Foro Empresas Innovadoras

amper

DOCUMENTO COMPLEMENTARIO AL INFORME "SITUACIÓN Y TENDENCIAS EN EL USO DE LA IA EN EL SECTOR DE LA DEFENSA"

Evolución y tendencias del uso de la inteligencia artificial en la guerra híbrida y cognitiva

Gonzalo León
Catedrático emérito UPM, miembro del FEI

David Ramírez
Analista del Instituto Español de Estudios Estratégicos IEEE, CESEDEN

Ángel Gómez
Coronel (R) del Ejército del Aire y del Espacio
Director Europavia Middle East

Septiembre 2025

Tabla de contenido

"NATO faces a new kind of threat: a war fought not with bombs and missiles, but with lies and manipulation" (July 3, 2024)
https://www.act.nato.int/article/cogwar-concept/

"Pelear y conquistar en todas las batallas no es la suprema excelencia; la suprema excelencia consiste en romper la resistencia del enemigo sin luchar".
El arte de la Guerra (entre el 722 y el 481 a.c). Sun Tzu

1. Introducción

1.1. La tecnología como arma geopolítica

La historia enseña que la **confrontación entre naciones** puede tener fundamentos muy diversos. Desde el meramente económico, derivado de la asimetría de riqueza entre naciones, a otros estimulados por razones territoriales de expansión demográfica, por la necesidad de acceso a materias primas, alimentos o productos manufacturados, o por el control de las rutas marítimas o terrestres. A ellos se suman razones históricas de enemistad, con problemas latentes entre comunidades o países no resueltos a lo largo de la historia, que dependen de equilibrios inestables que pueden derivar periódicamente en conflictos armados.

Ante este panorama, todos los países han asumido como propio el objetivo de disponer de los medios necesarios, tecnológicos o no, dentro de sus posibilidades o con alianzas externas, para asegurar la **superioridad militar** ante potenciales conflictos convencionales de alta intensidad o en otros encubiertos.

Asegurar el acceso a bienes y servicios esenciales en situaciones de crisis se ha convertido, junto a la necesidad de disponer del **conocimiento sobre tecnologías emergentes**, en un factor clave para lograr la competitividad internacional y disponer de un alto nivel de bienestar en una **sociedad tecnológica** como son las que caracterizan a los países desarrollados. No es extraño, por ello, que la interacción entre todos los elementos citados haya provocado que el uso de estos factores se haya convertido en un **"*arma*" en la batalla geopolítica entre países**, afectando a la **"seguridad" colectiva**.

Como consecuencia, en la sociedad actual **"todo" se ha convertido o se puede convertir en un "arma"** empleada en la **confrontación** entre naciones, organizaciones o personas con diferentes formas de pensar o con intereses contrapuestos. Específicamente, el **control del acceso a múltiples componentes básicos necesarios para el funcionamiento de la sociedad** se ha empleado en los conflictos geopolíticos entre países para imponer determinadas posiciones de fuerza de manera permanente o coyunturales ante determinados eventos o procesos negociadores.

La amplitud del fenómeno y su diversidad pueden verse en la figura 1, en la que se indican algunos de estos factores empleados como *arma*. Los aspectos indicados en la figura se han empleado en la historia de la humanidad para **imponer determinadas visiones de las naciones que ostentaban el poder en cada momento**, y, con ello, alterar la forma en la que imponían sus propuestas políticas, religiosas, culturales, o económicas a otras naciones bajo su (pretendido) yugo, con objetivos de dominación a largo plazo o para conseguir objetivos tácticos a corto plazo; todos los aspectos indicados van más allá y se superponen, si fuera necesario, al mero empleo de la fuerza militar.

Muchos de ellos no son novedosos; tienen un **largo recorrido histórico,** como es el caso del **acceso al agua**, **a los alimentos**, o a **productos manufacturados**. En todos ellos, el control de las **infraestructuras de transporte** (terrestre, aéreo o marítimo) juega un papel destacado.

También se ha visualizado el empleo como arma geopolítica del **control provocado del movimiento masivo de personas**; hecho no solo acaecido muchas veces en la historia pasada, sino que está, desgraciadamente, omnipresente en la actualidad en la evolución del conflicto de Gaza con Israel, pero también con menor repercusión en los medios occidentales en Sudán, Birmania, Ruanda, y otros lugares de África, Asia y América. La UE también lo ha sufrido con movimientos de migrantes en la frontera con Bielorrusia (en 2021 contra Lituania) y en el Mediterráneo, implicando a mafias de tráfico de personas cuyo apoyo por países se inserta en los vaivenes de la confrontación geopolítica. A ello se une el uso de **civiles como "escudos humanos"** en operaciones militares que sigue empleándose en zonas de conflictos.

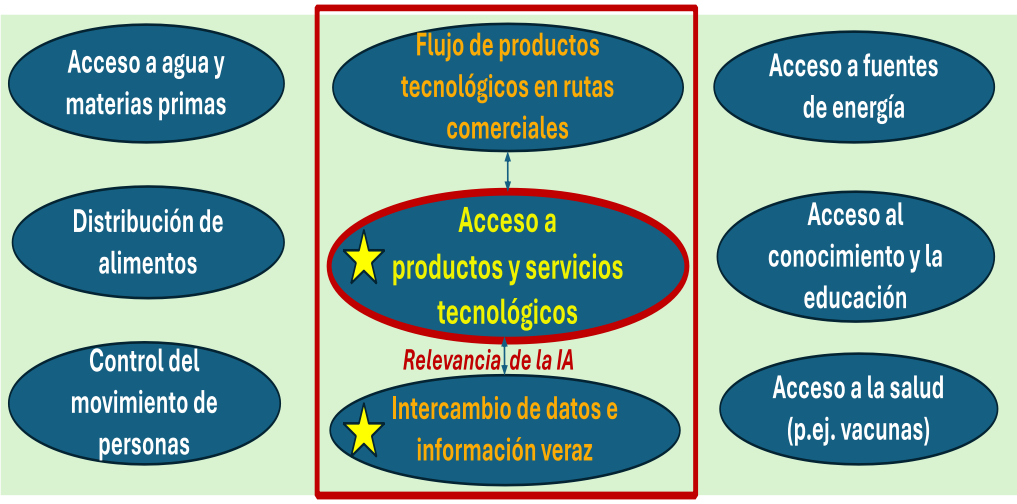


Figura 1. Componentes básicos para el funcionamiento de la sociedad empleados en la confrontación geopolítica. Fuente: elaboración propia

Otros factores de base tecnológica han surgido más recientemente afectando al desarrollo de una sociedad avanzada, como son el **acceso a la salud** (p.ej. mediante el control de la generación y difusión de medicamentos y vacunas), **a la energía** (p.ej. actuando sobre las fuentes de materias primas o sobre sus redes de distribución) o, con un impacto más alejado en el tiempo, el **acceso al conocimiento y a la educación** (p.ej. con medidas discriminatorias sobre el acceso de las mujeres a las escuelas o universidades).

En una sociedad altamente tecnificada como es la actual, algunos de los **elementos directamente ligados a la tecnología**, ya sean productos, servicios, procesos o el conocimiento tecnológico asociado a ellos, han adquirido una relevancia fundamental para **asegurar el funcionamiento de la sociedad** en su conjunto. No es extraño, por ello, que las **infraestructuras tecnológicas**, ya sean energéticas, de telecomunicaciones, de producción, o de transporte, se conviertan para actores estatales o no en **objetivos de la acción directa e indirecta en conflictos** como parte de procesos de desestabilización frente al "*enemigo*", que afecta a la sociedad en su conjunto, sin que eso implique la existencia de un conflicto militar abierto de alta intensidad, aunque pueda ser la antesala de ello.

La experiencia reciente en la **guerra de Ucrania**, valorado como un conflicto militar de alta intensidad desde la invasión rusa en 2022, visualiza este hecho con consecuencias que afectan

a otros muchos países no directamente implicados en el campo de batalla. Recuérdese, por ejemplo, el problema geopolítico derivado de los **envíos de grano ucraniano en el Mar Negro** (afectando a muchos otros países de África) o de **suministros del gas ruso** (afectando fuertemente a la UE), o los **ataques a cables submarinos**, ya sean de electricidad o de comunicaciones. En esos casos, bienes físicos como las terminales portuarias, los buques de carga, los oleoductos o los cables submarinos internacionales, se han visto convertidos en armas con objetivos geopolíticos.

Su amplitud permite valorar la **relevancia del fenómeno de "armamentización" de las infraestructuras tecnológicas más allá de las zonas de conflicto.** Los prolongados **ataques rusos a la infraestructura energética de Ucrania** o los de Ucrania a las equivalentes de Rusia constituyen un ejemplo en este sentido, porque también afectan a países limítrofes, empleando el posible ataque a centrales nucleares como una amenaza constante de peligro mayor[1].

No se trata únicamente de resaltar su relevancia en conflictos militares declarados, sino también en la **importancia adquirida por la tecnología en la situación de guerra híbrida** extendida y facilitada por la propia evolución de la tecnología, con repercusiones que afectan a todos los países. Una consecuencia de ello es la creciente perpetración de **sabotajes** sobre infraestructuras tecnológicas, de **ciberataques** contra entidades gubernamentales o grandes empresas (inutilizando servicios públicos esenciales y accediendo o robando datos sensibles) o a las sofisticadas campañas de **desinformación** alentadas por gobiernos u organizaciones paraestatales, con **actuaciones de desestabilización altamente tecnificadas** que han estallado en múltiples zonas del planeta (Neuberger, 2025).

En el contexto del presente documento, se desea prestar especial atención a la forma en la que el **acceso a productos y servicios tecnológicos**, tanto desde una faceta física (como es el control del acceso y flujo de productos tecnológicos a través de rutas comerciales) como desde una faceta lógica (como asegurar sin interferencias el intercambio de datos a través de cables submarinos o comunicaciones espaciales, y la veracidad de la información generada y diseminada mediante múltiples servicios digitales) se ven implicadas en conflictos denominados **"híbridos"** en los que **la detección y atribución de acciones es un problema** en sí mismo, como se caracterizará posteriormente en el presente documento.

La **faceta lógica**, basada no solo en asegurar el intercambio de datos y el acceso a información, sino también la **modulación intencionada de la opinión** de personas e instituciones ha cobrado una dimensión de gran relevancia al hilo del **concepto de "guerra cognitiva"**; se trata de influir en las narrativas colectivas y en las tomas de decisión derivadas. Es este aspecto en el que se focaliza específicamente el presente documento, en su interacción con conceptos próximos como son los de *guerra híbrida*, *zona gris*, *guerra informativa* (*desinformación)*, todos ellos en su relación con la *guerra cognitiva*.

Desde una perspectiva tecnológica, las herramientas de ejecución de todos los elementos indicados están profundamente potenciadas por el desarrollo y despliegue de herramientas de **inteligencia artificial (IA),** cuyo objetivo está directamente ligado a la generación y procesamiento de información en una escala e impacto muy superior al existente con las herramientas digitales habituales. El **doble carácter dual y habilitador de la IA** convierte su acceso y gobernanza en un problema de carácter estratégico para el que no existe un marco

[1] Es relevante indicar que los primeros acuerdos promovidos por Estados Unidos desde marzo de 2025, sin un cumplimiento efectivo, por ahora, se hayan centrado en "*treguas temporales*" de no atacar infraestructuras energéticas. La situación en septiembre de 2025 no ha cambiado sustancialmente.

global consensuado. Actualmente, **el uso de la IA se ha convertido en un factor esencial en la "batalla tecnológica" entre grandes potencias**.

Por supuesto, una estrategia basada en **utilizar la tecnología disponible para "influir" en la población objetivo y cambiar su percepción de una situación no es nueva**; se relaciona con el antiguo concepto de "*propaganda*" empleado por los imperios en la Antigüedad con los medios disponibles en su momento; asumiendo que el objetivo no tiene por qué ser toda la población sino aquélla con capacidad de decisión en zonas alcanzables con los medios disponibles[2].

Pero lo que no tiene precedentes, y explica la urgente atención que ha alcanzado en el discurso político y de seguridad en Europa, es el **desarrollo y ubicuidad de los medios digitales de comunicación**, ahora **apoyados por la IA** (Lahman, 2024) empleados por los gobiernos, que permite acelerar la toma de decisiones con efectos sobre las personas[3].

Para algunos autores (Impiombato et al., 2024), no se trata solo de considerar el impacto de la IA de manera aislada, sino de su convergencia con otras tecnologías emergentes. La rápida adopción de lo que denominan **"tecnologías persuasivas"** (entendidas como cualquier sistema digital que moldee las actitudes y comportamientos de los usuarios explotando las reacciones o vulnerabilidades fisiológicas y cognitivas) pondrá a prueba la seguridad nacional de formas difíciles de predecir: "*Las tecnologías persuasivas emergentes, como la inteligencia artificial (IA) generativa, las tecnologías y la neurotecnología interactúan con la mente y el cuerpo humanos de formas mucho más íntimas y subconscientes, y a una velocidad y eficiencia mucho mayores que las tecnologías anteriores*"[4].

En la figura 2 se representa un esquema conceptual de los elementos que enmarcan la **gobernanza tecnológica** desde la perspectiva de su dualidad y desde los pilares técnico, regulatorio y geopolítico (León, 2024). En la figura se ha remarcado el ámbito específico de interés del presente documento focalizado en el **papel de la IA en la guerra híbrida y cognitiva**. Obsérvese en la figura 2 la importancia que adquieren en términos de **intensidad** de la influencia y del **impulso** para un desarrollo tecnológico dos elementos clave:

- el ya referido previamente de la **tecnología**, no solo la digital, aunque sea la más relevante, como parte de la batalla por la supremacía en términos geopolíticos, y
- la influencia de **comportamientos asimétricos entre contendientes** en los que las consideraciones de defensa toman protagonismo en una visión sociopolítica más amplia, modulada por el tipo y capacidades de los actores intervinientes.

---

[2] En épocas históricas en las que el porcentaje de población alfabetizada era muy reducido, el uso de las obras públicas, la numismática y lo eventos culturales de masas (como el teatro o los acontecimientos deportivos) eran un medio apropiado. Actualmente, no solo han perdido vigencia en una sociedad alfabetizada, sino que, transformados en eventos digitales, algunos creados sintéticamente, tienen un objetivo similar.

[3] En un estudio para el Parlamento Europeo (Unvar, 2024) se acuña el término de ***Autoritarismo algorítmico*** (*Algorithmic authoritarianism*) referido a un sistema más amplio de gobernanza o control en el que los algoritmos (de IA) desempeñan un papel central en los procesos de toma de decisiones y en los mecanismos de vigilancia y control.

[4] La relación entre la IA y las neurotecnologías serán tratadas en una sección posterior del presente informe.

Estos dos elementos citados se encuentran, además, en la **batalla por el control del desarrollo futuro de la IA** con impactos relevantes sobre el conjunto de la sociedad en los tres ejes verticales indicados en la figura: el **técnico**, el **regulatorio** y el **geopolítico,** con influencias crecientes sobre el regulatorio como herramienta a disposición de los Estados para encauzar su desarrollo y limitar sus potenciales efectos perniciosos.

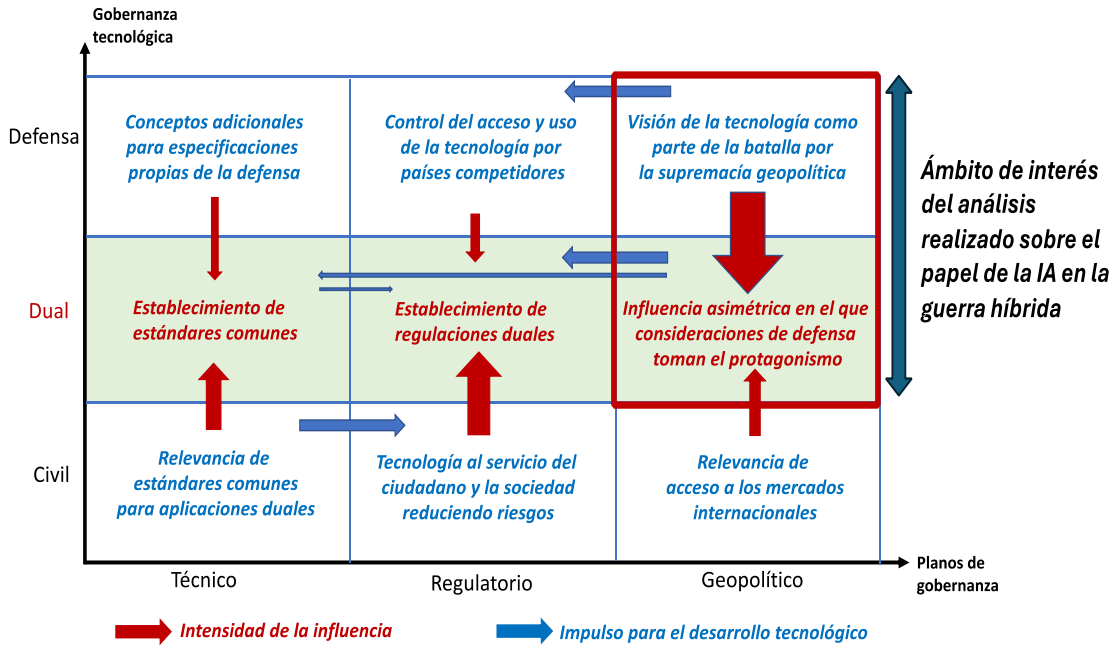


Figura 2. Pilares de la gobernanza de la tecnología desde la perspectiva de su dualidad. Fuente: elaboración propia adaptada de León (2024).

