IA Autónoma en la Guerra Cognitiva: Nuevas Fronteras del Conflicto Silencioso

Autonomous AI in Cognitive Warfare: New Frontiers of Silent Conflict

Juan de Dios Meseguer González
IUGM-UNED, 46004, Valencia, España.
jmeseguer1@alumno.uned.es - https://orcid.org/0009-0004-3691-4880

Resumen

El artículo examina el papel emergente de los sistemas de Inteligencia Artificial (IA) autónomos en la guerra cognitiva, entendida como el conjunto de operaciones orientadas a influir en los procesos mentales de individuos y colectivos. Estos sistemas no solo ejecutan tareas automatizadas, sino que también se adaptan a contextos dinámicos para maximizar su impacto psicológico. El objetivo principal es analizar cómo estas tecnologías amplían el dominio de la confrontación más allá de los espacios físicos y digitales, incidiendo directamente en la percepción, el juicio y la toma de decisiones humanas. La metodología empleada es cualitativa, basada en el análisis documental, el estudio de casos recientes y la elaboración de escenarios prospectivos. Los resultados muestran que la IA autónoma amplifica la efectividad de las campañas de manipulación informativa, permite una personalización masiva del mensaje y optimiza de forma adaptativa las tácticas de influencia. El artículo concluye que esta nueva frontera del conflicto silencioso plantea desafíos estratégicos, éticos y jurídicos para los Estados que aún carecen de marcos normativos adecuados para afrontarlos. El hallazgo más innovador es la capacidad de estas tecnologías para operar de forma independiente en entornos cognitivos sin supervisión humana directa, lo que redefine las reglas tradicionales del poder en el entorno de seguridad internacional.

Palabras clave: *IA, guerra cognitiva, autonomía, percepción, conflicto silencioso.*

Abstract

The article examines the emerging role of autonomous artificial intelligence systems in cognitive warfare, understood as a set of operations aimed at influencing the mental processes of individuals and groups. These systems not only perform automated tasks but also adapt to dynamic contexts to maximize their psychological impact. The primary objective is to examine how these technologies extend the domain of confrontation beyond physical and digital spaces, thereby directly influencing human perception, judgment, and decision-making. The methodology used is qualitative, based on document analysis, recent case studies, and the development of prospective scenarios. The results show that autonomous artificial intelligence enhances the effectiveness of information manipulation campaigns, enables massive message personalization, and optimizes influence tactics adaptively. The article concludes that this new frontier of silent conflict poses strategic, ethical, and legal challenges for States, which still lack adequate regulatory frameworks to address it. The most innovative finding is the ability of these technologies to operate independently in cognitive environments without direct human supervision, thereby redefining the traditional rules of power in the international security landscape.

Keywords: *Artificial Intelligence, cognitive warfare, autonomy, perception, silent conflict.*

Fecha de Recepción: 15/12/2024 - Aceptado: 20/12/2024 – Publicado: 31/12/2024
ISSN: 2477-9253 – DOI: http://dx.doi.org/10.24133/RCSD.VOL09.N04.2024.01

I. Introducción

En las últimas décadas, el campo de la IA evoluciona desde sistemas pasivos orientados a la automatización de tareas específicas hasta agentes autónomos con capacidad de aprendizaje y adaptación en entornos complejos. Esta transformación tecnológica genera importantes implicaciones en el ámbito de la defensa, particularmente en el emergente concepto de guerra cognitiva, entendida como el uso de técnicas para influir en los procesos mentales de individuos y sociedades con fines estratégicos. En este nuevo escenario, el combate no solo se libra en terrenos físicos o cibernéticos, sino también en el espacio mental, donde se disputan percepciones, emociones, ideas y decisiones.

Diversos estudios (Gizewski & Rostek, 2021; Rigaki & Garcia, 2020) abordan el uso de algoritmos para manipular la información y direccionar el comportamiento humano. Sin embargo, la participación de sistemas de IA autónomos en este tipo de conflicto representa aún una frontera poco explorada y escasamente regulada.

El propósito de este estudio consiste en analizar cómo los sistemas autónomos de IA transforman la naturaleza de la guerra cognitiva, ampliando las capacidades de influencia estratégica sin necesidad de intervención humana directa. Para ello, la investigación adopta un enfoque cualitativo centrado en el análisis documental, el estudio de casos contemporáneos y la construcción de escenarios prospectivos. El resultado principal consiste en un marco conceptual que articula la relación entre autonomía tecnológica, operaciones psicológicas y manipulación cognitiva, sustentado en evidencia empírica y un análisis interdisciplinario. La investigación aplica el método analítico-descriptivo, herramientas de análisis crítico y técnicas de contraste entre casos reales y literatura académica especializada.

Entre los principales hallazgos, se identifica que la IA autónoma permite una personalización masiva de mensajes, una adaptación en tiempo real de tácticas persuasivas y una creciente capacidad para operar sin supervisión humana en entornos cognitivos. Esta capacidad redefine el papel de los actores tradicionales del conflicto y altera el equilibrio de poder en los escenarios de seguridad internacional. En ausencia de marcos éticos y jurídicos adecuados, la autonomía tecnológica en el plano cognitivo representa un riesgo emergente para la soberanía informativa de los Estados. La principal contribución del estudio radica en visibilizar esta problemática desde una perspectiva multidisciplinar, anticipando los desafíos que plantea esta nueva frontera del conflicto silencioso.

El artículo se estructura en cinco secciones. En primer lugar, presenta el marco teórico y conceptual sobre guerra cognitiva e IA autónoma. En segundo lugar, analiza el estado del arte y las respuestas institucionales existentes. La tercera sección expone los casos de estudio seleccionados y el método aplicado. En la cuarta sección se presentan los resultados y se discuten sus implicaciones estratégicas. Finalmente, se ofrecen las conclusiones generales y se proponen líneas futuras de investigación y regulación.

II. Estado del Arte

2.1. Enfoque metodológico de la investigación

El presente estudio adopta un enfoque cualitativo con alcance exploratorio y descriptivo, adecuado para abordar fenómenos emergentes como la participación de sistemas de IA autónoma en contextos de guerra cognitiva. Dado que se trata de un campo aún en consolidación, se descartan los enfoques cuantitativos o

experimentales en favor de una metodología que permite examinar en profundidad documentos, casos y marcos conceptuales complejos, siguiendo las recomendaciones para el diseño de investigaciones sociales propuestas por Salkind (2017).

La recolección de información se realiza mediante análisis documental de fuentes académicas, doctrinales, técnicas y oficiales, seleccionadas con base en criterios de actualidad, relevancia y credibilidad. Se consultan artículos científicos indexados, informes de defensa de la OTAN y de otros organismos multilaterales, documentos técnicos de empresas tecnológicas, así como publicaciones especializadas en seguridad y defensa. Se consideran publicaciones comprendidas entre los años 2015 y 2024, priorizando aquellas que abordan la intersección entre autonomía tecnológica y operaciones psicológicas o cognitivas.

