marta Rosa Vigevano (2021) publicó:

La IA tiene aplicación en varios aspectos de vital importancia en la conducción de hostilidades. En este contexto, se utiliza en tareas de reconocimiento de objetivos militares, vigilancia, comunicación, logística, en la manipulación de la información y el desarrollo de nuevas armas y los métodos de combate. Sin embargo, esta misma tecnología también encuentra su uso humanitario en la protección de las víctimas. En el ámbito terrestre, por ejemplo, se utilizan artefactos autónomos en el seguimiento de vehículos o, para evitar obstáculos, coordinando sistemas no tripulados con vehículos tripulados. En el ámbito marítimo esta tecnología posibilita que los navíos de superficie no tripulados, ligeros y pequeños estén equipados con instrumentos de barrido de minas marinas. En el ámbito aéreo, son utilizados con fines humanitarios los sistemas autónomos de aeronaves de vigilancia o reconocimiento. En cuanto a las redes de información, el caso del acceso a una gran cantidad de datos y conocimiento en el ciberespacio permite a individuos y grupos un acceso inmediato a recursos estratégicos.

Tal y como lo plantea la docente, el futuro de la IA presenta diferentes aristas, pero señala también que serán los mecanismos para generar información falsa como método de combate quienes tomen mayor relevancia en la Guerra Cognitiva, ya que con su avance resulta cada vez más complicado distinguir la información real de la falsa (Vigevano 2021).

### **Conclusión parcial**

La GC como tal no es algo nuevo, sino que evidentemente ha cambiado el medio a través del cual actúa. Para el tiempo presente es claro que la herramienta por excelencia es la inteligencia artificial. Se ha observado que la generación de noticias falsas requiere de un conocimiento detallado de los posibles sesgos cognitivos que pudieran presentarse de manera consciente o inconsciente.

En las elecciones presidenciales de Estados Unidos durante el año 2016 se habría configurado principalmente las operaciones de desinformación orientadas a dividir a la sociedad sobre la plataforma de las redes sociales y los medios de comunicación social. Algo similar se pudo haber desarrollado en el estallido social chileno, pero con efectos nocivos en la sociedad al ocasionar serios daños personales y materiales.

Un párrafo aparte merece el conflicto Rusia – Ucrania donde, en la lógica evolución por el tiempo transcurrido, la IA tuvo un vector destacado el cual es la manipulación de la opinión pública por ambos bandos. Se observó que el objetivo final era obtener la legitimidad de las acciones.

Los tres hechos diferentes en su naturaleza presentan un posible denominador común el cual es la participación de actores estatales y no estatales. Todos con el objetivo de modificar la percepción de la realidad empleando diferentes métodos.

Se considera que lograr estabilizar y equilibrar; la ética, el empleo de la IA y el plexo normativo legal es uno de los mayores desafíos para el presente y el futuro.

No es posible determinar el grado de desestabilización a lograr con el empleo de la IA debido a la multiplicidad de factores que deberían analizarse y lo cual excede ampliamente la profundidad del presente estudio. Pero si se determinó que está influyendo considerablemente en la ética porque a la luz de los hechos obliga a replantear permanentemente los límites para el desarrollo y empleo de la tecnología de la IA.

## Conclusión

Se identificó que la combinación de redes neuronales y aprendizaje de las máquinas son herramientas de uso masivos en los conflictos recientes. Dentro de ello un elemento fundamental es el uso específico de *bots* para operar en los medios de comunicación y las redes sociales orientados a crear las noticias falsas.

Si bien la información consultada dificulta aseverar con precisión es posible identificar los siguientes aspectos:

En las elecciones presidenciales norteamericanas del 2016 se infiere que un actor estatal externo habría empleado redes neuronales dirigido contra la población. A la opinión pública se la intentó modificar mediante generación de noticias y desinformación.