En un contexto más amplio, el **uso de la IA en el diseño y operación de sistemas de defensa** ha crecido fuertemente en la última década en diversas áreas relacionadas con el procesamiento de la información relevante para la defensa como son las siguientes:

- Sistemas parcialmente automatizados de **apoyo a la toma de decisiones** en la planificación y gestión del campo de batalla basados en la captura, análisis y gestión de grandes volúmenes de información (p.ej. telemetría, imágenes, datos), muchos de ellos capturados en tiempo real desde diversos tipos de sensores integrados, para la toma de decisión desde un enfoque de operación multidominio.
- Desarrollo de **sistemas robóticos inteligentes,** como vehículos militares autónomos (terrestres, marítimos o aéreos) operando individualmente o en "enjambres" de sistemas autónomos con comportamiento cooperativo.
- Sistemas inteligentes de **guerra electrónica** para conseguir el dominio del espectro electromagnético (protegiendo el propio y perturbando el del oponente) en redes de comunicaciones, ya sean fijas, móviles o satelitales (individuales o formando constelaciones), o sistemas de navegación u observación, empleando e integrando diversas tecnologías.
- Sistemas de **ciberdefensa inteligente** diseñados para hacer frente (evitar, anticipar, reducir el impacto o recuperarse) a ataques cibernéticos lanzados contra sistemas críticos para la sociedad (no necesariamente militares) en los que se sustenta el funcionamiento de la sociedad.

- Inteligencia militar basada en el **análisis de información** multifuente procedente de fuentes abiertas o clasificadas contenidas en servidores, redes sociales y otros tipos de servicios digitales.
- Generación y gestión automatizada de **narrativas y campañas de desinformación** focalizadas en objetivos o grupos de interés, y herramientas de detección y minimización del impacto de esas campañas.

En todos los ámbitos indicados, la adopción de sistemas basados en tecnologías de IA permite **mejorar las prestaciones** de los sistemas preexistentes, la **velocidad** a la que pueden procesar ingentes volúmenes de **datos**, y la posibilidad con ello de permitir el **análisis y la toma de decisiones** en entornos multidominio competidos, factores que son imposibles de lograr con las capacidades de un operador humano.

En un estudio llevado a cabo por *RAND Corporation* (Black et al., 2024) se analizan los **riesgos del uso de la IA en conflictos militares y las acciones necesarias para mitigarlos**, y explotar las oportunidades generadas. En todo caso, será un proceso paulatino con una fuerte componente de competición estratégica entre grandes potencias. La tabla 1 presenta la visión de RAND (Black et al., 2024) sobre el tipo de herramientas a emplear (adaptado al caso del Reino Unido) y las acciones prioritarias.

CATEGORÍA	ACCIONES PRIORITARIAS
Mecanismos para impulsar la adopción de la IA y beneficios para la Defensa del RU	*Acelerar la inversión y la adopción de la IA en toda la defensa, al tiempo que aumenta la resiliencia frente al uso indebido hostil o accidental de la IA*
Mecanismos para restringir la adopción de la IA y beneficios para adversarios	*Adoptar un enfoque de campaña para restringir, ralentizar o aumentar los costos para los adversarios (estatales o no estatales) del despliegue de la IA militar.*
Mecanismos para dar forma a los nuevos acuerdos de gobernanza de la IA militar	*Desempeñar un papel de liderazgo en la concienciación, la detección de problemas y el intercambio de conocimientos sobre los riesgos de la IA militar.*
	Desarrollar medidas de transparencia y fomento de la confianza con aliados clave (por ejemplo, EE. UU.) y competidores (por ejemplo, China) para reducir los riesgos de escalada
	Promover un enfoque inclusivo y participativo para construir un consenso global emergente sobre normas de comportamiento responsable en torno a la IA militar, como preludio a acuerdos vinculantes más sólidos en el futuro.
	Promover el desarrollo paralelo de mecanismos multilaterales para reducir los riesgos urgentes de IA relacionados con la energía nuclear y biológica
	Investigar formas de incorporar la IA en los mecanismos de verificación y cumplimiento, y viceversa
	Con el tiempo, consolidar el actual panorama fragmentado de las iniciativas de gobernanza de la IA en una arquitectura más concreta

Tabla 1. Acciones prioritarias para el uso de la IA en defensa. Fuente: adaptada de https://www.rand.org/pubs/research_reports/RRA3295-1.html .

Debe tenerse presente que **el nivel de madurez y la tasa de adopción de sistemas basados en la IA varía significativamente de un país a otro**, y que, además, todas ellas evolucionan en el tiempo muy rápidamente; sobre todo, cuando esta evolución se produce en periodos de conflictos geopolíticos y militares de creciente virulencia en los que el propio conflicto militar muta, como ha ocurrido en el pasado con grandes revoluciones tecnológicas en el campo de batalla. En el momento actual, se mantienen significativas incertidumbres sobre su evolución futura puesto que las tecnologías que provocan este cambio, y su adopción, son aún inmaduras.

En el contexto del presente documento, se desea reflejar como aspecto prioritario que la importancia de la IA no solo ha crecido en su uso en **conflictos militares convencionales entre Estados** como los indicados previamente, sino también en **conflictos asimétricos** y, especialmente, ante la denominada **"guerra híbrida"**. Posteriormente, se analizará su evolución en un concepto ya consagrado en el contexto de la OTAN como es el de "**guerra cognitiva",** sobre el que se focaliza el presente documento.

1.2. Concepto y contexto de la guerra híbrida

La **guerra híbrida** es un término que intenta capturar la complejidad de la guerra del siglo XXI, que involucra una multiplicidad de actores y difumina las distinciones tradicionales entre tipos de conflictos armados y entre guerra y paz. Responde a una visión perturbadora: *"El principal campo de batalla está en la mente y, como resultado, las guerras de nueva generación deben estar dominadas por la información y la guerra psicológica. ... El objetivo principal es reducir al mínimo necesario el despliegue de un poder militar duro".*

El **origen de la guerra híbrida** es antiguo, y se puede ver ya reflejado en contiendas del siglo XIX[5]. Otro hecho histórico destacado de la evolución de la guerra híbrida se puede asociar a la revuelta árabe provocada durante la I Guerra Mundial, en donde las fuerzas militares británicas idearon una formulación de fuerzas irregulares y operaciones convencionales que T.E. Lawrence hizo famosas. En Cominotto (2025) pueden encontrarse muchos más ejemplos históricos.

Más recientemente, el concepto fue enunciado por Thomas Huber (1996) y definido inicialmente por William J. Nemeth en 2002, que planteó la hipótesis de que la guerra híbrida consiste en una **combinación creativa de despliegue sincronizado no militar y militar-estratégico**. Los esfuerzos no militares incluyen el combate psicológico y social de la propaganda, las noticias falsas, la diplomacia y la intervención electoral (Querishi, 2020).

Aunque fue Frank Hoffman (2007) quien, en su libro *"Conflict in the 21st Century: The Rise of Hybrid Wars"*, lo reformula en términos cercanos a la situación actual. Para Hoffman, la **guerra híbrida** puede definirse como la *"combinación en el tiempo y en el espacio de fuerzas convencionales e irregulares, tácticas terroristas, desórdenes criminales y métodos no militares como la guerra cibernética, la desinformación y la coerción económica. Con ella se busca obtener ventajas psicológicas y físicas mediante el uso integrado de herramientas militares y no militares,*

[5] En la denominada *Guerra de la Independencia* española al comienzo del siglo XIX las guerrillas españolas que luchaban contra el ejército francés actuaban "coordinadamente" con el ejército español y/o con el ejército inglés desplegado en España, hostigando, apoyando en operaciones logísticas propias o perturbando las adversarias, y proporcionando información. Pero no actuaban fusionadas en el campo de batalla, ni tampoco gozaban de una estructura jerarquizada con mando único.

incluyendo operaciones de información, electrónicas y cibernéticas, así como presión económica y acciones de inteligencia".[6]

Esta definición tan amplia incluye muchos tipos y niveles distintos de *hibridación*. Algunos autores[7] distinguen entre **tres tipos de situaciones "híbridas"** (amenaza, conflicto y guerra) que aluden a una **gradación en intensidad y en el impacto social y militar perseguido**:

- **Amenaza híbrida (***hybrid threat*): Fenómeno resultante de la convergencia e interconexión de diferentes elementos que, en conjunto, constituyen una amenaza más compleja y multidimensional.
    - La "amenaza" es empleada como elemento de disuasión o de posicionamiento ante procesos negociadores en los que se pretende sacar provecho de la reacción intimidada del contrincante.
    - Un ejemplo clásico es la amenaza del uso del arma nuclear frente a otro país o países que no la poseen[8].
- **Conflicto híbrido (***hybrid conflict*): Situación en la cual las partes se abstienen del uso abierto de la fuerza (armada) y actúan combinando la intimidación militar (sin llegar a un ataque convencional) y a la explotación de vulnerabilidades económicas, políticas, tecnológicas y diplomáticas.
    - En este nivel se produce una planificación y sincronización de acciones diversas en un contexto de control no solo militar, sino también político, puesto que es necesario integrar áreas competenciales muy diversas.
    - Generalmente, se combinan con acciones en organismos internacionales para conseguir los apoyos necesarios de terceros países para obtener resoluciones favorables a una determinada narrativa sobre el conflicto de que se trate.
- **Guerra híbrida (***hybrid war*): Situación en la que un país recurre al uso abierto de la fuerza (armada) contra otro país o contra un actor no estatal, junto al uso de otros medios de coerción (por ejemplo, económicos, políticos o diplomáticos) combinados con operaciones, encubiertas o no, de menor intensidad, que afectan a servicios básicos de la sociedad.
    - Se pretende con ello afectar al conjunto de la sociedad del país afectado para acelerar su derrota e, indirectamente, a otros que le apoyan.
    - Generalmente, va ligado y complementa a amenazas y conflictos híbridos[9].

---

[6] https://www.marshallcenter.org/sites/default/files/files/2020-05/pC_V10N1_en_Wither.pdf

[7] https://campus-stellae.com/que-son-las-guerras-hibridas/

[8] Como ejemplo, el recurso en una confrontación al empleo de la fuerza nuclear como una amenaza "*si se ponen en riesgo aspectos considerados como existenciales*" de un país relega su interpretación por terceros a términos subjetivos en los que los errores de cálculo pueden tener consecuencias devastadoras. En el caso de la guerra de Ucrania, Rusia actualizó su doctrina nuclear en noviembre de 2024, dos días después de que el presidente Biden aprobara el uso de misiles estadounidenses para ataques ucranianos dentro de Rusia. La nueva doctrina establece que "*la agresión de cualquier Estado de una coalición, bloque o alianza militar contra la Federación Rusa o sus aliados se considerará una agresión",* expresamente que el uso de misiles occidentales no nucleares por parte de las Fuerzas Armadas ucranianas contra Rusia en virtud de la nueva doctrina podría conllevar una respuesta nuclear. https://es.euronews.com/2024/11/19/putin-advierte-que-el-uso-de-misiles-contra-rusia-puede-implicar-una-respuesta-nuclear

[9] El caso de Rusia en el conflicto de Ucrania es un caso paradigmático de su uso para evitar el despliegue de determinados sistemas de armas proporcionadas por Occidente o su uso en territorio ruso, al mismo tiempo que sistemas similares sí son empleados diariamente por Rusia en todo el territorio ucraniano.

Murray (2021) analiza la **relación entre la guerra híbrida y la tecnología** asumiendo una evolución paralela. Para este autor *"la evolución de la guerra ha introducido nuevas metodologías en el combate, lo que supone un cambio significativo en la trayectoria de la evolución de la guerra, ya que se asignan recursos para contrarrestar la influencia de los actores no estatales en lugar del potencial militar de poder duro de los Estados. En resumen, la teoría explora el desarrollo de la guerra híbrida a medida que los avances tecnológicos han evolucionado el curso de la guerra mediante la introducción de medios irregulares de combate social y psicológico".*

Muy brevemente (Murray, 2021) la evolución de la guerra puede verse en forma de "generaciones" sucesivas.

- La ***1ª generación*** emergió del Tratado de Westfalia en 1648, con el concepto de integridad territorial y el monopolio del Estado en la conducción de las guerras. Además, se estableció una cultura de orden y organización.
- La ***2ª generación*** fue introducida por el ejército francés y terminó después de la Primera Guerra Mundial. La cultura del orden continuó, mientras que la mano de obra masiva fue reemplazada por la potencia de fuego masiva para dominar el campo de batalla. La razón de esta transición fue la introducción de artillería, aviones y fuego pesado. Este desarrollo resultó en la transición de la mano de obra a la potencia de fuego.
- La ***3ª generación*** fue desarrollada por Alemania en la Segunda Guerra Mundial. Se introdujo el concepto de guerra de maniobras con tácticas de sorpresa para eludir y socavar al enemigo a través de la velocidad, el sigilo y la sorpresa (campañas relámpago) con el uso de tanques, artillería y aviones de combate. Las tácticas se centraron en la relevancia del "tiempo" en comparación con las centradas en el "espacio físico".
- La ***4ª generación*** ha ganado protagonismo en las últimas seis décadas, con la introducción de actores no estatales como parte de la guerra y la emergencia del terrorismo alterando la dinámica de la guerra. Como resultado, terminó el monopolio estatal sobre la guerra, las fuerzas militares tradicionales fueron pasadas y las poblaciones civiles se convirtieron en un objetivo directo.
- La ***5ª generación*** ha remodelado claramente el arte y la filosofía de la guerra al introducir una batalla de percepciones e información. Los avances tecnológicos han alterado claramente la conducción de la guerra. Bajo este sistema, los Estados rivales son manipulados a través de la información, lo que en última instancia genera inestabilidad en los Estados.

En esta visión evolutiva, el concepto de guerra hibrida emerge desde la 4ª generación de la conducción de la guerra y se apoya y consolida con el enorme desarrollo tecnológico que ha tenido lugar en la 5ª generación.

Relacionado con el concepto de guerra híbrida, algunos autores[10] incluyen el concepto (más borroso) de **"zona gris"** definida como: "*zona del espectro de los conflictos donde predominan las actuaciones situadas al margen del principio de buena fe entre Estados (bona fide) que, pese a alterar notablemente la paz, no cruzan los umbrales que permitirían o exigirían una respuesta armada"*. El concepto de "zona" no se refiere, al menos no exclusivamente, a un espacio (territorio) definido geográficamente, sino a un **espacio lógico de confrontación**.

[10] https://www.ieee.es/Galerias/fichero/docs_opinion/2022/DIEEEO34_2022_LUISHER_Zona.pdf y/o enlace bie3

La **zona gris es un espacio funcional entre la guerra y la paz**, donde las jurisdicciones se difuminan, se disputan o no se dejan claras y donde las responsabilidades y la rendición de cuentas son vagas y puestas formalmente en duda. Es un espacio (físico y lógico) en el que florece la guerra híbrida y las operaciones por debajo del umbral, porque es más difícil saber si se ha producido un ataque y quién podría ser el responsable[11].

El término, en aspectos geográficos, suele emplearse actualmente en el caso de **jurisdicciones disputadas y superpuestas**, por ejemplo, en el Ártico o en el Mar de China Meridional, en las que **se dificulta la identificación clara de las responsabilidades**. También surge en zonas que han sido "*anexionadas*" formalmente por uno de los contendientes, aunque su ocupación real no sea absoluta ni exista un reconocimiento internacional (es el caso de atolones ocupados y militarizados por China o en zonas de Ucrania invadidas por Rusia)[12].

Kormych et al. (2023) abogan por la necesidad de abrir una vía que permita resolver los conflictos de la zona gris y prevenir la escalada mediante la **creación de mecanismos y procedimientos internacionales de supervisión para establecer hechos y atribuir acciones a los diferentes actores**. Esto ayudaría a superar la ambigüedad jurídica y a establecer una norma común de reconocimiento de las fuentes jurídicas pertinentes, lo que ayudaría a formular y aplicar medidas para resolver el conflicto.