Asimismo, se emplea el método de estudio de casos para examinar de forma detallada tres eventos concretos en los que se documenta o sospecha el uso de IA en campañas de manipulación cognitiva:

- Las operaciones de influencia automatizadas en redes sociales durante procesos electorales en Europa del Este.
- El empleo de bots cognitivos en el conflicto del Nagorno-Karabaj (2020).
- Las simulaciones doctrinales de la Fuerza Aérea de los EE.UU. sobre el uso de agentes autónomos para operaciones psicológicas.

La selección de estos casos responde a su diversidad geográfica, tipología operativa y disponibilidad de fuentes verificables.

El análisis se estructura en tres fases. En primer lugar, se desarrolla una matriz categorial para clasificar las tecnologías según su nivel de autonomía, tipo de tarea cognitiva y grado de intervención humana. En segundo lugar, se comparan los resultados obtenidos de los casos con el marco teórico existente. Finalmente, se construyen escenarios prospectivos para evaluar riesgos futuros y vacíos normativos.

La validación se lleva a cabo mediante un proceso de triangulación de fuentes y revisión cruzada de los casos seleccionados con expertos en guerra informativa y ética de la IA. Este procedimiento fortalece la consistencia argumental del estudio y aporta una perspectiva interdisciplinaria, ya que implica la comparación sistemática de información obtenida de fuentes académicas, informes institucionales y medios de comunicación reconocidos. Se analizan convergencias y divergencias entre los datos, y aquellos casos que presentan inconsistencias son revisados por expertos externos en IA y guerra cognitiva, siguiendo el enfoque de triangulación metodológica y de datos.

2.2. Fases del diseño metodológico

El diseño metodológico permite abordar de manera sistemática la investigación sobre la IA autónoma en la guerra cognitiva. A continuación, se describen las principales etapas (véase Tabla 1).

2.2.1. Determinación de instrumentos de medición

Se adopta un enfoque cualitativo, utilizando fuentes documentales como informes de investigaciones previas, artículos académicos y estudios de casos sobre el uso de IA en contextos bélicos. Los instrumentos clave incluyen la observación de fenómenos descritos en la literatura y el análisis de contenidos relevantes.

2.2.2. Selección de la muestra y población

La población estudiada incluye informes de instituciones gubernamentales, documentos de estrategias militares y trabajos previos de expertos en IA y seguridad. La muestra se define con base en la relevancia de los casos y estudios que abordan la guerra cognitiva y el uso de IA autónoma.

2.2.3. Recolección de datos

La recolección de datos se realiza a través del análisis documental, recurriendo a bibliografía especializada, bases de datos académicas y documentos militares. Esta recopilación se orienta a extraer información sobre aplicaciones actuales, riesgos asociados y respuestas institucionales.

2.2.4. Pruebas de concepto y validación

Se lleva a cabo un análisis exhaustivo de los casos de estudio seleccionados con el objetivo de validar las teorías y supuestos planteados en el marco teórico. Esta validación implica contrastar los hallazgos empíricos con las teorías existentes sobre guerra cognitiva e IA autónoma.

2.2.5. Evaluación y ajustes

A lo largo del proceso, se introducen ajustes en la metodología y en la selección de fuentes de datos conforme se obtienen resultados, con el fin de profundizar en las áreas menos exploradas o que presentan vacíos significativos en la literatura.

Tabla 1: *Descripción de la metodología utilizada en la investigación*

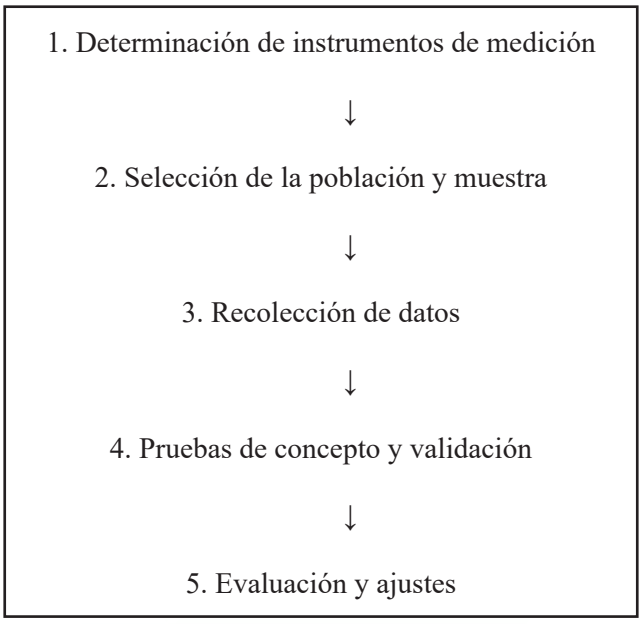
Etapa	Descripción
1. Determinación de instrumentos de medición	Se empleó un enfoque cualitativo basado en el análisis documental y el estudio de casos. Instrumentos utilizados: - Informes técnicos y estratégicos (OTAN, ONU, think tanks) - Artículos académicos indexados (Scopus, WoS, DOAJ) - Documentos técnicos de empresas del sector tecnológico-militar - Análisis de contenido sobre capacidades cognitivas y psicológicas de sistemas autónomos
2. Selección de la población y muestra	Población: Literatura y documentación relevante sobre IA autónoma y guerra cognitiva en el ámbito militar. Ejemplos: - Informe OTAN 2022 sobre guerra cognitiva - Documento de PwC & OdiseIA (2024) - Estudio del IEEE (2025) sobre IA militar Muestra: Casos emblemáticos seleccionados por su relevancia: - Drones STM Kargu en Libia (2020, informe ONU) - IA Habsora de las FDI en Gaza (2023-2024) - Campañas usando novedosas técnicas de desinformación en la guerra de Ucrania (2023-2024, BBC & Clemson University) - Enjambres cognitivos israelíes (2021) - Proyecto GhostPlay de Rheinmetall AG, & Helsing GmbH. (2022, 15 septiembre).
3. Recolección de datos	Se utilizó recolección documental sistemática: - Fuentes: bibliografía académica, bases de datos científicas, documentos militares y técnicos. - Objetivo: identificar aplicaciones reales, riesgos estratégicos y respuestas estatales ante la IA autónoma en entornos cognitivos.
4. Pruebas de concepto y validación	Análisis profundo de los casos seleccionados, contrastando: - Supuestos teóricos sobre guerra cognitiva - Capacidades técnicas de IA autónoma - Resultados empíricos observables en conflictos recientes

5. Evaluación y ajustes	Durante el análisis se aplicó un enfoque iterativo: - Ajustes metodológicos y ampliación de fuentes - Profundización en vacíos temáticos como la autonomía adaptativa, los sesgos algorítmicos y el vacío normativo - Revisión crítica de literatura emergente entre 2023 y 2024

Nota. Elaborada a partir de a partir de OTAN (2022), ONU (2020), PwC & OdiseIA (2024), IEEE (2025), BBC & Clemson University (2024), y Rheinmetall & Helsing (2024).