Del conflicto social chileno del 2019 se pudo concluir que necesariamente se debió haber empleado machine learning para poder procesar la big data. La automatización de la información mediante uso de personas digitales habría sido el factor multiplicador dentro de las redes sociales.

Es para destacar del enfrentamiento RUSIA – UCRANIA la clara tensión por dominar la guerra de propaganda. Se entiende que es utilizado por el más débil como un elemento potenciador de poder en esta desequilibrada relación. Las principales acciones desarrolladas se pueden clasificar dentro del grupo machine learning empleando procesamiento del lenguaje natural direccionada hacia los sentimientos de las personas.

Los tres conflictos analizados presentan un objetivo de similares características hacia donde se direccionan sus acciones más relevantes: la población civil. Estratégicamente se puede considerar como maniobra indirecta entre los actores.

Ha sido difícil mensurar el grado preciso de desestabilización estratégica alcanzado por los oponentes debido a la poca profundidad y escaso tiempo para el análisis.

Por otro lado, la aplicación de IA por parte de agentes estatales y no estatales escaló a niveles que los obligaron a rever la aceptabilidad ética de su uso, estableciendo regulaciones en tratados y acuerdos para que la IA esté en consonancia con los valores humanos, independientemente de su aplicación en la Guerra Cognitiva con el fin de desestabilizar adversarios.

Finalmente concluimos que se deberá controlar exhaustivamente el uso de esta herramienta y su proliferación en nuestras actividades diarias, ya que su excesivo uso distorsiona la realidad, generando conflictos y favoreciendo a aquellos que se benefician con la polarización de la población

## **BIBLIOGRAFÍA**

Kaleman Levitín, M (2024). *Guerra Cognitiva* [Diapositivas de Powerpoint]. Instituto de Ciberdefensa de las Fuerzas Armadas.

[https://drive.google.com/drive/folders/1cYaEF5WWLSDqByPiGGAwL7d5TCrOo8\\_4](https://drive.google.com/drive/folders/1cYaEF5WWLSDqByPiGGAwL7d5TCrOo8_4)

Coeckelbergh, M. (2020). *AI Ethics*. Cambridge: MIT Press.

Muñoz Iturrieta, P. (2023). *Apaga el celular y encendé el cerebro*. Hojas del sur S.A.

Liang, Q. y Xiangsui, W. (1999). *Guerra por encima del límite*. La Casa de Publicaciones de Arte y Literatura del EPL, Beijing.

Rid, T. (2020). *Active Measures: The Secret History of Disinformation and Political Warfare*. Farrar, Straus and Giroux

Siu, A. (30 de octubre de 2018). *Cómo funcionan los bots (y cómo contribuyen a difundir información falsa)*. International Journalists' Network.

<https://ijnet.org/es/story/c%C3%B3mo-funcionan-los-bots-y-c%C3%B3mo-contribuyen-difundir-informaci%C3%B3n-falsa>

Amazon Web Services (2023). *¿Qué es el análisis predictivo?* AWS.

<https://aws.amazon.com/es/what-is/predictive-analytics/>

Concepto (2024, sección sociedad). *Opinión pública*. <https://concepto.de/opinion-publica/>

Burda, R. (4 de junio de 2023). *Cognitive Warfare as Part of Society: Never-Ending Battle for Minds*. The Hague Centre for Strategic Studies. [https://hcss.nl/wp-content/uploads/2023/06/04-](https://hcss.nl/wp-content/uploads/2023/06/04-Cognitive-Warfare-as-Part-of-Society-Never-Ending-Battle-for-Minds.pdf)

[Cognitive Warfare as Part of Society Never Ending Battle for Minds.pdf](https://hcss.nl/wp-content/uploads/2023/06/04-Cognitive-Warfare-as-Part-of-Society-Never-Ending-Battle-for-Minds.pdf)

Bostick, R. (21 de julio de 2011). *Initiating the Cognitive Revolution: An Examination of Special Operations Military Information Support Operations*. U.S. Army War College Civilian Research Project.