Otra **definición relevante de guerra híbrida**, abundando en la ambigüedad que le caracteriza, es la procedente del ***Centro Europeo de Excelencia para la Lucha contra las Amenazas Híbridas*** (*Hybrid CoE*)[13], establecido en 2017 en Helsinki conjuntamente por la OTAN y la Unión Europea. El *Hybrid-CoE* describió tales acciones como aquéllas: *"caracterizadas por la ambigüedad, ya que los actores híbridos difuminan las fronteras habituales de la política internacional y operan en las interfaces entre lo externo y lo interno, lo legal y lo ilegal, y la paz y la guerra".*

Se trata, por tanto, de un **término borroso y difícil de interpretar** puesto que los "*umbrales*" de aplicación no están definidos *a priori*, de forma objetiva, sino que varían en el tiempo y dependen del contexto de cada bando en conflicto que permite valorar su gravedad. Además, como señala Ålander (2024), no debe considerarse únicamente la gravedad de un incidente aislado, sino la **reiteración de decenas o centenares de pequeños incidentes en un breve lapso** para detectar una situación de zona gris y la existencia de una guerra híbrida de suficiente entidad.

Con la figura 3 se intenta visualizar este **espectro de actuaciones** representado como una secuencia de acciones que van desde la **paz** al **conflicto armado**, y desde este al **postconflicto**. Como se indica en la figura 3, se trata de un proceso temporal en el que el **nivel de escalada** (incremento de la intensidad del conflicto) fluctúa, y se ve salpicado de diversas "crisis" y acciones de "desestabilización" en el transcurso del tiempo. Esta gradación no es una cualidad de la propia guerra híbrida, donde se utilizan medios y recursos de uno u otro dominio a criterio del decisor, sino del **nivel de enfrentamiento que se desea alcanzar con esas acciones**. Para

---

11 https://theconversation.com/russia-and-the-west-are-entering-the-grey-zone-of-warfare-and-the-oceans-are-a-key-battleground-244668

12 El caso de las cuatro provincias de Ucrania que Rusia ha considerado como propias, pero no reconocidas así por muchos otros países, es un ejemplo de esta situación. También puede serlo zonas de Cisjordania ocupadas por Israel por asentamientos de colones que tampoco son reconocidas internacionalmente.

13 https://www.hybridcoe.fi/

estos autores, hablar de zona gris no tiene sentido cuando se ha desencadenado un conflicto armado, aunque puede seguir siendo "gris" respecto al ámbito jurídico legal aplicable.

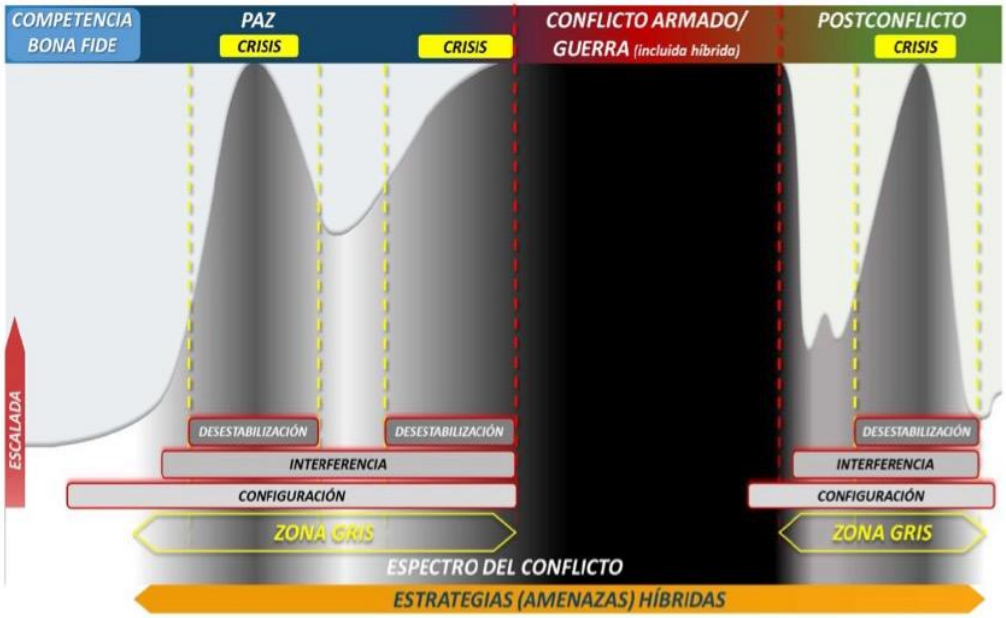


Figura 3. Espectro de actuaciones híbridas. Fuente: https://www.ieee.es/Galerias/fichero/docs_opinion/2022/DIEEEO34_2022_LUISHER_Zona.pdf y/o enlace bie3

La tabla 2 (Pandey, 2023) resume la relación y diferencias entre un conflicto en zona gris y una guerra hibrida atendiendo al nivel, el uso de operaciones militares, convencionales o no, y la implicación prolongada o no de los actores.

Characteristic	Grey-Zone Conflict	Hybrid Warfare
Level	Tactical, operational, strategic	Tactical and operational
Use of conventional military operations	Used alongside non-conventional operations.	Used alongside non-conventional operations. Usually the dominant element.
Use of non-conventional military operations	May be used standalone or alongside conventional operations.	Used alongside conventional operations as auxiliary tactics.
Protracted engagement	One of the dominant characteristics.	May be protracted or short

Tabla 2. Comparación entre conflicto en zona gris y guerra híbrida. Fuente: Pandey (2023)

Un factor que alcanza gran relevancia en la implementación de una estrategia de guerra híbrida se refiere a la **sincronización y escalada de acciones**, factor necesario para que tenga éxito, tanto desde una perspectiva de *escalada horizontal* (es decir, aquélla que tiene lugar entre diferentes dominios del poder en una sociedad: militar, político, económico, civil, o de información), como de *escalada vertical* (con diferentes niveles de intensidad y visibilidad de la actuación híbrida en el conjunto de la población).

En la figura 4 se indican diversos **instrumentos de poder que deben sincronizarse** para asegurar el éxito en la guerra híbrida: los **militares**, **políticos**, **económicos**, **civiles** y de **información**. No se trata de elementos disjuntos y, de hecho, para obtener el éxito pretendido, es necesario

asegurar su interacción y adaptación continua en función de la evolución temporal de la intensidad del conflicto.

En la figura 4 se han incorporado, asimismo, dos **umbrales de visibilidad** que son relevantes para caracterizar con precisión este fenómeno:

El denominado ***umbral de detección***, que se refiere al nivel mínimo de intensidad de la acción híbrida que permite a un gobierno u organización detectar que se está produciendo una amenaza, conflicto o guerra de carácter hibrido; si el umbral se sitúa en un punto demasiado alto, habrá situaciones no detectadas a tiempo, pero si el umbral está demasiado bajo, "todos los eventos" que sucedan serán considerados como guerra híbrida, saturando los sistemas de respuesta. Téngase en cuenta que un gobierno puede detectar acciones de guerra híbrida en su contra, pero, no necesariamente puede desear hacer partícipe de ello a la población en su conjunto, con objeto de controlar el nivel de alarma social. También influye en este caso la actividad y posicionamiento de los medios de comunicación.

Y el ***umbral de atribución***, referido al nivel mínimo de intensidad que permite atribuir (una vez detectado) quién es el actor que la provoca para poder ejercer acciones (como las de represalia) con solidez y no rebatibles[14]. Este umbral se utiliza por los actores implicados de una forma ligada a su propia estrategia y los datos obtenidos.

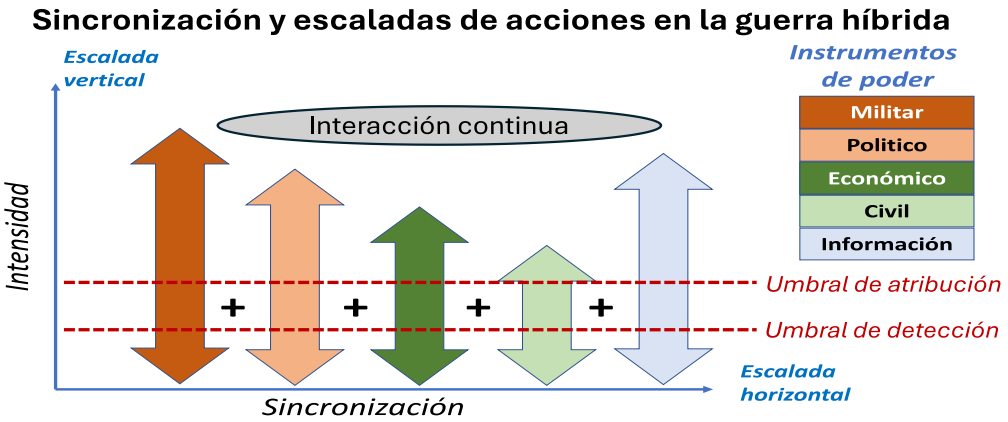


Figura 4. Sincronización y escalada de acciones en la guerra híbrida. Fuente. Adaptada de https://assets.publishing.service.gov.uk/media/5a8228a540f0b62305b92caa/dar_mcdc_hybrid_warfare.pdf

Debe tenerse en cuenta que, **para el actor que provoca una guerra híbrida**, generalmente gubernamental, paraestatal o ligado con un gobierno, **no siempre es deseable que su acción sea visible**, y no necesariamente desea que sea atribuible; comportamiento diferente al exhibido habitualmente por un grupo u organización terrorista que puede estar **interesado en reclamar la autoría de una acción**, incluso en casos en los que no la haya realizado, con el fin de incrementar su visibilidad, o como herramienta de fidelización, proselitismo y atracción de nuevos adeptos. Es, de nuevo, un ejemplo del **escenario de borrosidad** imperante y, como tal, se desea preservar por muchos actores implicados.

[14] Obsérvese que el umbral de atribución suele implicar un nivel de intensidad más alto que la mera detección para poder trazar su origen, dado que hay un interés explicito en que la acción no se pueda "atribuir" de manera explícita, como ocurre en muchas acciones sobre el ciberespacio.

Un concepto asociado al de guerra hibrida es el de "**actor no estatal**". Se refiere a *"entidades que desempeñan un papel en las relaciones internacionales y que ejercen suficiente poder para interferir, influir y provocar cambios sin ninguna afiliación a las instituciones establecidas de un Estado".* Las clasificaciones de los actores no estatales incluyen desde individuos hasta empresas privadas, instituciones religiosas, organizaciones humanitarias, grupos armados y regímenes de facto que tienen el control real del territorio y la población (Rauta, 2025). El ejemplo prototípico en el caso de Ucrania es el "*Grupo Wagner*"[15]. Las situaciones en el Sahel o en Oriente Medio implican a este tipo de actores con fuerte capacidad de desestabilización.

Se ha comentado previamente que **el concepto de guerra híbrida** ha adquirido más relevancia en las últimas dos décadas con la denominada 5ª generación **potenciada por la tecnología**. Es el momento de profundizar en ello.

Para el *Centro Europeo de Excelencia para la Lucha contra las Amenazas Híbridas* (*Hybrid-CoE*)[16], **la tecnología es uno de los principales impulsores de la guerra híbrida y de las teorías de la guerra híbrida**; de hecho, ha sido la rápida evolución de las tecnologías de la información y las comunicaciones (TIC) y, en especial de la inteligencia artificial (IA), la que ha posibilitado e impulsado muchas de las actuaciones implicadas y su alto nivel de resonancia en una sociedad digitalizada como la actual.

El *Hybrid-CoE* caracteriza las amenazas híbridas como "*acciones coordinadas y sincronizadas que se dirigen deliberadamente a las vulnerabilidades sistémicas de los Estados e instituciones democráticas; actividades que explotan los umbrales de detección y atribución"*. El análisis de la situación en la región del Báltico derivado de la actuación hibrida de Rusia tras la invasión de Ucrania en 2022 es significativo para entender la **variedad de técnicas utilizadas y el uso sistemático y combinado de todas ellas** (Praks, 2024)[17].

Concretamente, el proyecto *Hybrid Warfare*: *Future and Technologies (HYFUTEC II)*[18] impulsado por el *Hybrid-CoE* tiene como objetivo **ampliar las teorías existentes de las tecnologías de doble uso para abordar los riesgos y oportunidades tecnológicos para actores híbridos**. El análisis

[15] En enero de 2023, el Ministerio de Defensa del Reino Unido había estimado que había hasta 50.000 mercenarios de Wagner en Ucrania. Cerca de 20.000 de ellos, muchos de ellos exprisioneros, pudieron morir en la batalla para tomar la pequeña ciudad de Bajmut en Ucrania. Tras la muerte de su líder Yevgeny Prigozhin en un accidente aéreo, algunos se han incorporado al ejército ruso y otros nutren las filas de grupos mercenarios en otras partes como en el Cáucaso, Siria o el Sahel. https://www.bbc.com/mundo/articles/c4gle5991dyo

[16] El *Hybrid-CoE* fue establecido en 2017 como una organización autónoma internacional con funcionamiento en red. Su origen se remonta en la Comunicación conjunta de la Comisión Europea y de la Alta Representante al Parlamento Europeo y al Consejo titulada: "Marco conjunto para luchar contra las amenazas híbridas: una respuesta de la Unión Europea", del 6 de abril de 2016 y aprobado por el Consejo de la Unión Europea y el Consejo de la OTAN el 6 de diciembre de 2016. Arrancó formalmente el 11 de abril de 2017, cuando los primeros nueve Estados participantes (Finlandia, Suecia, Reino Unido, Letonia, Lituania, Polonia, Francia, Alemania y Estados Unidos) firmaron el memorando de entendimiento. https://www.hybridcoe.fi/about-us/

[17] Operaciones de desinformación e influencia; Instrumentalización de las diásporas y de las instituciones vinculadas a Rusia; Sabotaje, vandalismo y actividades de inteligencia; Ataques cibernéticos; Instrumentalización de la migración; Amenazas a las infraestructuras energéticas y de comunicaciones.

[18] https://www.hybridcoe.fi/wp-content/uploads/2020/09/20200915_HYFUTEC_info.pdf

realizado en el proyecto *HYFUTEC* comprende **19 tecnologías** identificadas como particularmente relevantes para la evolución de los desafíos híbridos, el conflicto y la guerra[19]. Un informe específico de HYFUTEC se elaboró en 2021 sobre el uso de los drones en la guerra híbrida (Sprengel, 2021)[20].

El presente informe se centrará especialmente en el papel que juega la **IA y su relación sinérgica con otras tecnologías implicadas en la guerra hibrida**, distribuidas en tres grupos de tecnologías, tal y como se indica en la figura 5. Aunque ya existen algunos trabajos sobe el posible uso de las tecnologías cuánticas en el contexto de las amenazas híbridas (Bruze, 2021), no será abordado expresamente en el presente informe.

- Tecnologías orientadas a la **manipulación del acceso radioeléctrico a la información** mediante el uso de técnicas de guerra electrónica para generar interferencias de la señal radioeléctrica, suplantación de identidad u otros ciberataques como denegación de servicios en redes móviles (p.ej. en redes 5G, IoT), o en sistemas espaciales (p.ej. interfiriendo en la señal de navegación GPS o en las comunicaciones). En estos casos no se actúa sobre el contenido de la información sino sobre la capacidad de recibirla correctamente.

- Orientadas a la **manipulación de la información y su narrativa**, actuando sobre los servicios ofrecidos por las plataformas digitales (p.ej. interfiriendo en las redes sociales o los accesos a información contenida en la nube mediante cuentas falsas, bots, agentes inteligentes), generando "desinformación" mediante la generación de noticias falsas o tendenciosas apoyados por imágenes, texto, voz, vídeo, etc. que facilitan herramientas de IA generativa), y el incipiente uso de realidad extendida (p.ej. realidad virtual, realidad aumentada).

- Incipiente uso de **tecnologías emergentes** como las neurotecnologías (p.ej. mediante el uso de técnicas invasivas o no de modulación cerebral para la mejora de prestaciones), sistemas autónomos (p.ej. robots empáticos en equipos híbridos) o, en un próximo futuro, la interacción de la IA con las tecnologías cuánticas. Se ha incluido entre las tecnologías emergentes a la neurotecnología cuyo uso se empieza a explorar no solo para la manipulación de la narrativa, sino que también se plantean técnicas más allá como infundir directamente miedo, sin narrativa, alterar el estado de ánimo o incluso modificar la reacción.

[19] Las tecnologías inicialmente consideradas en HYFUTEC son: 5G; fabricación aditiva; inteligencia artificial; sistemas autónomos; biotecnología; computación en la nube; redes de comunicación; guerra cibernética y electrónica; libro mayor distribuido; energía dirigida; realidad extendida; hipersónicos; internet de las cosas; microelectrónica; nanomateriales; modernización nuclear; ciencias cuánticas; activos espaciales; y sensores ubicuos.

[20] El informe de 2021 analizó su uso en Yemen, Libia o Nagorno-Karabakh, pero no todavía en Ucrania.