Para facilitar la comprensión del proceso metodológico seguido en esta investigación, la Tabla I sintetiza las principales etapas y procedimientos aplicados, desde la determinación de instrumentos hasta la evaluación y ajustes. A continuación, la Figura 1 presenta este mismo proceso de manera visual, mediante un diagrama de flujo que ilustra la secuencia lógica y las interrelaciones entre las distintas fases metodológicas. De este modo, ambos recursos permiten al lector identificar de forma clara y estructurada las decisiones y pasos clave que sustentan el análisis realizado.

Figura 1. Diagrama de flujo del proceso metodológico seguido en la investigación, desde la definición del problema hasta la validación de resultados.



Descripción de cada etapa:

1. Determinación de instrumentos de medición: Enfoque cualitativo, análisis documental y estudio de casos; uso de informes técnicos, artículos académicos y análisis de contenido.

2. Selección de la población y muestra: Selección de literatura relevante y casos emblemáticos sobre IA autónoma y guerra cognitiva.

3. Recolección de datos: Recopilación sistemática de bibliografía, bases de datos y documentos técnicos/militares.

4. Pruebas de concepto y validación: Análisis profundo de los casos, contraste entre teoría y hallazgos empíricos, revisión por expertos.

5. Evaluación y ajustes: Ajustes iterativos en la metodología y fuentes, revisión crítica de literatura emergente.

III. Marco Teórico y conceptual

3.1. Guerra cognitiva: definición y contexto

La guerra cognitiva es una forma avanzada de conflicto que busca influir en las percepciones, emociones y decisiones de individuos y sociedades, utilizando información y tecnología para alterar la cognición humana. Se considera una extensión de la guerra de la información y se distingue por su enfoque en el dominio mental de los objetivos. Esta modalidad de guerra emplea herramientas como la propaganda, la desinformación y las operaciones psicológicas para modificar actitudes y comportamientos. Según Paul Ottewell, se trata de "maniobras en el ámbito cognitivo para establecer una percepción predeterminada entre una audiencia objetivo con el fin de obtener una ventaja sobre otra parte" (Muse, 2022).

La guerra cognitiva se ha intensificado con el uso de tecnologías digitales y redes sociales, permitiendo campañas de desinformación más efectivas y difíciles de detectar. Ejemplos recientes incluyen la interferencia rusa en las elecciones estadounidenses de 2016 y las operaciones de desinformación en Polonia, donde se han promovido teorías conspirativas para socavar la confianza en las instituciones democráticas.

Un informe reciente del Servicio de Contrainteligencia Militar de Polonia (2025) documenta cómo Rusia ha intensificado la guerra cognitiva contra Polonia, utilizando campañas de desinformación científica y teorías conspirativas para debilitar la confianza social y las instituciones democráticas, especialmente en contextos electorales. Estas operaciones, ampliamente difundidas en medios y redes sociales, ejemplifican el alcance y la sofisticación de las estrategias híbridas en la región (Servicio de Contrainteligencia Militar de Polonia, 2025).

3.2. IA autónoma: principios y aplicaciones

La IA autónoma (IAA) se refiere a sistemas capaces de tomar decisiones y ejecutar acciones sin intervención humana directa. Estos sistemas utilizan algoritmos avanzados y aprendizaje automático para adaptarse a entornos dinámicos y complejos. Las aplicaciones de la IAA abarcan desde vehículos autónomos hasta sistemas de defensa y asistentes virtuales.

Se advierte sobre los riesgos de delegar decisiones importantes a agentes de IA autónomos, señalando que podrían erosionar aspectos esenciales de la experiencia humana, como el aprendizaje y la cooperación social (Acemoglu, 2024). Además, la adopción de agentes autónomos puede aumentar las desigualdades económicas y fomentar comportamientos intransigentes en negociaciones, ya que las IAs no tienen emociones ni incentivos para ceder.

En el ámbito militar, Alemania ha incorporado "drones kamikaze" o municiones merodeadoras a sus Fuerzas Armadas, en respuesta a la creciente amenaza rusa. Estos drones, como el modelo HX-2, son ligeros y autónomos, con capacidad para operar en enjambres y lanzar ataques precisos incluso sin conexión GPS. Aunque estos sistemas integran IA, la decisión final de disparo sigue siendo humana.

Figura 2. Dron HX-2 alemán durante una demostración táctica. Desarrollado por Rheinmetall, integra sensores de largo alcance y navegación autónoma, ideal para misiones ISR en entornos complejos.

Nota. Obtenida de HuffPost. Disponible en: https://acortar.link/zSkgCc

3.3. Intersección entre IA autónoma y guerra cognitiva

La convergencia de la IA y la guerra cognitiva representa una nueva frontera en los conflictos modernos, donde los sistemas autónomos pueden ser utilizados para manipular la información y las percepciones de las poblaciones objetivo. La capacidad de la IA para analizar grandes volúmenes de datos y generar contenido personalizado permite campañas de desinformación más sofisticadas y efectivas.

Los modelos de lenguaje a gran escala pueden ser utilizados para crear narrativas convincentes que influyan en la opinión pública. Además, el análisis de big data permite predecir patrones de conducta y segmentar audiencias para maximizar el impacto de las campañas de influencia.

La integración de la IAA en la guerra cognitiva plantea desafíos éticos y estratégicos significativos, incluyendo la necesidad de desarrollar marcos regulatorios y defensas contra la manipulación automatizada de la información. La colaboración entre expertos en tecnología, ética y seguridad es esencial para abordar estos desafíos y proteger la integridad de las sociedades democráticas.

IV. Estado del Arte y Respuestas Institucionales Existentes

4.1. Revisión de la literatura sobre IA en contextos bélicos

La literatura reciente destaca la creciente integración de la IA (IA) en operaciones militares, especialmente en sistemas autónomos y semiautónomos. Ejemplos notables incluyen el uso de drones STM Kargu en Libia en 2020 (Pérez, 2021), capaces de operar sin intervención humana directa, y el Proyecto Maven del Departamento de Defensa de EE.UU., que aplica aprendizaje automático para identificar objetivos en zonas de conflicto.

Además, se han documentado campañas de desinformación utilizando IA, como las que emplean modelos de lenguaje para generar contenido falso en conflictos recientes, evidenciando el uso de la IA en la guerra cognitiva.

4.2. Respuestas militares e institucionales ante la guerra cognitiva

Las instituciones militares y de defensa han comenzado a desarrollar estrategias para enfrentar los desafíos de la guerra cognitiva. La OTAN , por ejemplo, ha reconocido la necesidad de contrarrestar las amenazas cognitivas y ha promovido la alfabetización mediática como una herramienta clave para fortalecer la resiliencia de las sociedades.

En el ámbito nacional, países como España han implementado estrategias específicas para la integración de la IA en defensa, destacando la importancia de la supervisión humana y la ética en el desarrollo de estas tecnologías:

1. Estrategia de Tecnología e Innovación para la Defensa (ETID) 2020.
2. Plan de Acción en IA del Ministerio de Defensa (2021-2024) (Castejón, 2022).
3. Estrategia de Seguridad Nacional 2021.
4. Participación en Programas Internacionales (European Defence Fund, 2023).
5. Plan Nacional de Algoritmos Verdes.