<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c82c2e81c63c5ae70bf4b09439d2841059525c0b>

Olejnik, L. (2025). *Propaganda: From Disinformation and Influence to Operations and Information Warfare*. Google Books.

[https://books.google.com.ar/books?hl=es&lr=&id=Cq4XEQAAQBAJ&oi=fnd&pg=PT18&dq=The+Cognitive+War:+Strategic+Influence+and+Information+Operations.&ots=CfaPpYiK4j&sig=\\_wgStVNZIDJPNczkqooDZXdSio8#v=onepage&q&f=false](https://books.google.com.ar/books?hl=es&lr=&id=Cq4XEQAAQBAJ&oi=fnd&pg=PT18&dq=The+Cognitive+War:+Strategic+Influence+and+Information+Operations.&ots=CfaPpYiK4j&sig=_wgStVNZIDJPNczkqooDZXdSio8#v=onepage&q&f=false)

Peláez Martínez, J. (4 de mayo de 2022 sección opinión). *Evolución hacia la guerra cognitiva*. Movimiento Antorchista Nacional

<https://mail.movimientoantorchista.org.mx/evolucion-hacia-la-guerra-cognitiva>

Carnegie Council por la Ética en los Asuntos Internacionales, (9 de abril de 2024).

*Algorithms of War: The Use of AI in Armed Conflict*. Carnegie Council.

<https://es.carnegiecouncil.org/media/video/algorithms-war-ai-armed-conflict>

Arreola, A. (15 de agosto de 2024). *Inteligencia Artificial y Desinformación: Papel*

*en los Conflictos del Siglo XXI*. Centro de Estudios Estratégicos del Ejército del

Perú. [https://ceep.mil.pe/2024/08/15/inteligencia-artificial-y-desinformacion-papel-en-](https://ceep.mil.pe/2024/08/15/inteligencia-artificial-y-desinformacion-papel-en-los-conflictos-del-siglo-xxi/)

[los-conflictos-del-siglo-xxi/](https://ceep.mil.pe/2024/08/15/inteligencia-artificial-y-desinformacion-papel-en-los-conflictos-del-siglo-xxi/)

López J. (4 de mayo de 2023). *Campañas de Desinformación y respuesta democrática:*

*el caso de las elecciones estadounidenses de 2016*. [Tesina de investigación,

Universidad Nacional de San Martín] *Repositorio Institucional de la UNSAM*.

[https://ri.unsam.edu.ar/bitstream/123456789/2169/1/TLIC\\_EPYG\\_2023\\_LJP.pdf](https://ri.unsam.edu.ar/bitstream/123456789/2169/1/TLIC_EPYG_2023_LJP.pdf)

Gardel, L. (19 de julio de 2023). *Qué es un deepfake y en qué se diferencia de un*

*video manipulado*. Chequeado.

[https://chequeado.com/el-explicador/que-es-un-deepfake-y-en-que-se-diferencia-de-](https://chequeado.com/el-explicador/que-es-un-deepfake-y-en-que-se-diferencia-de-un-video-manipulado/)

[un-video-manipulado/](https://chequeado.com/el-explicador/que-es-un-deepfake-y-en-que-se-diferencia-de-un-video-manipulado/)

Parra Sepúlveda, D. y Concha Machuca, R. (29 de octubre de 2021). *Inteligencia*

*artificial y derecho. Problemas, desafíos y oportunidades*. Pontificia Universidad

Javeriana. <https://www.redalyc.org/journal/825/82569129005/html/>

Ibrahim, S. (23 de agosto de 2021). *Cómo los 'deepfakes' están*

*cambiando nuestra visión de la realidad*. Swissinfo.ch.

<https://www.swissinfo.ch/spa/ciencia/c%3%b3mo-los-ultrafalsos-cambian-nuestra-visi%3%b3n-de-la-realidad/46866008>

Vigevano, M. (abril de 2021). *Inteligencia artificial aplicable a los conflictos armados: límites jurídicos y éticos*. *Revista Arbor*.