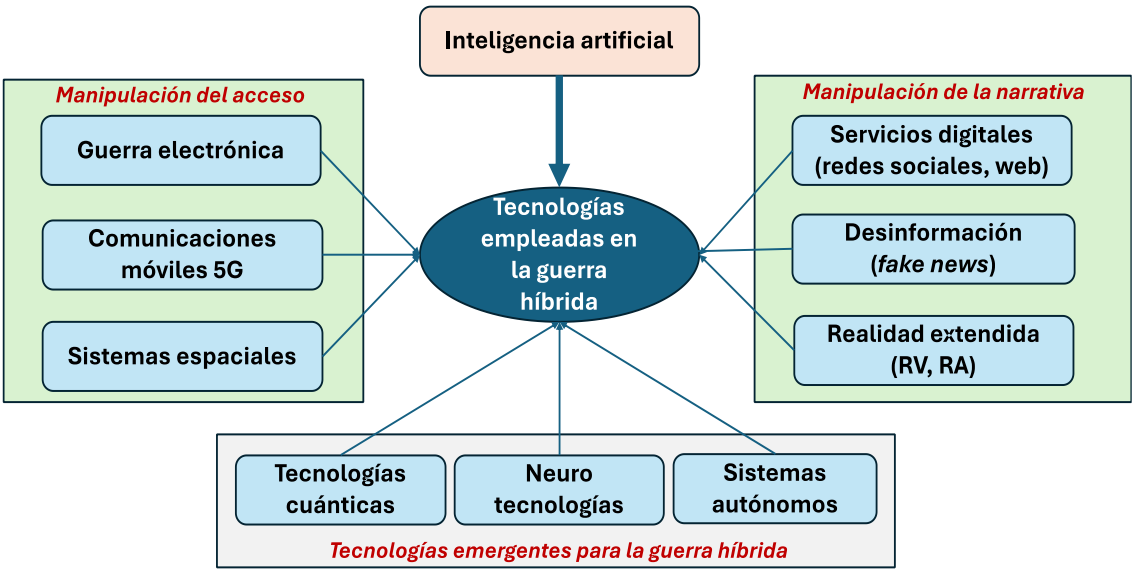


Figura 5. Relación entre las tecnologías implicadas en la guerra híbrida. Fuente: elaboración propia

Para lograr sus fines no basta con disponer de sistemas que integren la IA y tecnologías relacionadas; desde un **punto de vista operativo militar**, es decir, referido a la forma de llevar a cabo contrarrestar acciones híbridas desde la perspectiva de su uso en operaciones militares, es necesario que se incorporen a las tácticas y estrategias militares desde una perspectiva más amplia. Para ello, debe disponerse de la capacidad de **combinar la acción** (no necesariamente armada en un conflicto híbrido) **en el máximo número de los cinco dominios militares convencionales** (espacio, cibernético, marítimo, aéreo y terrestre) dentro de campañas híbridas más amplias que involucran dimensiones político-económicas, socioculturales e informativas. Se trata, por tanto, de una **acción que se ejecuta en un contexto multi-dominio**. Más tarde, se analizará la emergencia paulatina de un sexto **dominio sociocognitivo**, que se asocia más directamente a la guerra cognitiva[21].

En la figura 6 se han integrado todos los elementos indicados en las páginas anteriores. Obsérvese la posición de los **dos umbrales de detección y de atribución** señalados previamente, que van a variar de posición en el transcurso del tiempo. Un elemento relevante es la visibilidad que estas acciones tienen en la población no combatiente. Debe tenerse en cuenta que, sin frentes geográficos de batalla definidos, **toda la sociedad está en el frente**.

La figura 6 asume como eje evolutivo un **proceso gradual de escalada** desde situaciones de paz, aunque ello no implique que no exista confrontación ideológica y económica, a una situación de **amenazas híbridas** que se suma a las confrontaciones anteriores, no necesariamente con atribución del origen; de ésta al **conflicto híbrido**, sin olvidar las amenazas híbridas y la lucha ideológica y económica; y, finalmente, a la **guerra híbrida**, fase en la que a las actuaciones anteriores se suma un enfrentamiento armado militar concentrado en el tiempo y el espacio.

[21] Este sexto dominio no está todavía firmemente asentado. En algunos casos se le conoce como dominio cognitivo y, en otros, se le asocia al dominio de ciberdefensa.

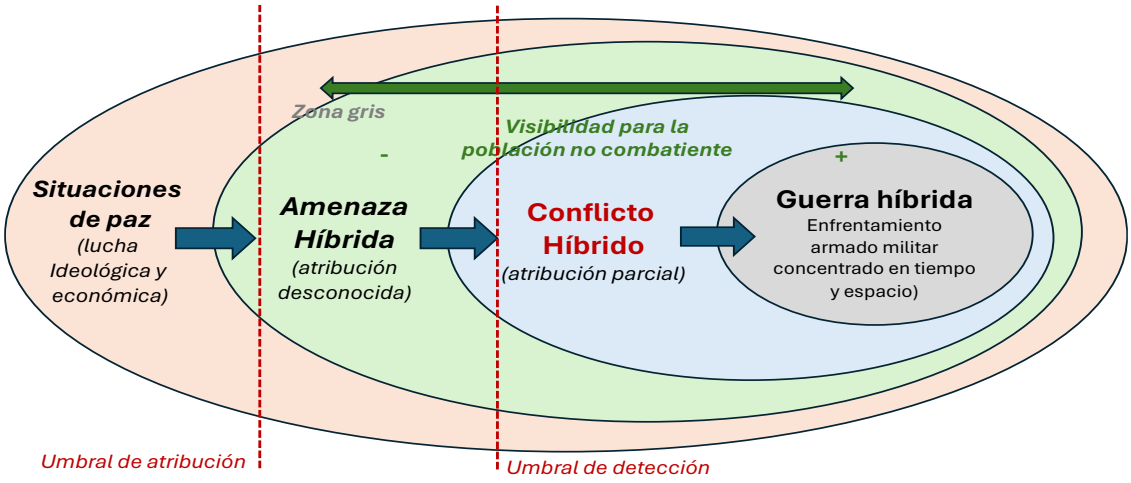


Figura 6. Relación entre amenazas, conflictos y guerra híbrida. Fuente: elaboración propia

Este proceso puede evolucionar en el tiempo, como se puede ver en el caso de la guerra en Ucrania, en dos fases diferenciadas en las que puede hablarse de conflicto militar (véase figura 7): **antes de la invasión** de Rusia (desde 2014 a 2022) y **después de la invasión** (desde 2022 a 2025).

En la **primera fase (2014-2022)**, salvo enfrentamientos esporádicos, la situación de guerra híbrida estaba limitada a algunas zonas ruso-parlantes del Donbas[22] reclamadas desde la independencia de Ucrania, pero el conflicto híbrido ya era visible en áreas de Ucrania y países limítrofes. Por otro lado, la amenaza híbrida se extiende a países que apoyan a Ucrania, aunque sin que Rusia se implique militarmente ni tampoco asuma la atribución de múltiples acciones.

Para muchos países europeos, la situación en esa fase se percibía como de "paz"; es verdad que compatible desde el punto de vista político con el apoyo a Ucrania y con la aprobación de sanciones por la UE (débiles y poco visibles para la mayoría de la población europea) y un inicio de confrontación geopolítica en el seno de las democracias occidentales sobre cómo responder de la mejor manera.

Ciertamente, en esa primera fase existen **acciones típicas de guerra híbrida**, con niveles de detección y de atribución muy cercanos entre sí, que no se consideraron suficientemente graves y que, probablemente, alentaron a Rusia a pasar a una segunda fase. Es ilustrativo en este contexto, analizar las respuestas políticas generadas en la fase 1 en la UE sobre el **gasoducto entre Rusia y Alemania denominado *Nord Stream,*** cuya construcción hay que enmarcarla en un contexto de "acercamiento" a Rusia; entre otras cosas para acceder a gas ruso barato. En 2013 comenzó la planificación de *Nord Stream 2*, dos tuberías de 1.250 km de longitud, de recorrido paralelo a las ya existentes entre Rusia y Alemania[23].

La siguiente frase pronunciada por el *Gen. Valery Gerasimov*, jefe de las fuerzas armadas de Rusia en 2013 es casi premonitoria: "*En el siglo XXI hemos visto una tendencia a difuminar las líneas entre los estados de guerra y de paz. Las guerras ya no se declaran y, una vez comenzadas,*

[22] La situación en Crimea no se saldó con enfrentamientos militares destacables.

[23] En 2010, comenzó la construcción del Nord Stream 1. El gasoducto, de doble tubería, con 1.224 kilómetros de longitud cada una, conecta Víborg, en Rusia con Lubmin, en Mecklemburgo-Pomerania Occidental.

proceden de acuerdo con un modelo desconocido"[24]. Un año después, Rusia ocupó Crimea y el Donbas con miles de "*hombrecitos verdes*" (personal militar sin distintivos, pero con armamento ruso y en camiones con placas de matrícula rusas)[25] provocando lo que, a la vista de lo sucedido posteriormente, fue una **tibia reacción de la UE**, deseosa de no romper los lazos con Rusia (sobre todo, los energéticos), que alimentó los siguientes pasos en 2022, esperando una reacción similar.

En 2015 se firmaron los primeros contratos del *Nord Stream 2* con *Gazprom* y varios proveedores de energía europeos. Apenas un año antes, Rusia se había anexionado la península ucraniana de Crimea. Para Alemania, esto no fue motivo para detener el proyecto. En septiembre de 2022, seis meses después del inicio de la guerra de Ucrania, "desconocidos", probablemente con apoyo de Ucrania (Hasselbach, 2025), hicieron explotar el gasoducto *Nord Stream*.

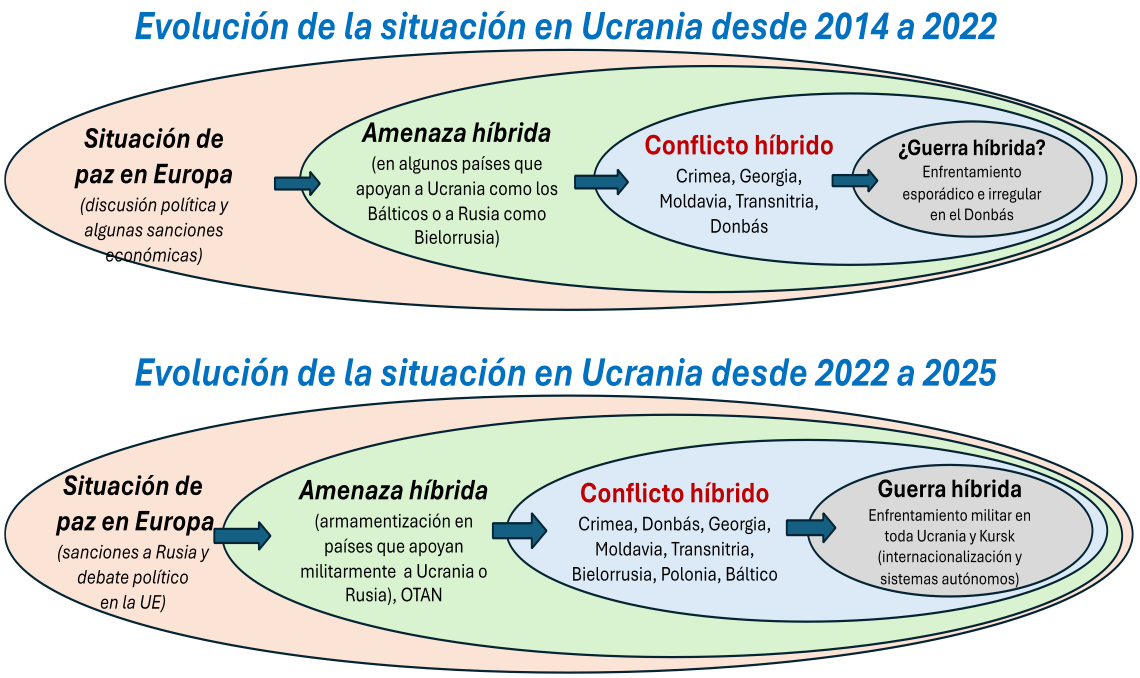


Figura 7. Evolución temporal de la situación en Ucrania. Fuente: elaboración propia

En la **segunda fase (2022-2025)**, esta situación da un vuelco significativo desde febrero de 2022, cuando estalla un **conflicto militar convencional** con invasión terrestre en cuatro regiones limítrofes con Rusia y ataques aéreos en toda Ucrania, que se **superpone al conflicto híbrido** limitado en la fase anterior a una parte del territorio de Ucrania (con diferentes niveles de

[24] https://www.economist.com/the-economist-explains/2022/02/22/what-is-hybrid-war-and-is-russia-waging-it-in-ukraine

[25] El término surgió por primera vez durante la adhesión de Crimea a Rusia en febrero de 2014, cuando tales fuerzas ocuparon y bloquearon el Aeropuerto Internacional de Simferópol y la mayoría de las bases militares de las Fuerzas Armadas de Ucrania en Crimea, como así también el Edificio del Consejo Supremo de Crimea en Simferópol https://www.bbc.com/news/world-europe-26379722 . En abril de 2014 una escuadrilla de "hombrecitos verdes" (soldados sin ninguna insignia nacional) tomó el control de una estación de policía en Sloviansk, una pequeña ciudad del óblast de Donetsk en el este de Ucrania. https://www.project-syndicate.org/commentary/putin-russia-failure-in-ukraine-by-carl-bildt-2019-04/spanish

intensidad). Posteriormente se extiende a Rusia (invasión de una parte de la región rusa de Kursk por Ucrania y ataques aéreos esporádicos que llegan hasta Moscú) como un elemento de reducción de presión en otras partes del frente y como una posible baza negociadora futura.

En esta segunda fase, el **conflicto híbrido supera el umbral de visibilidad para la población de los países europeos**[26] y, casi siempre, **el de atribución**, extendiéndose a Crimea, Georgia, Moldavia, Transnistria, Bielorrusia, Polonia o los países alrededor del Mar Báltico, con incremento masivo de ciberataques desde los dos bandos. Es revelador indicar la denominación dada por Rusia a esta invasión como una "**operación militar especial**" para desnazificar Ucrania. No se apela directamente a una confrontación militar abierta[27]. **Comienza la batalla por la narrativa.**

La **amenaza híbrida**, incluyendo la retórica de la amenaza nuclear, se extiende a otros países que apoyan militarmente a Ucrania y que, en el caso de la UE, aceleran su "*armamentización*" con fuertes incrementos de los presupuestos militares y con un incremento de las sanciones a Rusia y Bielorrusia.

Asimismo, el conflicto se internacionaliza militarmente con la presencia de "*voluntarios*" procedentes de otros países del Cáucaso y, en 2024, con la presencia de **unidades regulares de Corea del Norte** en los frentes de combate (al menos en *Kursk*, dentro de Rusia). El umbral de atribución baja claramente en el caso de Rusia, se adapta la intervención iniciada como "operación especial" a un conflicto militar, pero se desplaza o mantiene en países limítrofes con incremento de las acciones híbridas. Bielorrusia, cuya ayuda a Rusia es clara, intenta mantenerse fuera del conflicto abierto.

En los tres años y medio transcurridos en septiembre de 2025 desde el comienzo de la invasión, **todas las decisiones y actuaciones han estado condicionadas por el riesgo de escalada del conflicto.** Ello ha afectado expresamente a las condiciones establecidas para el uso de nuevos tipos de armas (p.ej. misiles de largo alcance) proporcionados por los países occidentales que apoyan a Ucrania, o el ámbito geográfico de su empleo.

Los **riesgos de escalada como consecuencia de los ataques híbridos** no son siempre fáciles de determinar *a priori*. Como se indica en la figura 8, adaptada de Cluzel (2017), un ataque militar de alto impacto en el dominio de la información puede conducir a un efecto de primer orden en los dominios políticos (desestabilización) y económicos (pérdidas empresariales y cambios bruscos en las valoraciones bursátiles) que **pueden conducir a una crisis sistémica**. El problema está en el uso del término "puede". ¿Ocurrirá? Y, en ese caso, ¿hasta qué punto se debe

[26] Puede argumentarse que el umbral de visibilidad estaba claramente superado desde 2014 (o, incluso antes) en las zonas rusoparlantes de Ucrania y, obviamente, en el Donbas y Crimea. Sin embargo, para la mayor parte de la ciudadanía de la UE alejada de la zona de conflicto no existía una consciencia clara de que existía una guerra híbrida. De hecho, la débil respuesta, en opinión de los autores, de la UE desde 2014 contribuyó precisamente a que se perpetuara un umbral de visibilidad elevado.

[27] Sputnik, medio de comunicación ligado al gobierno ruso, anunciaba: "*El 24 de febrero de 2022 el presidente de Rusia lanzó una operación militar especial en el territorio de Ucrania, luego de que las repúblicas de Donetsk y Lugansk le pidieran ayuda frente a la agresión de Kiev. El objetivo de la operación es "la desmilitarización y la desnazificación" de Ucrania".* https://noticiaslatam.lat/category_operacion-militar-especial-de-rusia-en-ucrania/

considerar que la exposición al riesgo[28] es suficientemente elevada como para actuar? ¿Debería hacerse de forma preventiva?