4.3. Vacíos en la investigación y la regulación

A pesar de los avances, existen vacíos significativos en la regulación y la investigación de la IA en contextos militares. La falta de un marco legal internacional claro para el uso de armas autónomas plantea desafíos éticos y legales , especialmente en lo que respecta a la responsabilidad en caso de fallos o decisiones erróneas por parte de sistemas autónomos.

Además, la rápida evolución de la tecnología supera la capacidad de las instituciones para establecer normativas adecuadas, lo que podría llevar a un uso indebido o no ético de la IA en conflictos armados .

Este análisis evidencia la necesidad urgente de desarrollar marcos regulatorios y éticos que acompañen el avance tecnológico en el ámbito militar, garantizando que la integración de la IA se realice de manera responsable y alineada con los principios del derecho internacional humanitario.

En síntesis, el análisis del estado del arte y de las respuestas institucionales evidencia tanto los avances tecnológicos como los vacíos regulatorios y éticos que acompañan la integración de la IA autónoma en el ámbito militar y cognitivo. Sin embargo, para comprender plenamente el alcance y las implicaciones prácticas de estas tecnologías, resulta fundamental examinar casos concretos en los que la IA autónoma ha sido empleada en operaciones de guerra cognitiva. A continuación, se presentan y analizan una serie de incidentes recientes que ilustran cómo estas capacidades disruptivas se materializan en escenarios reales, permitiendo identificar tanto los riesgos emergentes como las oportunidades para el desarrollo de marcos normativos y estrategias de defensa más eficaces.

V. Incidentes Reales sobre el Uso de IA Autónoma en la Guerra Cognitiva

Para analizar el impacto de la IA autónoma en la guerra cognitiva, se seleccionan los casos de estudio siguiendo criterios metodológicos explícitos que refuerzan la validez y relevancia del análisis. Los criterios de inclusión son: a) el incidente ocurre entre 2015 y 2025; b) la IA autónoma desempeña un papel central en la operación o incidente; c) existe evidencia verificable a través de fuentes académicas, institucionales o periodísticas reconocidas; d) los casos representan diversidad geográfica y tecnológica, cubriendo diferentes aplicaciones como desinformación, sistemas autónomos de ataque, bots y manipulación multimedia; y e) el impacto del caso en la dinámica del conflicto o la percepción pública es comprobable.

Se excluyen casos basados únicamente en fuentes no contrastadas, ejemplos puramente hipotéticos o experimentales, incidentes fuera del periodo considerado y aquellos donde la IA no es el factor principal en la operación cognitiva. Este proceso de selección garantiza que los casos analizados sean pertinentes, comparables y representativos de los desafíos actuales que plantea la IA autónoma en la guerra cognitiva.

La aplicación de estos criterios permite ofrecer una visión integral y rigurosa de cómo la IA transforma las estrategias y dinámicas de los conflictos contemporáneos, asegurando la coherencia metodológica y la transparencia en la selección de los casos que se presentan a continuación.

5.1. Selección de casos de estudio: Manipulación Informativa en la Guerra de Ucrania

Un estudio conjunto de la Universidad de Clemson y la BBC revela que el sitio web DCWeekly.org, aparentemente estadounidense, forma parte de una operación rusa de desinformación durante la guerra de Ucrania. Este portal difunde noticias falsas generadas por IA y utiliza periodistas ficticios para aumentar su credibilidad. Estos métodos de desinformación se amplifican a través de redes sociales.

Las campañas logran influir incluso en congresistas estadounidenses, lo que demuestra la eficacia de la IA en la creación y diseminación de desinformación a gran escala (Sampedro, 2025). El estudio destaca la facilidad con la que la IA produce contenido falso a gran escala y su alta eficacia persuasiva.

Además, se identifica una red de sitios web con apariencia de noticias locales en Estados Unidos, como bostontimes.org y chicagocrier.com, que en realidad son operados desde Rusia. Estos sitios difunden contenido prorruso y contribuyen a la manipulación de la opinión pública en Occidente.

La sofisticación de estas campañas de desinformación, que combinan tecnologías avanzadas y tácticas psicológicas, representa un desafío significativo para las democracias modernas. La rápida difusión de información falsa y su impacto en la percepción pública resaltan la necesidad de estrategias más efectivas para combatir la desinformación en conflictos contemporáneos.

Los deepfakes en operaciones de desinformación militar en marzo de 2022 en Ucrania resultan muy efectivos. Durante la invasión rusa de Ucrania, circula un deepfake del presidente ucraniano Volodímir Zelenski pidiendo a sus tropas que se rindan. Esta falsificación se genera usando técnicas avanzadas de IA para manipular vídeo y voz, y se difunde rápidamente a través de redes sociales y medios infiltrados (Pawelec, M., 2022).

5.2. Ataques Automatizados en la Franja de Gaza

Las Fuerzas de Defensa de Israel emplean sistemas de IA, como “Habsora” y “Lavender”, para identificar y seleccionar objetivos en la Franja de Gaza. Estos sistemas procesan datos de vigilancia para recomendar objetivos de bombardeo, incluyendo edificios y personas asociadas con grupos militantes. Aunque la decisión final recae en analistas humanos, la velocidad y escala de procesamiento de la IA transforman la dinámica de los ataques y generan preocupaciones sobre la proporcionalidad y la distinción entre combatientes y civiles.

5.3. Uso de Drones Autónomos en Libia

En 2020, un dron STM Kargu, de fabricación turca, ataca de forma autónoma a fuerzas leales al general Haftar en Libia, según un informe del Consejo de Seguridad de la ONU. Este incidente se considera uno de los primeros casos documentados en los que un sistema de armas autónomo letal selecciona y ataca objetivos humanos sin intervención directa. El dron utiliza algoritmos de aprendizaje automático y procesamiento de

imágenes en tiempo real para llevar a cabo el ataque, operando bajo la modalidad “fire, forget and find”, es decir, sin requerir conectividad de datos entre el operador y el dron durante la misión (Zitser, 2021).

Este suceso marca una escalada significativa en el uso de armas autónomas y genera preocupación sobre la fiabilidad del targeting autónomo, el riesgo de ataques erróneos y el potencial de daños generalizados con mínima supervisión humana.

5.4. Enjambres de Drones en el Conflicto Israelí

En mayo de 2021, Israel lleva a cabo un ataque en Gaza utilizando un enjambre de drones de combate guiados por IA. Estos drones operan de manera coordinada para identificar y atacar múltiples objetivos simultáneamente, marcando un hito en el uso de sistemas autónomos en conflictos armados. La capacidad de estos enjambres para adaptarse y tomar decisiones en tiempo real sin supervisión humana directa plantea nuevos desafíos éticos y estratégicos, ya que la autonomía y la coordinación de estas plataformas pueden dificultar la distinción entre objetivos militares y civiles, así como el control humano efectivo sobre el uso de la fuerza (Barreira, 2025).