<https://arbor.revistas.csic.es/index.php/arbor/article/view/2417/3638>

Cruz Roja Internacional (junio de 2019). *La inteligencia artificial y el aprendizaje automático en los conflictos armados: un enfoque centrado en las personas*.

International Review of the Red Cross, N.º 913 Las tecnologías digitales y la guerra.

[https://international-review.icrc.org/sites/default/files/reviews-pdf/2022-05/913\\_IA\\_OK.pdf](https://international-review.icrc.org/sites/default/files/reviews-pdf/2022-05/913_IA_OK.pdf)

Lozano, M. (s.f.). *Guerra Cognitiva: Narrativas y Desinformación en el siglo 21*. IT CONNECT LATAM. <https://itconnect.lat/portal/guerra-cognitiva-001/>

Mártir Hidalgo, J. (21 de febrero de 2023). *Pensamiento Crítico. Guerra cognitiva: un ataque a la verdad y lucha por la información - Resumen Latinoamericano*.

*Resumen Latinoamericano*.

<https://www.resumenlatinoamericano.org/2023/02/21/pensamiento-critico-guerra-cognitiva-un-ataque-a-la-verdad-y-lucha-por-la-informacion/>

Moressi, A. (2 de mayo de 2023). *El impacto de la Inteligencia Artificial en la Defensa*. *Pucará Defensa*. <https://www.pucara.org/post/el-impacto-de-la-inteligencia-artificial-en-la-defensa>

Crosby, C.(s.f.). *Operativizar la inteligencia artificial para la guerra algorítmica*. *Army University Press*.

<https://www.armyupress.army.mil/Journals/Edicion-Hispanoamericana/Archivos/Primer-Trimestre-2021/Operativizar-la-inteligencia-artificial-para-la-guerra-algoritmica/>

Upbe, (s.f.). *Procesamiento del Lenguaje Natural (NLP). ¿Qué es?*.

<https://www.upbe.ai/blog/procesamiento-del-lenguaje-natural-nlp-que-es/>

Amazon Web Services (s.f.). *¿Qué es el procesamiento de lenguaje natural (NLP)?*. AWS. <https://aws.amazon.com/es/what-is/nlp/>

LISA Institute (s.f.). *¿Qué son las Fake News o noticias falsas? Guía Práctica de Falsas técnicas y herramientas para detectarlas*. LISA Institute.

<https://www.lisainstitute.com/blogs/blog/como-detectar-fake-news-que-son-noticias-falsas>

Siu, A. (30 de octubre de 2018). *Cómo funcionan los bots (y cómo contribuyen a difundir información falsa)*. International Journalists' Network.

<https://ijnet.org/es/story/c%C3%B3mo-funcionan-los-bots-y-c%C3%B3mo-contribuyen-difundir-informaci%C3%B3n-falsa>

Universidad de Malaga (s.f.). *Los primeros programas capaces de procesar lenguaje Natural*. Master de Formación Permanente en Big Data e Inteligencia Artificial.

<https://www.bigdata.uma.es/los-primeros-programas-capaces-de-procesar-lenguaje-natural/>

Matiz-Rojas, A. y Fernandez-Camargo, J. (2023). *Del uso de la inteligencia artificial como medio y método en los conflictos armados* [Revista Científica General José María Córdova](#).

<https://revistacientificaesmic.com/index.php/esmic/article/view/1151>

Reuters (16 de octubre de 2024). *Rusia utiliza la IA generativa para aumentar la desinformación: alto funcionario ucraniano*, La Nación.

<https://www.lanacion.com.ar/agencias/rusia-utiliza-la-ia-generativa-para-aumentar-la-desinformacion-alto-funcionario-ucraniano-nid16102024/>

Google Cloud (s.f.). *¿Qué es el análisis predictivo?*

<https://cloud.google.com/learn/what-is-predictive-analytics?hl=es>