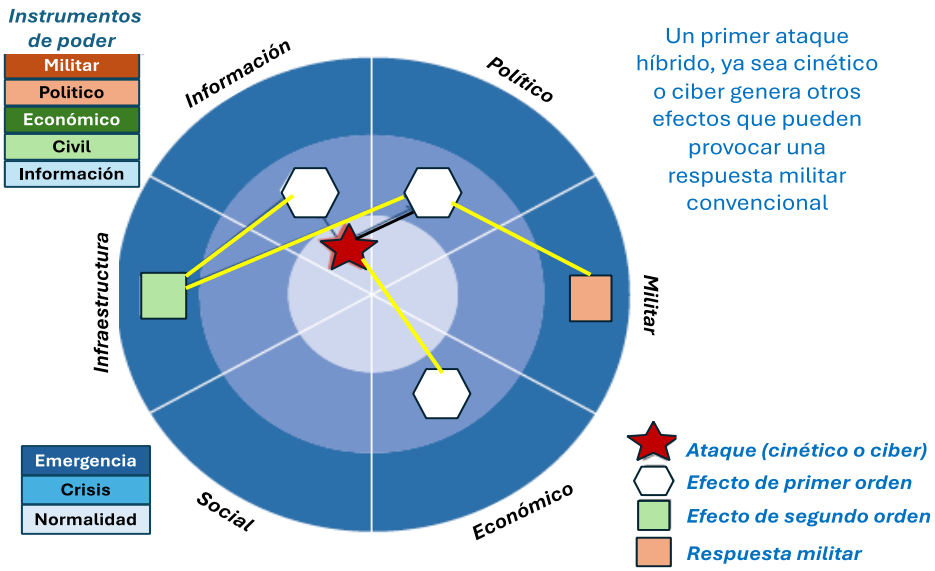


Figura 8. Efectos de escalada derivados de un ataque híbrido. Fuente: adaptada de Cullen (2017).

Esta situación también puede conducir a un **efecto de segundo orden en el dominio de la infraestructura de comunicaciones** (p.ej. con disrupción en servicios de internet si se atacasen cables submarinos, satélites de comunicaciones u observación, o centros de datos), o **energética** (p.ej. con ataques a los sistemas de control de las redes eléctricas o a las propias subestaciones o a las conducciones de gas) como está ocurriendo en Ucrania. La conversión de todas esas infraestructuras en objetivos militares obliga a pasar a un **estado de emergencia** con consecuencias directas sobre la población en un terreno en el que los inviernos son duros. Debe tenerse en cuenta que muchos sistemas satelitales son de aplicación dual, como se indica en la figura 9 con terminales de *Starlink* en Ucrania.

Relevancia geopolítica de Starlink en la guerra de Ucrania

¿Tiene sentido que un elemento clave en la lucha frente a la "*weaponization*" de las infraestructuras de telecomunicación dependa de la voluntad de una persona/empresa?

Figura 9. Uso dual de los servicios proporcionados por la constelación de satélites de Starlink. Fuente: Izquierda: uso civil (foto procedente de Financial Times). derecha: uso militar. Composición elaboración propia.

A su vez, y **como represalia**, el país atacado puede provocar la adopción de sanciones económicas e, incluso, una **respuesta militar empleando ataques híbridos similares o mediante**

[28] La exposición al riesgo de un evento suele medirse como el producto de la probabilidad de ocurrencia por su impacto.

el empleo de medios convencionales. Todo ello, con un inherente riesgo de escalada basada en el análisis con información parcial (datos históricos y de inteligencia) de las posibles reacciones del contrario[29].

El caso de la **guerra entre Rusia y Ucrania** ha sido analizado extensamente desde la perspectiva de guerra híbrida en el contexto de un conflicto militar de 5ª generación (Beznosiuk, 2025). Los esfuerzos de Rusia desde 2022 para desestabilizar y subyugar al resto de Ucrania han implicado una combinación de agresión militar convencional, sabotaje, ciberataques, campañas de desinformación y apoyo a actores prorrusos en Ucrania.

De hecho, **la actividad de Rusia en la guerra híbrida no se circunscribe a Ucrania** y atentados físicos (no incidentes cibernéticos) han proliferado en varios países. El *IISS (International Institute for Strategic Studies)* ha publicado un informe en agosto de 2025 (Edwards y Sedenstein, 2025) dando cuenta de casos de sabotajes en países europeos, en los que se ha podido verificar con alto nivel de confianza su origen ruso, como se indica en la figura 10, que indica la frecuencia y tipología de estos casos. Obsérvese que muchos de ellos están dirigidos contra infraestructuras críticas como comunicaciones, energía, transportes, agua o infraestructuras submarinas (p.ej. cables de comunicaciones o eléctricos)

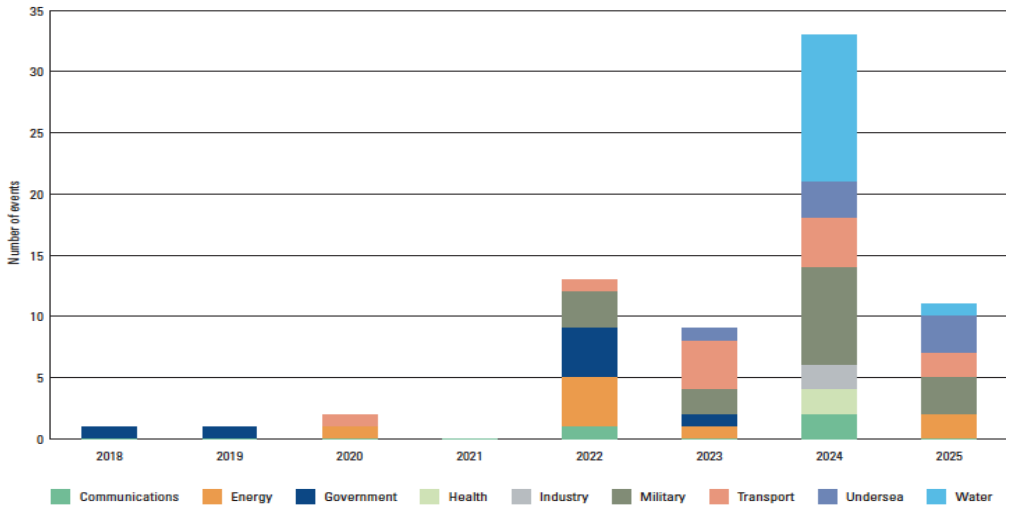


Figura 10. Frecuencia de sabotajes en Europa atribuidos a Rusia como parte de la guerra híbrida de enero de 2018 a junio de 2025. Fuente: Edwards y Sedenstein, 2025

Gracias a esta exposición prolongada a la guerra híbrida rusa, Ucrania ha podido desarrollar **contramedidas que han ayudado a desarrollar resiliencia y reducir el impacto de las operaciones híbridas de Rusia**. La respuesta de Ucrania se ha basado en un esfuerzo de colaboración que involucra al gobierno ucraniano, la sociedad civil y el sector privado. En la esfera cibernética, los esfuerzos para mejorar la seguridad digital de Ucrania han desempeñado un papel clave, con el establecimiento del *Ministerio de Transformación Digital*. Esta estructura permite al gobierno ucraniano **sincronizar posiciones** con el establecimiento de narrativas proactivas, al contrarrestar las campañas de desinformación de Rusia.

[29] Bruno Kahl, jefe del Servicio Federal de Inteligencia alemán (BND), indicó en noviembre de 2024 que las tácticas de guerra híbrida de Rusia contra Occidente podrían llevar a la OTAN a invocar la cláusula de defensa mutua de la Alianza (artículo 5). https://es.euronews.com/my-europe/2024/11/28/alemania-advierte-la-guerra-hibrida-de-rusia-puede-activar-la-clausula-de-defensa-de-la-ot

Ucrania también se ha beneficiado de un **enfoque descentralizado** que involucra a voluntarios digitales, sociedad civil y asociaciones público-privadas. Una amplia gama de grupos de tecnólogos civiles e investigadores de código abierto en Ucrania actúan detectando y contrarrestando la desinformación y narrativas rusas de manera eficiente, coordinando los mensajes entre el gobierno y la sociedad civil y manteniendo la coherencia durante las operaciones militares. Esta experiencia no puede quedarse reducida a Ucrania, sino que debe influir en la evolución de las estrategias y tácticas empleadas por los países miembros de la OTAN (Genini, 2025).

La relevancia e impacto potencial de estos ataques híbridos se ven acelerados por el **empleo de sistemas automatizados** con menor presencia del operador humano en su lanzamiento y gradación, debido al **uso de sistemas de IA capaces de analizar múltiples datos de contexto**, como se analiza en la siguiente sección, y actuar de acuerdo con unas reglas de respuesta predefinidas decididas e incorporadas a los algoritmos de IA. El uso de la IA en simulaciones de negociaciones diplomáticas parece que tiende también a acelerar la probabilidad de una escalada del conflicto (Rivera et al., 2024)[30].

El proceso de **escalada híbrida** ha venido acompañado de la creación de un discurso elaborado expresamente y mantenido en el tiempo, las denominadas ***"narrativas"*** como herramienta de apoyo a las posiciones de los bandos en conflicto, junto a un incremento y sofisticación de la desinformación con la generación de campañas específicas apoyadas por el uso de tecnologías como la IA.

Con el uso masivo de la IA se ha entrado en una explosión de la capacidad de la **"guerra cognitiva"** superpuesta a las actuaciones anteriores de naturaleza híbrida, cuyas características se abordarán en la siguiente sección.

1.3. Contexto de la guerra cognitiva

1.3.1. Conceptos y elementos básicos

La cita con la que se ha iniciado el presente documento: *"La OTAN se enfrenta a un nuevo tipo de amenaza: una guerra que no se libra con bombas y misiles, sino con mentiras y manipulación"*, extraída de un documento de la OTAN de julio de 2024[31] refleja un **cambio de paradigma potenciado por el uso de la IA** cuya relevancia es creciente.

Un concepto relacionado con el de guerra híbrida que es esencial para entender la evolución de los conflictos híbridos actuales es el de **"guerra cognitiva**", en el que las mentiras y la manipulación forman un sustrato esencial. Siguiendo la definición de la OTAN[32]: *la Guerra Cognitiva incluye actividades llevadas a cabo en sincronización con otros instrumentos de poder,*

[30] Según las conclusiones de un estudio publicado de enero de 2024 basado en simulaciones de conflicto empleando cinco grandes modelos de lenguaje, el uso de IA *"tiende a desarrollar una dinámica de carrera armamentística, que conduce a conflictos más importantes y, en casos contados, al despliegue de armas nucleares"*. https://www.swissinfo.ch/spa/la-ia-en-la-guerra%3A-un-avance-fulgurante-y-un-control-humano-dudoso/75522183

[31] https://www.act.nato.int/article/cogwar-concept/

[32] https://www.act.nato.int/activities/cognitive-warfare/

para afectar actitudes y comportamientos, influyendo, protegiendo o interrumpiendo la cognición a nivel individual, grupal o poblacional, para obtener una ventaja sobre un adversario. Diseñada para modificar las percepciones de la realidad, la manipulación de toda la sociedad se ha convertido en una nueva norma, y la cognición humana se ha convertido en un ámbito crítico de la guerra.

Desde nuestro punto de vista, la **guerra cognitiva** no debe tratarse como un concepto aislado de las actuaciones híbridas, sino de una **focalización y especialización del proceso de hibridación centrado en el manejo de la información con fines militares**, que contará con estrategias, recursos económicos y organizativos y, sobre todo, el uso de herramientas tecnológicas propias; entre ellas el uso preeminente de **herramientas de inteligencia artificial**, como se analizará posteriormente.

Aunque no exista una definición formal, la visión complementaria procedente de varios autores tomada conjuntamente permite hacerse una mejor idea del **concepto de guerra cognitiva**, como se indica seguidamente:

- *"Históricamente, el término "**guerra cognitiva**" se ha referido al uso de los medios de acción que un Estado o un grupo influyente fabrica para manipular los mecanismos espontáneos de la cognición de un enemigo o su pueblo, con el fin de debilitarlos, penetrarlos, influir o incluso someterlos o destruirlos"* (du Cluzel, 2021)
- *"La **guerra cognitiva** ya está con nosotros. El principal reto es que es esencialmente invisible; todo lo que ves es su impacto, y para entonces... a menudo es demasiado tarde"* (Claverie and du Cluzel, 2022).
- *"La **guerra cognitiva** busca cambiar lo que la gente piensa y su forma de actuar. Su objetivo es sembrar disonancia, instigar narrativas conflictivas, polarizar opiniones y radicalizar grupos. Esta se nutre de los avances tecnológicos y se enfoca en el entorno humano como objetivo de la manipulación cognitiva"* (Lupiáñez, 2024).
- *"La "**guerra cognitiva**" es una nueva forma de pensar sobre cómo los actores usan el poder, influyen en los paisajes geopolíticos y derriban a los oponentes. La guerra cognitiva tiene como objetivo manipular las percepciones y los procesos de toma de decisiones de los individuos y las sociedades explotando las vulnerabilidades inherentes a la psicología humana y al ecosistema de la información"* (Hagen, 2024).

Recientemente, Bebber y Marshal (2024) proponen una **definición de guerra cognitiva más elaborada**. Para estos autores se refiere a:

"El empleo de la ciencia y la tecnología que altera la cognición dentro de los individuos, grupos y poblaciones, lo que conduce a cambios en la comprensión, la emoción y el comportamiento. Su objetivo es incurrir en una influencia disruptiva de manera propagativa directa o indirecta alterando las sensaciones, percepciones, creencias, patrones de pensamiento, emociones y comportamientos resultantes de individuos y colectivos para desestabilizar y manipular direccionalmente el statu quo sociocultural, ecológico, económico, político y militar y, por lo tanto, permitir la influencia y el poder intencionales. Las características clave incluyen las aplicaciones de la comprensión avanzada y los métodos de las ciencias del cerebro, la dependencia de los datos y las ciencias y tecnologías computacionales, el uso del espectro electromagnético y las redes sociales impulsadas a diferentes velocidades y escalas para dirigirse a agentes, actores, grupos y poblaciones clave que, a su vez, pueden ejercer comportamientos que amplifican los efectos disruptivos en escalas y direcciones particulares".

A partir de esa definición la figura 11 tomada de los mismos autores (Bebber y Mashal, 2024) describe esquemáticamente los elementos de una **acción de guerra cognitiva** en el que el empleo de **la tecnología juega un papel predominante** (a la izquierda de la figura se indican algunas tecnologías cuyo objeto es alterar la cognición de individuos, grupos y poblaciones). Obsérvese que el objetivo es alterar (desestabilizar y manipular) el *statu quo* en ámbitos como el sociocultural, el ecológico, el económico, el político y el militar.

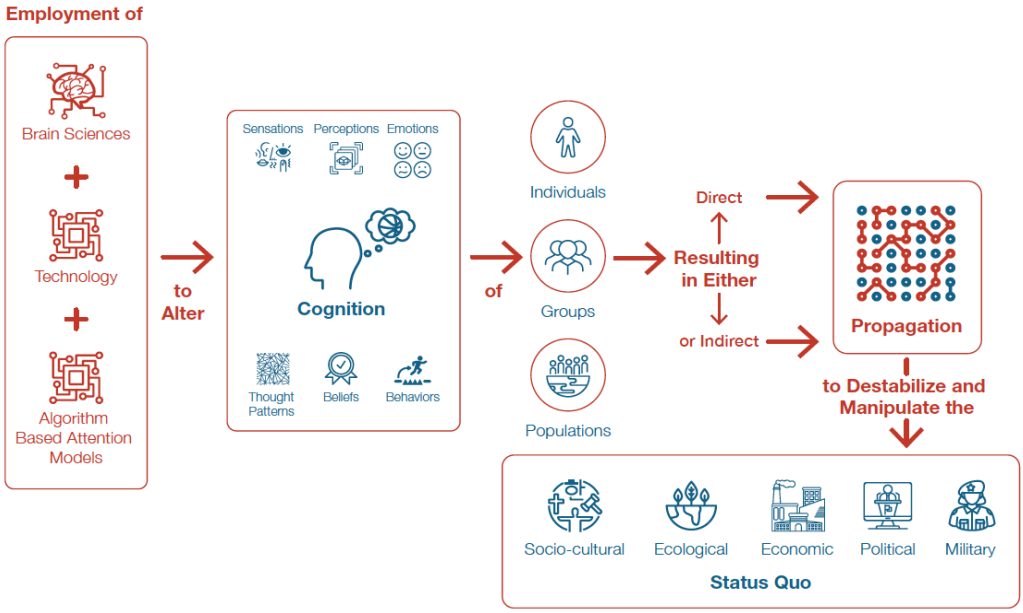


Figura 11. Elementos para una campaña de guerra cognitiva. Fuente: Bebber y Marshal (2024)

Algunos autores como Hung & Hung (2022), desde un enfoque de la guerra cognitiva modulado por la visión asiática, argumentan que no se trata de un concepto independiente, sino **subordinado al de guerra híbrida** combinada con otras formas tradicionales de guerra no cinética, como la **guerra de información** y la **guerra cibernética.**

Es interesante en este contexto referirse al marco de referencia propuesto por Nikola y McMahon (2024) (véase figura 12) cuyo objetivo es **enmarcar la guerra cognitiva**, orientada al control del cerebro de las personas, **como un subconjunto de la guerra híbrida**, incluyendo como parte de ella la **"guerra económica",** y considerar la "**guerra informativa**" (incluyendo la desinformación y control de los medios de comunicación) como un subconjunto de la guerra cognitiva.