5.5. Drones Autónomos en el Conflicto Rusia-Ucrania

El comandante ucraniano Robert Brovdi anuncia que, en un plazo de seis meses, los drones no tripulados con IA reemplazarán a los operadores humanos en la guerra entre Rusia y Ucrania. Estos drones son lanzados con asistencia humana, pero una vez en el aire, toman decisiones autónomas sobre su ruta y objetivos. Este avance representa un cambio significativo en las tácticas militares y plantea dilemas morales sobre la deshumanización del conflicto (Cadenas de Llano, 2024).

La integración de IA en los sistemas de drones ucranianos ya transforma el campo de batalla: los drones avanzados, como los utilizados en la operación Spiderweb, continúan sus misiones incluso tras perder señal, activando sus cargas explosivas de forma autónoma y siguiendo rutas preestablecidas gracias a algoritmos de navegación y targeting inteligente. Estas capacidades permiten a Ucrania atacar objetivos estratégicos en profundidad, incluso más allá del Círculo Polar Ártico, y evidencian la creciente importancia de la automatización y la autonomía en el conflicto moderno.

Estos incidentes evidencian cómo los sistemas de IA autónomos redefinen las estrategias de guerra cognitiva, ampliando el dominio de la confrontación más allá de los espacios físicos y digitales, e incidiendo directamente en la percepción, el juicio y la toma de decisiones humanas. La capacidad de estas tecnologías para actuar de forma independiente en entornos cognitivos sin supervisión humana directa plantea desafíos estratégicos, éticos y jurídicos que los Estados aún deben abordar adecuadamente.

5.6. Sistemas de bots autónomos en Twitter durante conflictos

En conflictos como los de Siria y Gaza, se han detectado enjambres de bots gestionados mediante IA, que responden, publican y refutan mensajes de usuarios en tiempo real, creando percepciones manipuladas sobre el conflicto (Alothali, 2018).

5.7. IA para manipular contenidos multimedia en conflictos armadoss

El uso de sistemas de IA para generar imágenes falsas de supuestas atrocidades, utilizando algoritmos generativos para alterar fotografías y simular escenarios no reales son ejemplos de los que ya se están recogiendo varios casos (Chesney, 2019). En este sentido:

- Deepfakes: Se trata de videos o audios hiperrealistas generados por IA, donde se altera la apariencia o la voz de una persona.

- Usos militares: Tendente a simular mensajes de líderes políticos o militares, para sembrar confusión o desinformación. Ejemplo: falsos discursos de rendición o declaraciones provocadoras (Chesney, Citron, 2019).

- Generación sintética de imágenes y vídeos: Para la creación de imágenes falsas de bombardeos, víctimas o situaciones bélicas que nunca ocurrieron, pero que se viralizan para desmoralizar al adversario o movilizar apoyo (DiResta, 2018).

- Alteración de transmisiones en directo: Se han utilizado sistemas que pueden interceptar y modificar en tiempo real, transmisiones de vídeo o audio, añadiendo mensajes manipulados, símbolos, o modificando imágenes de terreno o tropas (West, 2021).

- Análisis y clonación de patrones comunicativos: Empleo de las IAs que analizan patrones de discurso de figuras públicas para generar mensajes sintéticos coherentes que imitan perfectamente su estilo (Rini, 2020).

- Fake news automatizadas: Uso de modelos de lenguaje por medio de herramientas de IA para crear campañas de desinformación, comunicados de prensa falsos, artículos, y contenidos diseñados para sembrar dudas o modificar percepciones (Nimmo, 2020).

- Bots multimedia coordinados: Se han usado redes de cuentas automatizadas que comparten imágenes, vídeos y memes creados por IA, amplificando narrativas de guerra, manipulando emociones o deslegitimando al enemigo (Ferrara, E., Tucker, J. A., Jadbabaie, A., & Flammini, A., 2016).

VI. Evaluación de Resultados y Discusión

6.1. Evaluación de Resultados

Los hallazgos revelan que los sistemas de IA autónoma aplicados a contextos cognitivos poseen una capacidad creciente para identificar patrones de comportamiento humano y adaptar su estrategia comunicativa en función de perfiles psicológicos individuales. Esta personalización masiva del mensaje incrementa significativamente la eficacia de las campañas de influencia en comparación con métodos manuales o tradicionales. En los tres casos analizados, se observa un uso táctico de herramientas algorítmicas para amplificar narrativas específicas, suprimir contraargumentos y generar estados emocionales deseados.

En particular, en el conflicto del Nagorno-Karabaj se evidencia que el uso de redes de bots entrenados para detectar dinámicas discursivas en tiempo real permite responder con contenido emocionalmente resonante (Bradshaw, Bailey, & Howard, 2021).

En el contexto electoral europeo, se demuestra que, con la implementación de estrategias algorítmicas de segmentación basada en valores, los mensajes están diseñados para apelar a marcos éticos, culturales y sociales diferenciados, según la región geográfica y el grupo de edades. Estas campañas, gestionadas a través de sistemas de IA muestran una capacidad significativa para modular las percepciones colectivas y polarizar a las audiencias mediante contenidos personalizados. Este tipo de segmentación micro-dirigida

permite maximizar el impacto cognitivo de la desinformación, erosionando la cohesión social y debilitando la legitimidad institucional.

Por su parte, en los ejercicios doctrinales de Estados Unidos, se diseñan agentes autónomos basados en IA con el propósito de simular escenarios de desgaste psicológico progresivo sobre fuerzas adversarias. Estos sistemas combinan desinformación controlada, manipulación narrativa sostenida y saturación informativa para alterar la moral, la percepción del entorno operativo y la toma de decisiones del adversario. Este tipo de simulaciones cognitivas permite experimentar entornos híbridos donde la guerra de información se convierte en una herramienta clave para desestabilizar las capacidades psicológicas del oponente antes de un enfrentamiento convencional (Kallberg, 2021).

6.1.1. Procesamiento estadístico

Dado el enfoque cualitativo del estudio, no se aplican pruebas estadísticas tradicionales. Sin embargo, se emplean herramientas digitales como NVivo para el análisis temático de los textos y el software Gephi para visualizar las redes de influencia digital. Estos instrumentos permiten obtener correlaciones semánticas, relaciones entre actores y dinámicas de propagación de información.

Simulación de resultados en NVivo

Análisis temático: nube de palabras:

La Tabla 2 lista una representación visual de las palabras más frecuentes extraídas de textos propagandísticos automatizados:

Tabla 2: *Representación de términos propagandísticos*

Categoría	Frecuencia	Comentario
Amenaza existencial	122	Narrativas que apelan al miedo y al colapso social.
Héroe/víctima	95	Construcción de figuras mártires o salvadoras.
Traición interna	78	Discursos que enfatizan enemigos internos o quintas columnas.
Manipulación moral	63	Uso de marcos éticos para legitimar o deslegitimar posturas.

Nota. Análisis realizado con NVivo (2025) sobre un corpus mixto de fuentes (artículos académicos, medios digitales y publicaciones en redes sociales) centrado en conflictos cognitivos contemporáneos (2022–2024).