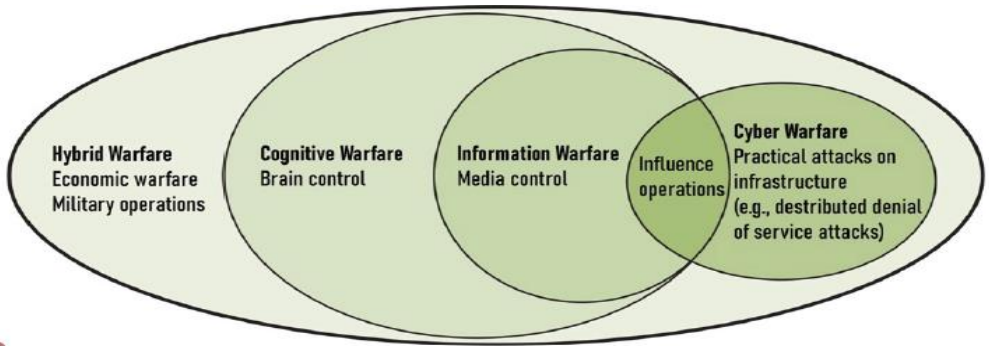


Figura 12. Relación entre guerra híbrida, cognitiva, informativa y cibernética. Fuente: Nikola y McMahon, 2024.

En la figura 12 se incluye también la **guerra cibernética** (*cyber warfare*) dentro de la guerra hibrida, pero solo parcialmente relacionada con la guerra cognitiva y, por tanto, con la informativa a través de las denominadas "*operaciones de influencia*", encubiertas o no. Se ha preferido dejar al margen del presente documento la guerra cibernética, cuyo ámbito técnico y del uso de la IA es diferente, para centrarse en el de guerra cognitiva (incluyendo la guerra informativa basada en la desinformación).

Con objeto de disponer de un **marco de referencia para el análisis**, la figura 13, desarrollada a partir de la figura 6 anterior, concibe la **guerra cognitiva** como una estrategia esencial cuyo objetivo es el **dominio de la información/desinformación** y la construcción y difusión de **narrativas orientadas a lograr objetivos políticos**. Desde esta perspectiva, surge como un elemento transversal en todos los conflictos, incluyendo los militares, y se expande a las situaciones denominadas "*de paz*".

Aunque el origen de la guerra cognitiva puede rastrearse muchos años atrás en la historia[33], realmente, no habría sido posible alcanzar la situación actual en la guerra cognitiva sin la **aceleración del proceso de digitalización**, fundamentalmente derivado del uso masivo de las **redes de banda ancha** y del procesamiento de **grandes volúmenes de datos** (*big data*) obtenidos tanto del campo de batalla como de la retaguardia. A ello se une la utilización preferente como "arma" de las **redes sociales**, y del uso de herramientas para el lanzamiento (automatizado) de sofisticados **ciberataques**.

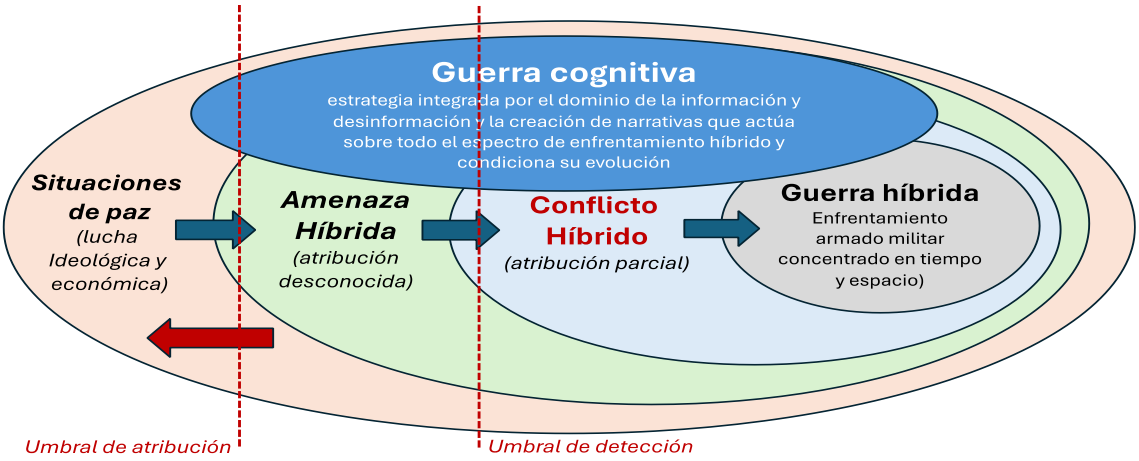


Figura 13. Interacción entre amenazas, conflictos y guerra híbrida, y la guerra cognitiva. Fuente: elaboración propia

Desde el punto de vista tecnológico, tampoco la **guerra cognitiva** hubiera podido alcanzar un impacto como el que está teniendo sin el concurso de la **inteligencia artificial**, puesto que todas las acciones asociadas a las tecnologías avanzadas mencionadas **dependen de la IA para ser**

[33] En 1950, durante la Guerra de Corea, comenzó a extenderse un rumor entre las tropas chinas que luchaban contra los estadounidenses: sus propios líderes los estaban traicionando. La moral se desplomó. Algunos desertaron, otros se entregaron sin disparar un solo tiro. Lo que parecía ser una crisis espontánea era en realidad una operación meticulosamente planificada. Mensajes filtrados, transmisiones de radio falsas y una red de desinformación habían sembrado la duda en la mente del enemigo. Esta fue una de las primeras demostraciones modernas del poder de la Guerra Cognitiva. https://grupogoberna.com/guerra-cognitiva-y-psyops-conceptos-fantasma/

efectivas como armas en la guerra cognitiva. En el próximo futuro, otras tecnologías, en especial las **tecnologías cuánticas** y la **neurotecnología**, se unirán a ella.

Durante la **Conferencia de Seguridad de Múnich** de febrero de 2025 el tema del uso de las "*tecnologías profundas*" estuvo presente en la mesa de discusión en palabras del presidente de la sesión, M. Muñiz: *"Estamos siendo testigos de cómo la Inteligencia Artificial, la Computación Cuántica, los Materiales Avanzados y la Biotecnología están redefiniendo las capacidades militares y las formas en que respondemos a las amenazas emergentes. A medida que estas tecnologías avanzan, las potencias mundiales están acelerando sus esfuerzos para liderar la supremacía tecnológica, lo que plantea cuestiones críticas sobre la seguridad, la soberanía y la competitividad económica, especialmente para Europa"*[34].

1.3.2. Uso militar por Rusia del dominio cognitivo

Todas las grandes potencias se han embarcado en la guerra cognitiva. Por su importancia para la UE en su apoyo a Ucrania, prestaremos espacial atención a Rusia. **Rusia** ha prestado creciente atención a la guerra cognitiva tomando como base la **consideración de las redes sociales como un arma** que debe emplearse como elemento clave para la difusión de **mensajes englobados en narrativas a largo plazo** (López-Garay, 2025).

No se trata de un hecho nuevo, ya existía con la *Unión Soviética* en todo el periodo de la *Guerra Fría*, pero su relevancia al calor del incremento de la tensión geopolítica entre grandes potencias es mayor y ha generado preocupación en la Unión Europea con repetidas declaraciones y medidas adoptadas por las instituciones comunitarias.

Como ejemplo de actuación rusa, semanas antes de las elecciones alemanas celebradas el 23 de febrero de 2025, Rusia reactivó sus tácticas de desinformación iniciadas en 2022 para impulsar al partido ultraderechista *Alternativa para Alemania (AfD)* y desacreditar a sus rivales. Se detectaron cientos de mensajes difundiendo afirmaciones engañosas sobre las elecciones en la red social X, con indicios de que pertenecen a la famosa campaña rusa conocida como *'Doppelgänger'* (doble de una persona)[35].

En la figura 14 pueden verse algunas de las actuaciones recientes ligadas a **campañas de desinformación generadas por Rusia en Europa** (Georgia, Crimea, Donbas, Moldavia, Rumanía) o fuera de ella (Siria, Mali, Níger) y los actores y medios implicados: comunicaciones oficiales del gobierno, mensajes de entidades financiadas por el Estado, cultivo de fuentes de proxies, utilización de las redes sociales como arma, o desinformación potenciada por ciberataques.

[34] https://www.ie.edu/cgc/news-and-events/news/discussing-deep-tech-defense-at-the-2025-munich-security-conference/

[35] El CeMAS, organización alemana sin ánimo de lucro, afirma haber identificado 630 mensajes en X en alemán con "patrones típicos de '*Doppelgänger*'" entre mediados de diciembre de 2024 y mediados de enero de 2025. Muchos incluían enlaces a sitios web falsificados que pretendían ser medios de comunicación alemanes como 'Der Spiegel' y el canal de televisión 'Welt'. https://es.euronews.com/my-europe/2025/01/22/las-campanas-de-desinformacion-rusas-resurgen-antes-de-las-elecciones-alemanas

Figura 14. Modelo de guerra cognitiva algorítmica de Rusia. Fuente: https://www.state.gov/russias-pillars-of-disinformation-and-propaganda-report/

La **estrategia de desinformación de Rusia** es bastante sofisticada: desde el diseño de una narrativa apropiada para cada objetivo, al uso de medios de comunicación controlados y convertidos en armas, pero creíbles[36], con objetivos temporales preidentificados y medición de la reacción. El proceso seguido se apoya en la **rápida difusión de los mensajes en la población objetivo**, que se convierte en el condicionante básico para el éxito, tal y como se describe esquemáticamente en la figura 15. El desarrollo de la **tecnología de redes sociales apoyada por la IA** permite segmentar la población objetivo, adaptando el mensaje a cada uno de los grupos para incrementar su eficacia. Así, el uso de redes sociales y plataformas de mensajería (p.ej. *WhatsApp* o *Telegram*) constituye la base de la actuación.

[36] "*New Eastern Outlook*" es una publicación pseudo-académica del Instituto de Estudios Orientales de la Academia Rusa de Ciencias que promueve la desinformación y la propaganda centrada principalmente en Oriente Medio, Asia y África. Combina los puntos de vista pro-Kremlin de los académicos rusos con los puntos de vista contrarios a países occidentales y favorables a los teóricos de la conspiración.

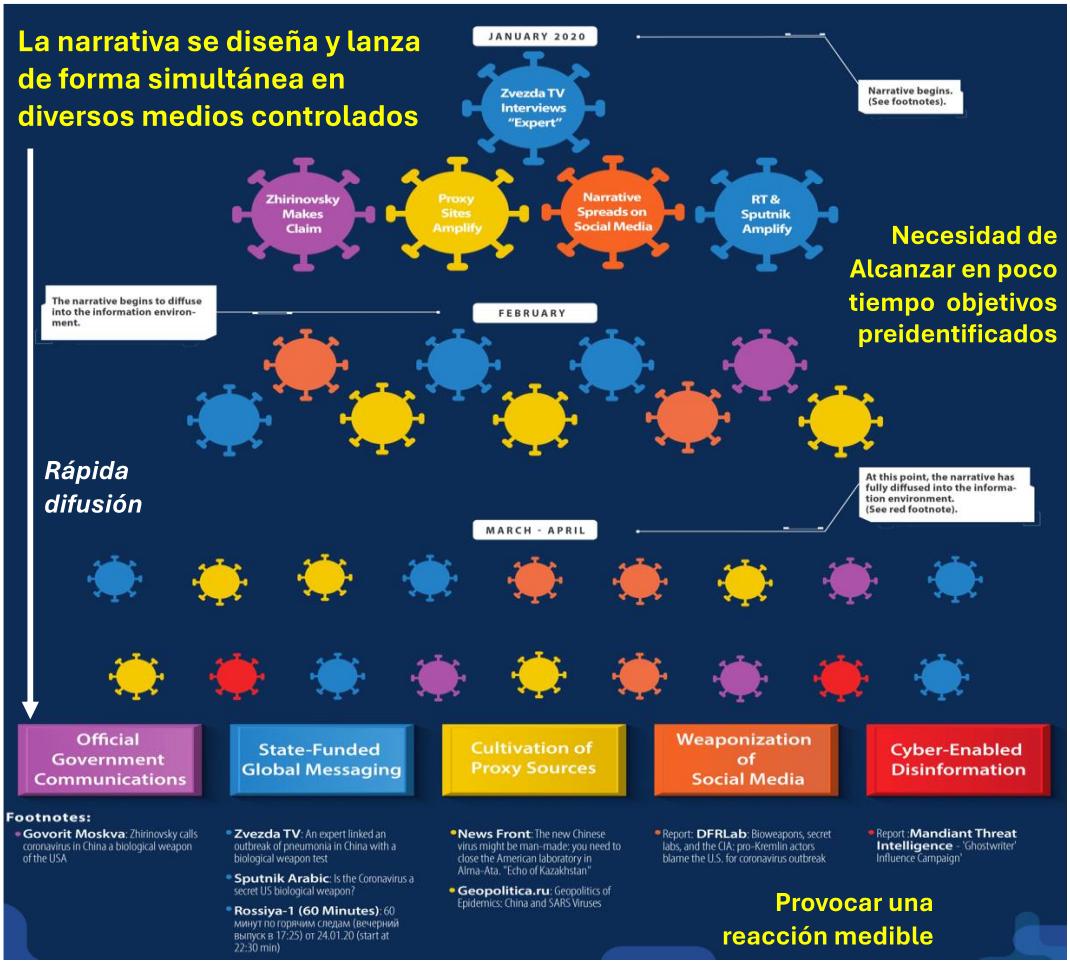


Figura 15. Estrategia de desinformación de Rusia. Fuente: adaptada de https://www.state.gov/russias-pillars-of-disinformation-and-propaganda-report/

Técnicamente, la forma de operar consiste en **clonar los sitios web de medios de comunicación de renombre, comprando dominios de aspecto similar e imitando su diseño y estilo.** Los artículos falsificados son luego amplificados por cuentas similares a bots en redes sociales. Esto permite que su propaganda y desinformación parezcan proceder de fuentes legítimas.

Volviendo a la situación de Ucrania, también puede analizarse la **evolución de la componente de guerra cognitiva en el conflicto** con la participación preeminente de Rusia y otras potencias. Inicialmente impulsada por Rusia, sus actuaciones se centraban en campañas de desinformación y construcción de narrativas focalizadas en Ucrania para modular la respuesta de la opinión pública (p.ej. centradas sobre el origen histórico e identidad rusa de Ucrania, el ataque sistemático de Ucrania sobre la población rusa de la zona, su comportamiento nazi y la necesidad de *"desnazificación"*, o la culpabilidad de Occidente en la generación del conflicto[37]). Sobre todo, esta actuación se llevaba a cabo en las zonas ruso-parlantes, pero también se extendió al conjunto de Ucrania y a otros países de la UE.

[37] Por ejemplo, en el ataque sistemático por el gobierno de Ucrania a la población ruso parlante y a la autonomía de regiones fronterizas, el no respeto a los acuerdos de Minsk, y el intento de extender la OTAN hasta la frontera de Rusia, integrando a Ucrania en ella.

El fenómeno no se limita a Ucrania. Un informe del servicio de contrainteligencia militar de Polonia, tras revisar más de 2.000 documentos clasificados y no clasificados desde 2004 hasta 2024, indica que Rusia ha estado difundiendo ***desinformación*** en Polonia como parte de una "**guerra cognitiva a largo plazo para sembrar división en la sociedad**"[38].

Vladimir Putin, presidente de Rusia, asegura repetidamente que el gobierno de Ucrania es "*abiertamente neonazi*", "*pronazi*" y está controlado por "*pequeños nazis*". La imagen de la figura 16, con vallas publicitarias con citas del presidente de Rusia en Simferopol, Crimea, en marzo de 2022, refleja el objetivo perseguido. En la de la derecha se lee: "*Queremos la desmilitarización y desnazificación de Ucrania*" (Troianovski, 2022)[39].

Figura 16. Campaña de Rusia en Crimea contra Ucrania en marzo de 2022. Fuente: https://www.nytimes.com/es/2022/04/17/espanol/nazi-ucrania-putin.html

Si bien no puede considerarse el uso de *vallas publicitarias* como una "tecnología avanzada de guerra cognitiva", sí se han usado muchos otros medios ligados a la batalla informativa y cibernética. De todas formas, aunque el uso de las vallas como medio de difusión no supone una técnica avanzada, la generación del mensaje y de la narrativa durante tiempo antes sí puede estar basada en el análisis de múltiples datos. El hecho de que el vehículo final no sea una tecnología avanzada no implica que esta no esté implicada en fases previas. De hecho, es común la **utilización de múltiples plataformas** para cada una de las diferentes funciones del ciclo de desinformación.

Desde el inicio de la guerra en Ucrania, los conflictos armados han estado marcados por la involucración de ***grupos hacktivistas*** que, motivados por su ideología, han defendido los intereses de uno de los bandos atacando a objetivos del bando opuesto. La manera de proceder

[38] https://www.huffingtonpost.es/global/un-informe-destapa-guerra-cientifica-putin-iniciado-vecino-ucrania.html

[39] Con el uso del término "nazi" Rusia utiliza el trauma de la Segunda Guerra Mundial en la población rusa para justificar su invasión a Ucrania. La guerra en Ucrania se presenta como una continuación de la lucha de Rusia contra el mal, en lo que se conoce en el país como la *"Gran Guerra Patriótica*". Hay que indicar también que la denominación de "nazi" se la atribuyen los propios grupos extremistas con sus objetivos, su ideología y, muy importante, su imagen y uso de símbolos relacionados con el nazismo

varía entre los distintos grupos aunque, en términos estadísticos, el método preferido de ataque ha sido la **denegación de servicio distribuida (DDoS).**

Su importancia se manifiesta en que los grupos de desinformación prorrusos superan en número a los proucranianos, y parece que poseen mejores capacidades de coordinación[40]. Cabe destacar el grupo *hacktivista* prorruso ***"NoName057"*** responsable de cientos de ataques de denegación de servicio (*DDoS*) contra objetivos en países críticos con la invasión rusa de Ucrania[41]. Su fuerza radica en la participación de un número elevado de personas físicas simpatizantes. ***NoName057*** también destaca por el desarrollo de una plataforma específica de IA denominada ***DDoSIA***[42], utilizada por los integrantes del grupo para lanzar sus ataques de denegación de servicio.

Las actuaciones de Rusia no se limitan a Ucrania y al lanzamiento de ciberataques. Sus actuaciones, junto con otros ataques de guerra híbrida, se extienden a otros países con nombres de ***"hackers"*** **identificados y buscados internacionalmente** (véase figura 17) dirigidos por la inteligencia militar de Rusia (*Russian military intelligence, GRU*) (Jones, 2025).

Figura 17. Ciberataques de Rusia. Fuente: Jones (2025), https://www.csis.org/analysis/russias-shadow-war-against-west

Según un reciente estudio de marzo de 2025 realizado por del CSIS, **Rusia está involucrada en una agresiva campaña de subversión y sabotaje contra objetivos europeos y estadounidenses** que complementan la guerra convencional en Ucrania. El número de ataques rusos en Europa casi se triplicó entre 2023 y 2024, después de cuadruplicarse entre 2022 y 2023 (Jones 2025)[43].

[40] *Telegram* es el canal preferido de los grupos hacktivistas rusos para coordinar ataques.

[41] En julio de 2025, Europa y Estados Unidos unieron fuerzas en la *Operación Eastwood* para desmantelar a NoName057(16) como parte de un operativo internacional. Bloquearon más de 100 servidores y arrestaron a parte del personal. https://www.cibersecurity.io/noname05716-operacion-eastwood-asesta-un-duro-golpe-al-cibercrimen-pro-ruso/

[42] https://www.sekoia.io/en/glossary/ddosia-project/

[43] Aproximadamente, el 27% de los ataques se dirigieron contra objetivos de transporte (como trenes, vehículos y aviones), otro 27% fueron contra objetivos gubernamentales (como bases militares y funcionarios), el 21% fueron contra objetivos de infraestructura crítica (como oleoductos, cables submarinos de fibra óptica y la red eléctrica) y el 21% fueron contra la industria (como empresas de defensa).

Para llevar a cabo estos ataques, Rusia utiliza una variedad de armas y tácticas. Los más comunes involucraron artefactos explosivos e incendiarios (35%), anclas utilizadas para cortar cables submarinos de fibra óptica (27%); ataque electrónico (15%); y la militarización de los inmigrantes ilegales (8%). En la figura 18 puede verse uno de los buques de Rusia en el Mar Báltico (el petrolero *Eagle S*) acusado de sabotaje a cables submarinos, que forma parte de la denominada "flota oscura" de Rusia.

Figura 18. Petrolero ruso Eagle S. Fuente: https://www.csis.org/analysis/russias-shadow-war-against-west

Un elemento clave para valorar el uso de la guerra cognitiva es conocer el **impacto real** que este tipo de campañas tiene en Occidente, tanto desde el punto de vista **económico**, relativamente sencillo de estimar, como también, más difícil de estimar, en el **efecto de "desánimo**" o reducción de la moral de combate **de la sociedad** (militares y civiles) tras un uso prolongado de este tipo de ataques.

Los líderes militares conocen bien que **el valor de combate** de sus tropas debe verse como una función de su **capacidad** (medios técnicos/armas) para llevar a cabo una misión **multiplicada por su motivación al cuadrado** (Marahrens and Schröfl, 2024). La guerra cognitiva puede degradar el primero de los factores, la capacidad, pero afecta, sobre todo, al segundo, la **motivación**. Como dice Bryc (2024) *"Rusia no tiene que atacar militarmente a Occidente utilizando su "poder duro"; Un medio mucho más eficaz de derrotarlo podría ser destruirlo desde dentro con su "poder blando". Todo lo que Moscú necesita hacer es atacar metódica y consistentemente los pilares democráticos de Occidente utilizando una guerra cognitiva barata y altamente efectiva".*

Esta idea se ha planteado de una manera más cruda en las palabras pronunciadas en noviembre de 2022 por *Yevgeny Prigozhin*, entonces jefe del *Grupo Wagner* y de una fábrica de *Trols* cerca de San Petersburgo, en las que confirmó que **Rusia había interferido en las elecciones en Occidente**: *"Interferimos, estamos interfiriendo y seguiremos interfiriendo. Cuidadosamente, minuciosamente, quirúrgicamente y a nuestra manera. Sabemos cómo hacerlo"* (Bryc, 2024). Una clara declaración de intenciones.

Si para Rusia la guerra cognitiva no es una operación más de desinformación o de disrupción del ciberespacio, sino *"operaciones subversivas planificadas y llevadas a cabo por el ejército y los*

servicios de inteligencia rusos para destruir Occidente desde dentro",[44] no bastará con mejoras tecnológicas en la identificación de las operaciones de guerra cognitiva, **es necesario disponer de "tecnología" para contrarrestar con decisión la agresión híbrida y tener éxito**.

España pasa a ser también un objetivo derivado de su apoyo militar a Ucrania. Como indica el último informe publicado en octubre de 2024 por el *Centro Criptológico Nacional (CCN)* referido a 2023[45], ***NoName057***, en su grupo de *Telegram*, anuncia ataques de denegación de servicio contra objetivos españoles por el envío de armas a Ucrania o como muestra de apoyo a unas protestas organizadas por bomberos españoles (véase la figura 19). Antes del segundo ataque de Ucrania, ya se había atribuido a Rusia la campaña de afectar a las elecciones de Cataluña, y en relación con el proceso del 1-O. En aquel entonces, todavía no se había posicionado de forma clara España, por no haber criticado de forma abierta lo ocurrido en Crimea.

Figura 19. España como objetivo de la guerra híbrida de Rusia. Fuente: CCN, 2024

A pesar de que diversos informes resaltan la importancia e impacto de las campañas de desinformación, la valoración subjetiva es que **el impacto e influencia de estas campañas han sido muy pequeños en la población española**, como también ha sucedido en otros países en diversas ocasiones[46]. Nada es casual: lograr limitar su impacto forma parte del trabajo de las unidades de defensa contra la desinformación ligada a la interferencia extranjera existentes tanto en la UE como en España, incrementado porque el nivel tecnológico empleado hasta la fecha es relativamente bajo.

[44] https://www.kew.org.pl/en/2024/11/27/destroy-from-within-russias-cognitive-warfare-on-eu-democracy/

[45] https://www.ccn-cert.cni.es/es/informes/informes-ccn-cert-publicos/7274-ccn-cert-ia-04-24-ciberamenazas-y-tendencias-edicion-2024/file.html

[46] Los mensajes que resaltan el gran efecto obtenido forman parte de la campaña de desinformación de quienes las desarrollan y, por tanto, deben analizarse cuidadosamente para evitar caer en una trampa. https://www.newyorker.com/news/daily-comment/avoiding-the-disinformation-trap

Podría ser que esta situación de relativa influencia sea más relevante cuando la **tecnología implicada para el diseño de estos ataques alcance un nivel mucho mayor de sofisticación con el uso de sistemas de IA** (p.ej. con la generación automática de mensajes multimedia personalizados con bots empáticos y su difusión masiva), como se verá seguidamente.

Desde esta perspectiva, la guerra cognitiva ha alcanzado en marzo de 2025 un nivel superior con los acontecimientos derivados del **cambio de posición de Estados Unidos y su acercamiento a las posiciones de Rusia culpabilizando a Ucrania** (León, 2025). Las consecuencias derivadas, como la de suspender la ayuda militar o el uso de información de inteligencia, son aún difíciles de evaluar, pero sí ha provocado una reacción de la UE, junto a otros países europeos, que deberá concretarse en los próximos meses.

La guerra cognitiva se extiende a las narrativas entre antiguos aliados que formaban parte del mismo bando, en el triángulo de Ucrania, Estados Unidos y la UE. Todos ellos, ante un cambio de contexto sobrevenido, deben **reconstruir sus narrativas.** En el caso de la UE**,** la narrativa a construir trata de convencer a su propia población de la **necesidad de actuar con determinación en apoyo de Ucrania, sumando recursos económicos** y sabiendo que hay poco tiempo para hacerlo si se desea que sea efectiva.

El 8 de octubre de 2024, la UE adoptó un **nuevo marco de sanciones contra los responsables de actividades desestabilizadoras contra la UE y sus Estados miembros**[47]. Este nuevo marco permite a la UE poner en su punto de mira a las personas y entidades que participan en acciones y medidas del Gobierno de la Federación de Rusia que menoscaban los valores fundamentales de la UE y sus Estados miembros, su seguridad, independencia e integridad, así como las de organizaciones internacionales y terceros países. Gracias al nuevo régimen de sanciones, **la UE puede hacer frente a diversas amenazas híbridas**, como:

- el menoscabo de los procesos electorales y del funcionamiento de las instituciones democráticas;
- las amenazas dirigidas contra actividades económicas, servicios de interés público o infraestructuras críticas y su sabotaje;
- el recurso coordinado a la desinformación, la manipulación de información e injerencia por parte de agentes extranjeros;
- las actividades informáticas malintencionadas;
- la instrumentalización de los migrantes.

1.3.3. Guerra cognitiva en China

Aunque se considera a **Rusia como el actor principal** en este "*nuevo espacio de batalla cognitivo*", otros Estados no democráticos como China, Irán o Corea del Norte han comenzado a probar tácticas similares contra sus adversarios, y se extiende a otros muchos países en la medida en la que acceden a herramientas más sofisticadas.

Esta sección aborda la situación en China bajo el término de "***guerra cognitiva algorítmica***". Su origen, sin embargo, deriva de otras doctrinas militares de China anteriores. La **doctrina de las «Tres Guerras»** (三种战法), desarrollada por China, constituye una estrategia integral de

[47] https://www.consilium.europa.eu/es/policies/sanctions-against-russia/

influencia y manipulación del entorno informativo[48]. Se compone de tres elementos fundamentales:[49]

- *Guerra Psicológica (心理战): Su objetivo principal es socavar la voluntad, cohesión y moral del adversario mediante el uso de propaganda, desinformación y técnicas de manipulación psicológica[50].*
- *Guerra Mediática (媒体战): Se orienta a controlar la narrativa y el flujo de información tanto a nivel nacional como internacional. Esto se logra a través del dominio de los medios de comunicación, la propaganda y estrategias de relaciones públicas.*
- *Guerra Legal (法律战): Se basa en la utilización y reinterpretación del marco jurídico internacional con el fin de legitimar las propias acciones y deslegitimar las del adversario.*

En la figura 20 puede verse de forma esquemática que el objetivo de la guerra cognitiva algorítmica es **obtener el máximo provecho de la combinación de algoritmos de análisis de datos y de recomendaciones en redes sociales** para influir eficazmente en el comportamiento del individuo. Obsérvese el papel que juega el uso de la segmentación de la información ("granularidad de los datos") para ejercer una influencia precisa e individualizada.

Figura 20. Componentes de la guerra cognitiva algorítmica de China. Fuente: https://scsp222.substack.com/p/algorithmic-cognitive-warfare-the

La clave de la **visión china de la guerra cognitiva algorítmica** se sustenta en el aprovechamiento intencionado de las enormes cantidades de datos disponibles (fácil en el contexto de la regulación de datos en China y con centenares de miles de cámaras), necesarios para entrenar algoritmos que puedan analizar el comportamiento de las personas y comprender sus

[48] El desarrollo doctrinal de este enfoque estratégico se formalizó en 2003 a través de una directiva de la Comisión Militar Central (CMC) del Partido Comunista Chino. Posteriormente, sus principios fueron incorporados en los manuales de la Academia de Ciencias Militares (AMS) y en diversos documentos internos del Ejército Popular de Liberación (EPL). https://grupogoberna.com/guerra-cognitiva-y-psyops-conceptos-fantasma/

[49] Como precedente se puede señalar el concepto de *"Guerra Irrestricta"* (en chino 超限战, que significa "guerra más allá de los límites"), escrito en 1999 por dos coroneles del Ejército Popular de Liberación chino, Qiao Liang y Wang Xiangsui. Plantea que la guerra ha cambiado de forma y ya no es solo un conflicto militar directo, sino que abarca todos los ámbitos del poder, incluyendo la economía, la tecnología, la ley internacional y la información. En aquel momento, identificaba métodos de combate para hacer frente a un adversario tecnológicamente más avanzado; ahora, probablemente, la asimetría sea menor, pero no la validez de los conceptos. https://legrandcontinent.eu/es/2023/06/03/la-guerra-irrestricta-un-aggiornamento/

[50] Dirigida al combatiente, daría cabida incluso a la neurotecnología una vez que madure y proporcione herramientas efectivas para su uso militar.

preferencias, estado mental y relaciones con otras personas de su entorno[51]. Dentro de este marco, los rasgos personales obtenidos no son solo datos estáticos, sino que se concibe como un **mosaico en constante cambio que se puede monitorizar y mapear frente a objetivos** a lo largo del tiempo.

El concepto de guerra cognitiva algorítmica se desarrolla en las actividades descritas en la figura 21. Todas las actividades indicadas en la figura están **potenciadas por un uso creciente de herramientas de análisis de IA y alimentadas por la ingente captura y clasificación de datos personales** que permite la legislación china. Desde el punto de vista de las necesidades de la defensa, se ve también favorecido por el concepto de **fusión civil-militar**[52], característico del entramado de uso dual de China, en el que las fronteras entre ambos ámbitos se superponen y se alinean con un metaobjetivo de estabilidad del sistema a largo plazo.

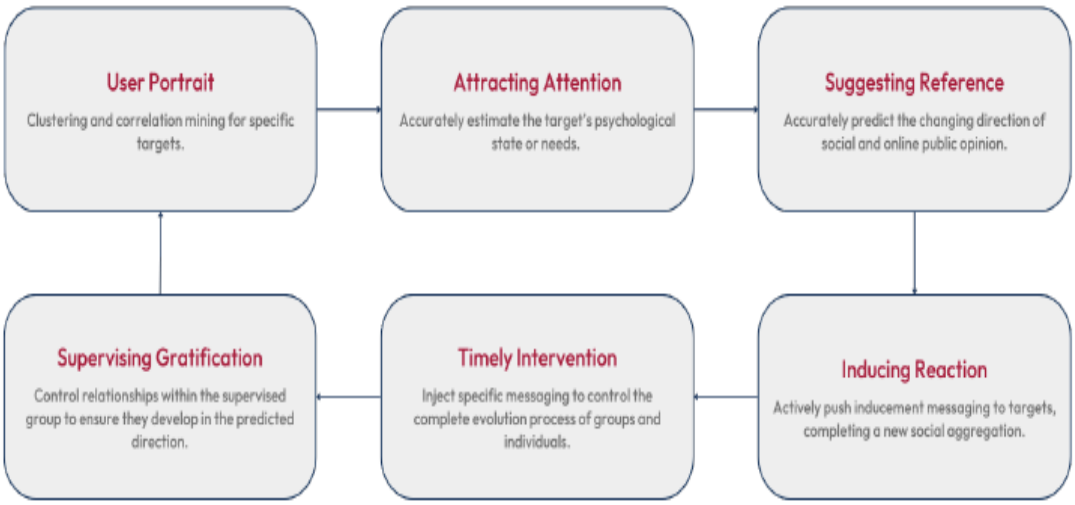


Figura 21. Modelo de guerra cognitiva algorítmica de China. Fuente: https://scsp222.substack.com/p/algorithmic-cognitive-warfare-the

Una vez descritos los elementos conceptuales que caracterizan la guerra cognitiva, y advertidos de sus efectos perniciosos sobre la sociedad, es necesario que unidades especializadas de los gobiernos responsables de las mismas puedan **elaborar una estrategia** que permita integrar todas las actuaciones necesarias para su defensa y neutralización. Esta estrategia requiere el desarrollo de un **modelo integrado de defensa cognitiva**. Este aspecto se trata en la siguiente sección.