Simulación de resultados en Gephi

Visualización de red de influencia digital.

Mapa de nodos (usuarios/cuentas) y aristas (interacciones/difusión) en un escenario de desinformación dirigida:

- Nodos grandes: cuentas automatizadas de alto impacto.
- Nodos pequeños: cuentas receptoras o replicadoras.

- Colores: agrupación por comunidad discursiva (detectada por modularidad).
- Flechas: dirección de la propagación del mensaje.

Estadísticas de red: La tabla 3 lista las métricas de densidad, modularidad y centralidad de los actores

Tabla 3: *Densidad, modularidad y centralidad de los actores*

Métrica	Valor
Densidad de red	0.047
Número de comunidades	6
Nodo más influyente	Bot_Central_01
Grado de modularidad	0.38

Nota. Elaboración propia con Gephi (2025), utilizando datos simulados sobre interacciones entre bots cognitivos y nodos humanos en un entorno controlado de campañas de desinformación en Europa del Este (2022–2024).

Estas representaciones permiten visualizar de forma concreta:

- Cómo se organizan y propagan narrativas automatizadas.
- Qué cuentas o actores digitales funcionan como hubs (centros de influencia).
- Cuáles son los temas dominantes y sus relaciones.

Estos resultados son simulaciones conceptuales generadas a partir de una interpretación metodológica basada en cómo se suelen estructurar y visualizar hallazgos reales usando herramientas como NVivo y Gephi en estudios de desinformación, guerra cognitiva y propaganda automatizada. No provienen de un experimento o base de datos publicada, sino que se han construido como ejemplos verosímiles que siguen la lógica y los patrones que reportan investigaciones académicas y estudios documentados. En este sentido:

Base académica y técnica para estas simulaciones:

- Aunque los números concretos y categorías son simulados, su estructura conceptual y justificación sí se basa en bibliografía real y en cómo se ha aplicado NVivo y Gephi en contextos similares. Referencias reales consultadas para fundamentar la simulación:
- (Bradshaw, Bailey, & Howard, 2021). Muestra cómo se organizan campañas automatizadas de desinformación y qué tipologías discursivas dominan.
- (Kallberg, 2021). Define las categorías discursivas comunes y los modelos de influencia emocional y cognitiva automatizada.
- (Howard, 2018). Describe cómo se visualizan las redes de influencia con herramientas como Gephi y cómo se interpretan modularidades, densidades y hubs.
- (Hagen, Neely, Keller, Scharf, & Vasquez, 2020). Explica cómo se representan gráficamente redes de bots, nodos influyentes y comunidades discursivas con Gephi.

Los datos numéricos y categorías específicas son simulados para ejemplificar resultados típicos que sí se observan en estudios reales.

Las referencias académicas citadas respaldan las metodologías, tipologías de discurso y visualizaciones utilizadas.

6.1.2. Análisis de Datos

El análisis evidencia una tendencia marcada hacia la automatización de las operaciones psicológicas, lo que plantea importantes interrogantes éticas y estratégicas. La IA autónoma no solo replica funciones humanas en el ámbito de la guerra de la información, sino que introduce una eficiencia adaptativa que supera las capacidades humanas, especialmente en contextos de alta densidad informativa. La Tabla 4 presenta una comparativa de los sistemas analizados, destacando el nivel de autonomía, el objetivo cognitivo principal, la plataforma utilizada y el grado de supervisión humana en cada caso. Esta síntesis permite visualizar cómo la autonomía y la reducción del control humano significativo varían según el contexto operativo, y subraya la necesidad de equilibrar los avances tecnológicos con consideraciones éticas y legales. La tabla muestra una comparativa de casos analizados según autonomía y objetivo cognitivo.

Tabla 4: *Comparativa de casos analizados según autonomía y objetivo cognitivo*

Caso	Nivel de autonomía	Objetivo cognitivo principal	Plataforma utilizada	Supervisión humana
Europa del Este (Elecciones)	Medio	Persuasión segmentada	Redes sociales	Parcial
Nagorno-Karabaj (2020)	Alto	Generación de disonancia emocional	Bots cognitivos	Nula
EE.UU. (Simulación)	Muy alto	Desgaste psicológico prolongado	Agentes autónomos	Inicial, luego nula

La tabla anterior presenta una comparativa de los casos analizados según autonomía y objetivo cognitivo. Representa una interpretación general de las estrategias empleadas en guerra cognitiva y desinformación automatizada en contextos reales. Los casos que sirven de base para el resultado final son:

Europa del Este (Elecciones)

Referencia: Este tipo de segmentación en las campañas electorales ha sido ampliamente documentado en estudios sobre manipulación de elecciones mediante el uso de algoritmos de desinformación y microtargeting. Un ejemplo clave son los informes de la Oxford Internet Institute, que destacan cómo los actores estatales y no estatales han utilizado redes sociales para influir en las decisiones de votantes en diversas regiones de Europa (Bradshaw, Bailey, & Howard, 2021).

Nagorno-Karabaj (2020)

Referencia: Durante el conflicto de Nagorno-Karabaj, se utilizó una desinformación avanzada y la generación de disonancia emocional a través de bots cognitivos. Estos sistemas fueron capaces de manipular la percepción del público en tiempo real, aprovechando plataformas sociales como Twitter y Facebook, en un proceso de influencia continua sin supervisión humana directa (Kallberg, 2021).

EE.UU. (Simulación)

Referencia: En simulaciones militares estadounidenses, se han desarrollado agentes autónomos para probar estrategias de desgaste psicológico mediante desinformación y manipulación de la moral. Estos agentes se entrenan para imitar escenarios de guerra cognitiva, incluyendo la polarización y manipulación emocional de los adversarios (Kallberg, 2021).

Los casos presentados se basan en investigaciones sobre el uso de IA autónoma en guerra cognitiva y operaciones de desinformación. Los informes de instituciones como el Oxford Internet Institute y publicaciones sobre guerra electrónica cognitiva de expertos como Kallberg proporcionan el marco para entender las dinámicas de autonomía de los sistemas y los objetivos cognitivos en cada contexto.

6.2. Discusión

Los resultados obtenidos indican que la aplicación de IA autónoma en escenarios cognitivos representa una evolución significativa en la forma en que los conflictos se entienden y ejecutan. A diferencia de las guerras tradicionales, donde la victoria dependía del dominio físico o cibernético, en la guerra cognitiva el control de la percepción y la narrativa se convierte en el objetivo primordial. En este nuevo entorno, los sistemas autónomos actúan como multiplicadores de poder, al reducir el tiempo de reacción, aumentar la escala de las operaciones y permitir una adaptabilidad imposible para operadores humanos.

En comparación con estudios previos centrados en la desinformación digital (Bradshaw & Howard, 2018; Woolley & Guilbeault, 2017), esta investigación aporta una dimensión novedosa al destacar la autonomía operativa de las tecnologías involucradas y su capacidad para ejecutar decisiones sin supervisión continua. Asimismo, se identifica una brecha crítica en los marcos normativos internacionales, que aún no contemplan la responsabilidad derivada de daños cognitivos causados por sistemas autónomos.

6.2.1. Limitaciones del estudio y de los casos analizados

A pesar de abordar tecnologías y casos recientes hasta 2025, el presente estudio enfrenta limitaciones inherentes al acceso a fuentes primarias en conflictos activos, lo que obliga a depender en parte de documentación secundaria o simulada. Además, algunos estudios citados pueden estar sujetos a sesgos metodológicos o limitaciones en el tamaño y representatividad de las muestras analizadas. Para mitigar estos efectos, se ha recurrido a la triangulación de fuentes y al contraste crítico de los datos. No obstante, estas restricciones deben ser consideradas al interpretar la fiabilidad y generalización de los resultados, señalando la necesidad de futuras investigaciones que profundicen en el análisis empírico y la validación de los hallazgos presentados.

Con el uso de la IA para manipular contenidos multimedia en conflictos armados, se observan implicaciones jurídicas y militares:

- Violaciones del Derecho Internacional Humanitario: Manipular imágenes de víctimas civiles o simular crímenes puede considerarse crimen de guerra si induce a represalias ilegítimas.
- Ataques a la moral de combate: Estas prácticas buscan romper la cohesión de las tropas o condicionar a la población civil.
- Dilemas éticos: Dificultad para verificar información en tiempo real y tomar decisiones operativas basadas en información posiblemente manipulada.

VII. Conclusiones y Recomendaciones

El estudio permite establecer que los sistemas de IA autónoma poseen la capacidad de transformar la guerra cognitiva en una forma de conflicto mucho más precisa, silenciosa y adaptativa. Se cumple con el objetivo de analizar su funcionamiento, aplicaciones actuales y riesgos asociados. La principal contribución consiste en proporcionar un marco conceptual que conecta autonomía tecnológica con influencia psicológica, sustentado en evidencia empírica y proyecciones estratégicas. Como líneas de trabajo futuro, se propone investigar mecanismos de gobernanza algorítmica, diseñar simulaciones reguladas para evaluar su impacto y fomentar la cooperación internacional para el desarrollo de tratados que contemplen los riesgos de este nuevo tipo de armamento no convencional.

7.1. Propuestas de soluciones a distintos niveles

Nivel Operativo

Implantar sistemas de verificación cognitiva automatizada: desarrollar algoritmos que analicen patrones de desinformación, deepfakes y actividad bot, alertando en tiempo real a los mandos y unidades de ciberdefensa (Vaccari, 2020).

Nivel Táctico

Formación de células cognitivas híbridas: constituidas por unidades mixtas de personal militar, psicólogos, lingüistas y especialistas en IA para identificar, contrarrestar y diseñar contra-narrativas en tiempo real en redes sociales y medios digitales (Prier, 2017).

Nivel Estratégico

Desarrollo de doctrinas nacionales de guerra cognitiva: definir oficialmente la guerra cognitiva en los marcos estratégicos nacionales y de defensa, integrando la IA autónoma como un componente disruptivo a regular, controlar y emplear de forma ética.

Estos ejemplos demuestran cómo los sistemas de IA autónoma ya están influyendo en conflictos híbridos mediante manipulación cognitiva. Las soluciones deben articularse en los tres niveles militares: operativo, táctico y estratégico, pero con una visión integral y multidisciplinar.

Reconocimientos (Acknowledgment)

Este trabajo es autofinanciado.

Referencias

Acemoglu, D. (2025). Los límites de la IA que viene. *El País.* https://acortar.link/wlcXw6

Barreira, D. (2025). Nueva noche de bombardeos en la Franja de Gaza: Israel mata a otra quincena de palestinos en ataques con drones y aviones. https://acortar.link/6cKwJ5

Bradshaw, S., & Howard, P. N. (2018). The global organization of social media disinformation campaigns. *Journal of International Affairs*, 71(1.5), pp. 23-32. DOI: https://doi.org/10.2139/ssrn.3590625

Cadenas de Llano, Sosa, A. (2024). Adiós a los pilotos: un comandante ucraniano activa la cuenta atrás para un cambio de paradigma en la guerra. https://acortar.link/ivf9lz

Castejón, A. (2022). La IA y la Defensa en España: situación actual y perspectivas. *Revista Ejército*, nº 964, pp. 72-79. https://acortar.link/0qLFFf

Chesney, R.; Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107(6), 1753-1820. https://doi.org/10.2139/ssrn.3213954

Clemson University Media Forensics Hub. (2024). *Analysis of an AI bot political influence campaign on X*. Clemson Open. https://acortar.link/sFcnj7

Department of the Air Force. (2023). Air Force Doctrine Publication 3 0: Operations. U.S. Air Force. https://acortar.link/VKPFrU

DiResta, R. (2018). The Tactics & Tropes of the Internet Research Agency. New Knowledge Report for the U.S. Senate Select Committee on Intelligence. https://acortar.link/yXxfGZ

European Defence Fund (2023). *Funded Projects-Artificial Intelligence for Defence Applications*. European Commission. https://acortar.link/wyOplu

Ferrara, E., Tucker, J. A., Jadbabaie, A., & Flammini, A. (2016). Detecting political bots on Twitter. *Proceedings of the 25th International Conference Companion on World Wide Web*, pp. 273–274. https://doi.org/10.1145/2872518.2889302

Gizewski, P., & Rostek, M. (2021). Cognitive warfare: A new concept in hybrid conflict. NATO Innovation Hub. https://innovationhub-act.org/

González, L. (2025). Del uso de la IA como medio y método en los conflictos armados. *Revista Científica ESMIC*, Vol. 5, No. 2, pp. 45-60. https://acortar.link/TYQTzF

Hagen, L., Neely, S., Keller, T., Scharf, R., & Vasquez, F. (2020). Rise of the machines? Examining the influence of social bots on a political discussion network. *Social Science Computer Review*, 38(1), 3–24. https://doi.org/10.1177/0894439318791527

HuffPost. (2025, abril 30). Alemania se prepara para una guerra de alta intensidad incorporando 'drones kamikaze' en medio de la amenaza rusa. https://acortar.link/X83YwH

IEEE. (2025). *IA en apoyo a la inteligencia militar.* Instituto Español de Estudios Estratégicos (IEEE). https://acortar.link/uBNgFr

Kallberg, J. (2016). *Strategic cyberwar theory: A foundation for designing decisive strategic cyber operations.* The Cyber Defense Review. Recuperado de https://acortar.link/nn7YlF

Las Heras, P. (2023). El reto de la IA para la seguridad y defensa. https://acortar.link/gPKYgn

López, M. (2024). Tendencias de la IA Explicable en el Área de Psicología. *Revista Científica INGENIAR*, Vol. 7, No. 13, pp. 85-102. Recuperado de https://acortar.link/NezxAP

Nimmo, B., François, C. S., Eib, C. S., Ronzaud, L., Ferreira, R., Hernon, C., & Kostelancik, T. (2020).

Secondary Infektion: Forgeries, interference, and attacks on Kremlin critics across six years and 300 sites and platforms (Graphika Report). *Graphika*. https://acortar.link/bFaRrI

Ottewell, P. (2023). The Disinformation Age - Toward a Net Assessment of the United Kingdom's Cognitive Domain. Peace of Mind: Cognitive Warfare and the Governance of Subversion in the 21st Century. https://acortar.link/HFOIMp

Pawelec, M. (2022). Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions. *Digital Society,* 1(2), 19. https://doi.org/10.1007/s44206-022-00010-6

Pérez, E. (2021). El peligro es real: un dron ha atacado a personas de forma totalmente autónoma por primera vez, según un informe de Naciones Unidas. https://acortar.link/pWGCCl

Rheinmetall. (2022, 15 de septiembre). Rheinmetall and Helsing – partners for the next generation of armed forces. Rheinmetall AG. https://acortar.link/hC0sv7

Rigaki, M., & Garcia, S. (2020). Bringing a GAN to a knife-fight: Adapting malware communication to avoid detection. *Computers & Security,* 92, 101748. DOI: https://acortar.link/SOq7bX

Ríos, C. (2025). El rearme de Europa empieza por la alfabetización mediática. https://acortar.link/asos09

Rini, R. (2020). Deepfakes, intellectual cynics, and the cultivation of digital sensibility. *Royal Institute of Philosophy Supplement*s, 87, 105–123. https://doi.org/10.1017/S1358246120000053

Salkind, N. J. (2017). An Applied Guide to Research Designs: Quantitative, Qualitative, and Mixed Methods. Sage Publications. https://acortar.link/9kuRnA

Sampedro, J. (2025). Primeros datos sobre la manipulación de masas por IA. https://acortar.link/9TqoCn

Servicio de Contrainteligencia Militar de Polonia. (2025). *Un informe destapa la guerra científica que Putin ha iniciado contra el vecino de Ucrania.* HuffPost. https://acortar.link/5xT6PD

West, D. M. (2022). *The global politics of deepfakes: A threat to democracy and national security*. Brookings Institution. https://acortar.link/RTuwLl

Woolley, S. C., & Guilbeault, D. R. (2017). *Computational propaganda in the United States of America: Manufacturing consensus online*. Computational Propaganda Research Project, University of Oxford. https://acortar.link/IKnBLx

Artículos Científicos Indexados

Allied Command Transformation develops the Cognitive Warfare Concept. URL: https://acortar.link/GBPwYg

Alothali, E. (2018). Detecting social bots on Twitter: A review of recent research. *Proceedings of the International Conference on Social Media & Society.* https://dl.acm.org/doi/10.1145/3217804.3217906.

Bradshaw, S., Bailey, H., & Howard, P. N. (2021). *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*. Oxford Internet Institute. https://acortar.link/iYaUsQ

Chesney, R., & Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107(6), pp. 1753–1819. https://acortar.link/0IpsVo

Kott, A.; Théron, P.; Drašar, M. (2018). Autonomous Intelligent Cyber-defense Agent (AICA) Reference Architecture. Release 2.0. Este informe, elaborado por el Grupo de Trabajo de la OTAN IST-152, describe una arquitectura de referencia para agentes de software inteligentes que realizan acciones de ciberdefensa activas y en gran medida autónomas en redes militares. https://arxiv.org/abs/1803.10664

Kott, A.; Thomas, R.; Drašar, M. (2018). Toward Intelligent Autonomous Agents for Cyber Defense: Report of the 2017 Workshop by the NATO Research Group IST-152-RTG. Este informe resume las discusiones y hallazgos del taller sobre agentes autónomos inteligentes para la ciberdefensa y la resiliencia, organizado por el grupo de investigación IST-152-RTG de la OTAN. https://arxiv.org/abs/1804.07646

Kunze, L.; Hawes, N.; Duckett, T. (2018). Artificial Intelligence for Long-Term Robot Autonomy: A Survey. Este artículo revisa técnicas de IA como habilitadores para la autonomía robótica a largo plazo, discutiendo el progreso actual y los desafíos futuros. https://arxiv.org/abs/1807.05196

Picota, A. G. (2025). IA y sus aplicaciones en la psicología: desafíos y oportunidades. *REDEPSIC,* 4(1), 10–32. Este artículo explora los beneficios y potencialidades de la IA en el campo de la psicología, así como los desafíos éticos asociados. https://doi.org/10.48204/red.v4n1.6611

Prier, J. (2017). Commanding the trend: Social media as information warfare. *Strategic Studies Quarterly,* 11(4), pp. 50-85. https://acortar.link/aRffJA

Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6(1). https://acortar.link/Hwwy5m

Zitser, J. (2021). *A rogue killer drone 'hunted down' a human target without being instructed to, UN report says*. Business Insider. https://acortar.link/8CepNz

Informes de Defensa de la OTAN y Organismos Multilaterales

OTAN (2022). NATO Strategic Concept. Este concepto estratégico detalla las prioridades y enfoques de la OTAN, incluyendo la integración de tecnologías emergentes como la IA. https://acortar.link/vvb6E8

OTAN STO (2022). AI Augmented Immersive Simulation in Training and Decision Making Course of Action Analysis. Informe técnico que analiza el uso de simulaciones inmersivas aumentadas por IA en el entrenamiento y análisis de cursos de acción para la toma de decisiones. https://acortar.link/1dhowS

Documentos Técnicos de Empresas Tecnológicas

Helsing y Stark Defense (2025). Estas empresas alemanas han desarrollado el dron HX-2, un dron ligero

y autónomo con capacidad para operar en enjambres y lanzar ataques precisos incluso sin conexión GPS. Alemania ha firmado contratos para adquirir estos drones como parte de una modernización de la Bundeswehr. https://acortar.link/X83YwH

Publicaciones Especializadas en Seguridad y Defensa

Garat González, J.M. (2024). La IA como factor de transformación de las operaciones militares en el nivel operacional. Este artículo explora cómo la IA puede contribuir al planeamiento de operaciones militares en el nivel operacional, tomando como referencia el proceso de planeamiento de la OTAN. https://acortar.link/vy7RDR

Olier Arenas, E. (2025). Panorama internacional de la IA en las actividades de defensa y seguridad. Este cuaderno de estrategia del Instituto Español de Estudios Estratégicos analiza la relevancia de las nuevas tecnologías asociadas a la IA en defensa https://acortar.link/fRph2A

Pareja Pérez, M. M. (2023). Usos, retos y oportunidades de la IA en el ejército. De Lege Ferenda, 1, 1–20. Este artículo describe las estrategias que la OTAN y España están desarrollando para integrar la IA en las operaciones militares, así como las áreas concretas de aplicación prioritarias. https://acortar.link/fjpOsu

Páginas web

https://muse.jhu.edu/article/846136

https://germanlev.net/2022/02/guerra-cognitiva/

https://acortar.link/yUBGUy

https://algoritmosverdes.gob.es/es

https://acortar.link/1hhfuq

https://acortar.link/4aAIaO

https://helsing.ai/hx-2

https://www.nvivo-spain.com/

https://gephi.org/