[51] Su uso no es solo válido para conducir una guerra cognitiva hacia "enemigos" situados en el exterior del país, sino también forma parte del control de la población en el interior.

[52] https://www.state.gov/wp-content/uploads/2020/05/What-is-MCF-One-Pager.pdf

2. Evolución del contexto de la guerra cognitiva

2.1. Guerra cognitiva en el contexto de la OTAN

El *Mando Aliado de Transformación de la OTAN* está desarrollando el **Concepto de Guerra Cognitiva** con la contribución directa de más de 20 países miembros de la OTAN, académicos y expertos de la industria[53].

La OTAN ha pretendido fortalecer este concepto partiendo del objetivo de "***explotar el error de racionalidad***", y avanzando hacia una visión operativa de la misma, como se indica en la figura 22, que combine la **influencia** con la **discapacidad cognitiva**. Obsérvese que se trata de integrar el dominio de las **operaciones psicológicas** (PSYOPS), promoviendo una influencia motivada, con el dominio de la **guerra cognitiva**, buscando la discapacidad cognitiva, mediante la combinación de diversas técnicas desde un enfoque multidisciplinar.

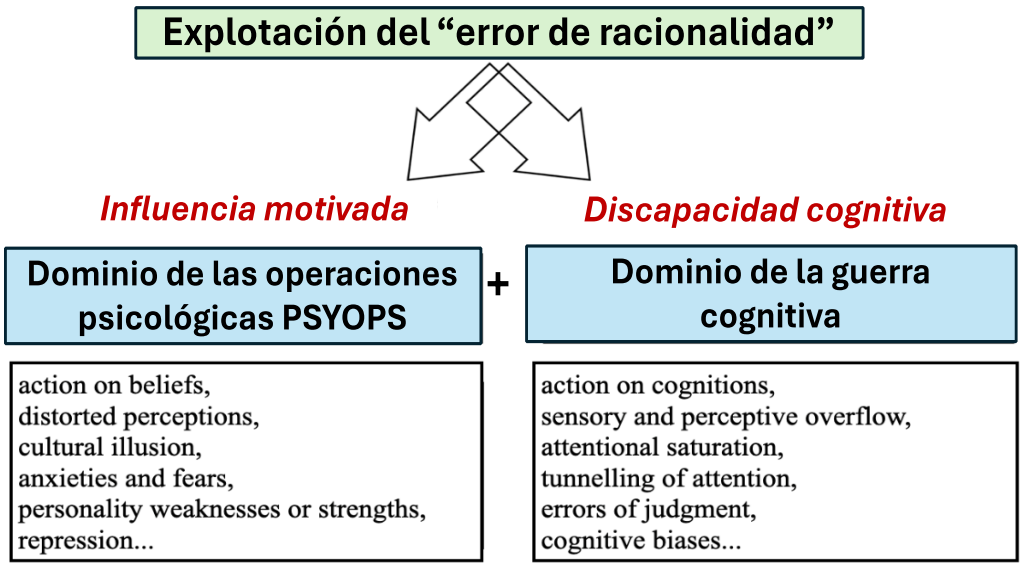


Figura 22. Contexto de la guerra cognitiva. Fuente: NATO STO-MP-AVT-211

Desde la visión de la OTAN (Deppe and Schaal, 2024), la guerra cognitiva emerge desde la **relación entre los tres elementos combinados** representados en la figura 23: **ataques cognitivos** (disrumpir el bucle de decisión, manipular la cognición, las actitudes y el comportamiento), **conflicto bélico** (para conseguir ventaja sobre el adversario, degradar sus capacidades (*Military Instrument of Power, MIoP*) y atacar al vector de la sociedad) y el uso de la **tecnología** existente y emergente (*Emerging and Disruptive Technologies, EDT*).

[53] https://www.act.nato.int/article/cogwar-concept/

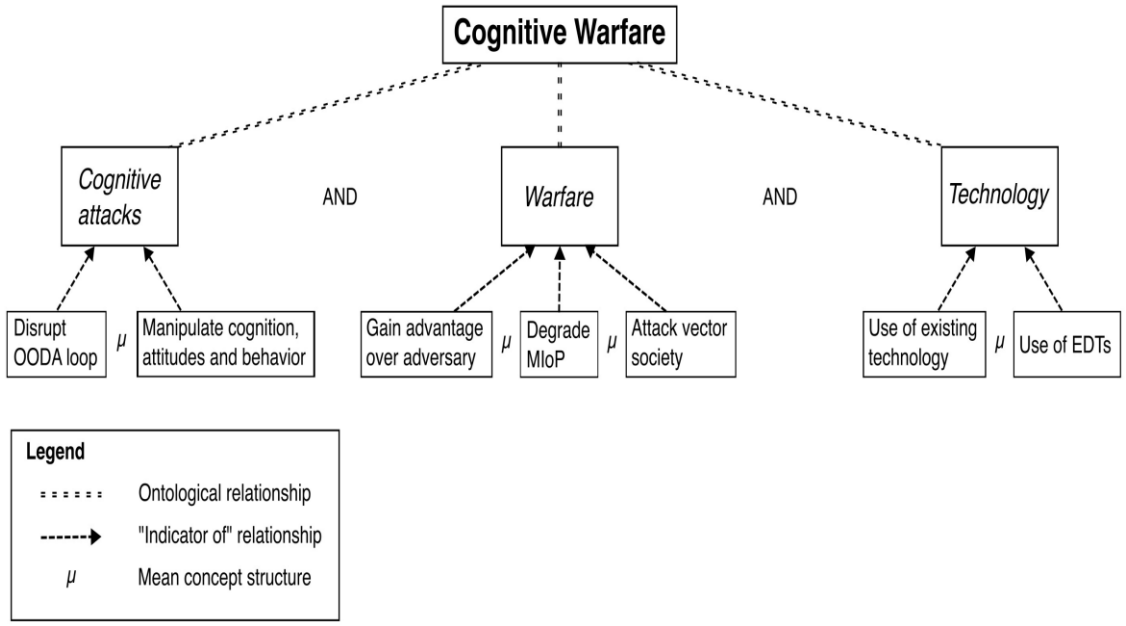


Figura 23. Componentes de la guerra cognitiva. Fuente: Deppe and Schaal, 2024

Para Deppe y Schaal (2024) el solape conceptual mencionado entre guerra híbrida y cognitiva sugiere que "*la guerra cognitiva podría integrarse efectivamente como un elemento táctico dentro de estrategias de amenaza híbrida más amplias, mejorando así la profundidad analítica y la comprensión de los efectos híbridos, al situar las operaciones cognitivas dentro de un contexto establecido y multifacético*".

Una visión más elaborada del **flujo del procesamiento de la información adaptada al conflicto** es la procedente de Lupiáñez (2024), en el que se plantean cinco fases partiendo del esquema mental de un individuo afectado por la información de un acontecimiento, lo que genera la elaboración de un pensamiento o su interpretación, supuestamente racionalizado, que conduce a unos sentimientos, y que, finalmente, llevan a tomar una acción o a no hacer nada (inacción).

Dentro de la doctrina de la OTAN, es probable que el concepto de guerra cognitiva, desde este **punto de vista conceptual amplio,** y su enfoque ligado al uso de tecnologías emergentes y disruptivas, como es la IA o la neurotecnología, ambas en rápida evolución, requiera efectuar **actualizaciones frecuentes y modificaciones** para que sigan siendo efectivas y pertinentes para los actores estatales implicados en la guerra cognitiva. No se trata, por tanto, todavía, de un dominio maduro y establecido.

2.2. Modelización del marco de actuación

A la vista de lo mencionado hasta el momento, un **modelo de guerra cognitiva debe estar centrado en la cognición** a la que se subordinan las actuaciones físicas en el terreno, y sobre los sistemas de información ligados a la **difusión de narrativas** en el dominio cognitivo a nivel nacional e internacional. La dificultad subyace en el elevado número de conceptos a tener en cuenta y sus complejas interacciones.

Para abordarlo de forma sistémica, Bebber y Marshal (2024) han elaborado una **ontología de la guerra cognitiva**, entendido como un sistema formal para organizar el conocimiento sobre este tema, construido en base a **cinco dimensiones instrumentales**:

1. **Herramientas que explotan los sesgos cognitivos y la percepción**: Estas amenazas manipulan los sesgos cognitivos y las vulnerabilidades perceptivas de las personas para dar forma a sus opiniones y comportamientos.
2. **Herramientas relacionadas con la neurociencia y la biología**: Los adversarios aprovechan los avances en neurociencia y biología para influir y controlar los procesos cognitivos de los individuos.
3. **Herramientas que explotan la psicología social y la dinámica de grupo**: Los adversarios aprovechan la psicología social y la dinámica de grupo para manipular el comportamiento del grupo, crear polarización o influir en la toma de decisiones colectivas.
4. **Herramientas que emplean aplicaciones tecno-sociales**: Los adversarios utilizan la tecnología de la información para difundir narrativas, participar en ingeniería social y realizar operaciones de información.
5. **Herramientas relacionadas con la tecnología de la información**: La tecnología de la información proporciona herramientas para los ciberataques, las campañas de desinformación y la interrupción de infraestructuras críticas.

Gran parte de la relevancia que ha adquirido la guerra cognitiva en los últimos años procede del reconocimiento de que **en las cinco dimensiones citadas se están produciendo avances tecnológicos muy rápidos, que alimentan la capacidad de integrar su uso en procesos específicos muy elaborados de guerra cognitiva**.

A partir de ello, es posible **elaborar modelos sobre los que definir procedimientos de acción,** incorporando diversos actores con el fin de permitir reducir el efecto de posibles ataques cognitivos. En la figura 24 se representa esquemáticamente un **modelo** bastante completo procedente de *Takagi* (2024).

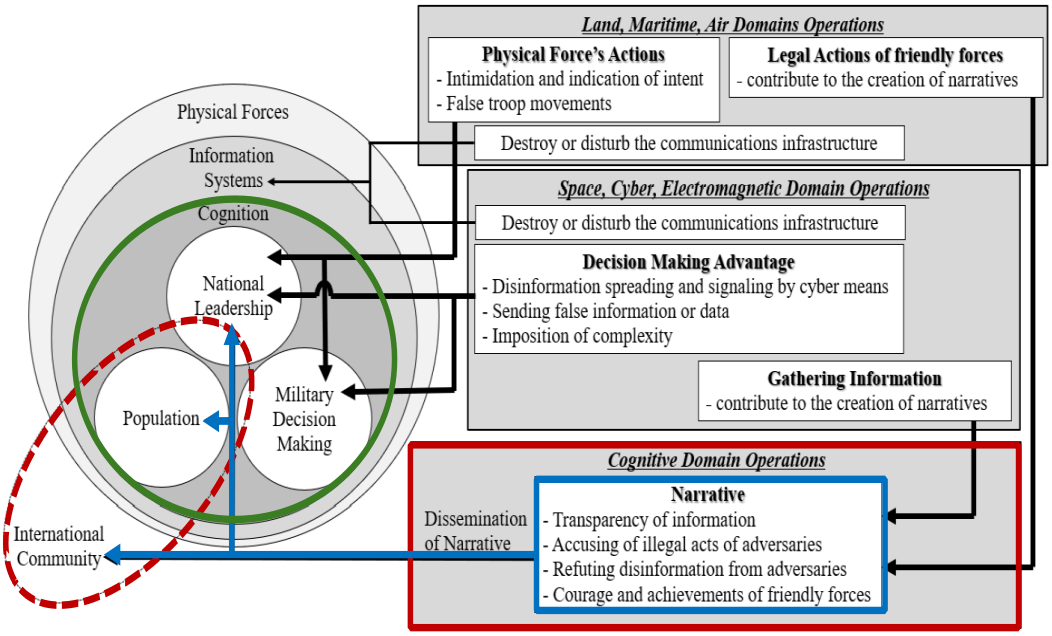


Figura 24. Modelo de guerra cognitiva. Fuente: https://www.hudson.org/corruption/cognitive-centric-warfare-modelling-indirect-approach-future-warfare-koichiro-takagi

Obsérvese en la parte de la izquierda de la figura 24 que el **nivel de cognición**, en la base del sistema de información y del de las fuerzas físicas, actúa sobre tres elementos clave: el **liderazgo nacional**, la **población**, y la **toma de decisiones militares**. Es sobre los dos primeros sobre los que influye de forma directa la diseminación de la **narrativa** como elemento clave de las operaciones del dominio cognitivo (a la derecha debajo de la figura); elemento que también influye en la opinión de la comunidad internacional.

La **construcción de la narrativa** cuya difusión es objeto de la guerra cognitiva se aprovecha de la información capturada en los dominios espacial, ciber y electromagnético, y de las actuaciones de fuerzas aliadas en operaciones en los dominios terrestre, marítimo y aéreo. Específicamente, las **operaciones del dominio cognitivo relacionadas con la narrativa** poseen cuatro elementos básicos.

- Transparencia de la información
- Acusación de actos ilegales realizados por los adversarios
- Refutación de desinformación generada por los adversarios
- Puesta en valor de los logros de fuerzas amigas.

La UE ha desarrollado un modelo de defensa elaborado con un **enfoque de pensamiento sistémico del ecosistema de la UE, con el fin de aumentar la resiliencia contra las amenazas híbridas,** incluyendo la anticipación a los objetivos estratégicos del adversario.

El enfoque de la UE (véase figura 25) se basa en el **fortalecimiento de un ecosistema de resiliencia integral** basado en promover el esfuerzo intersectorial de toda la sociedad, aprovechando las interrelaciones cruciales entre temas que, a menudo, se abordan por separado dentro de diferentes espacios de actuación.

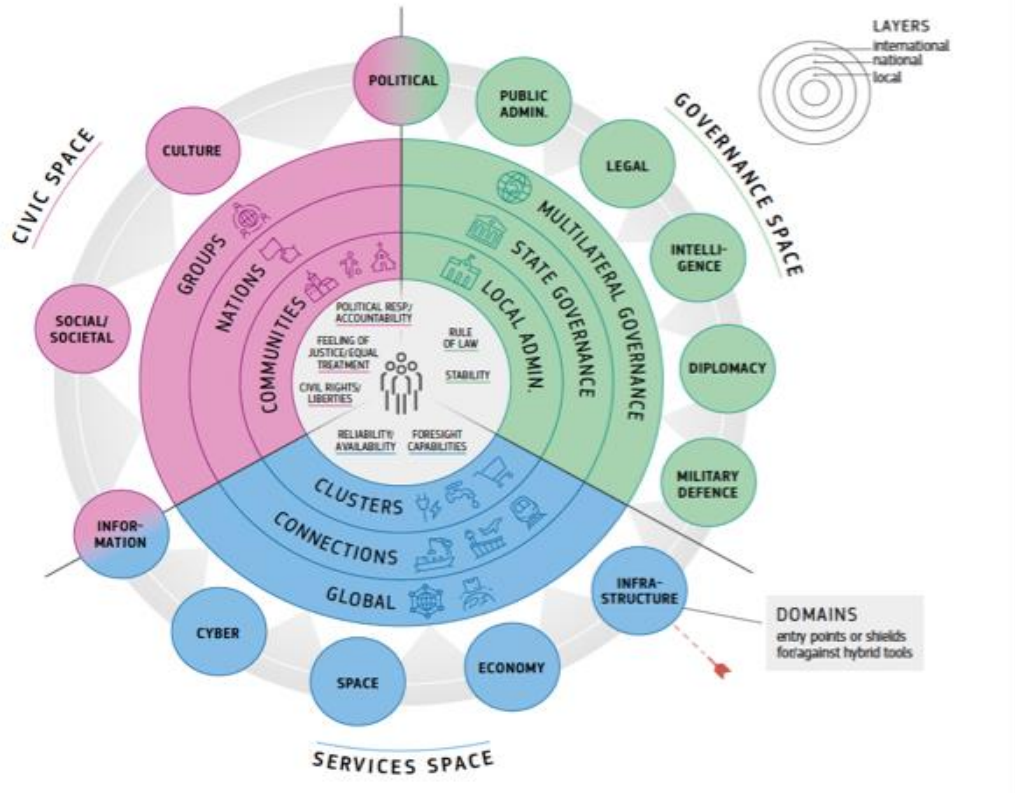
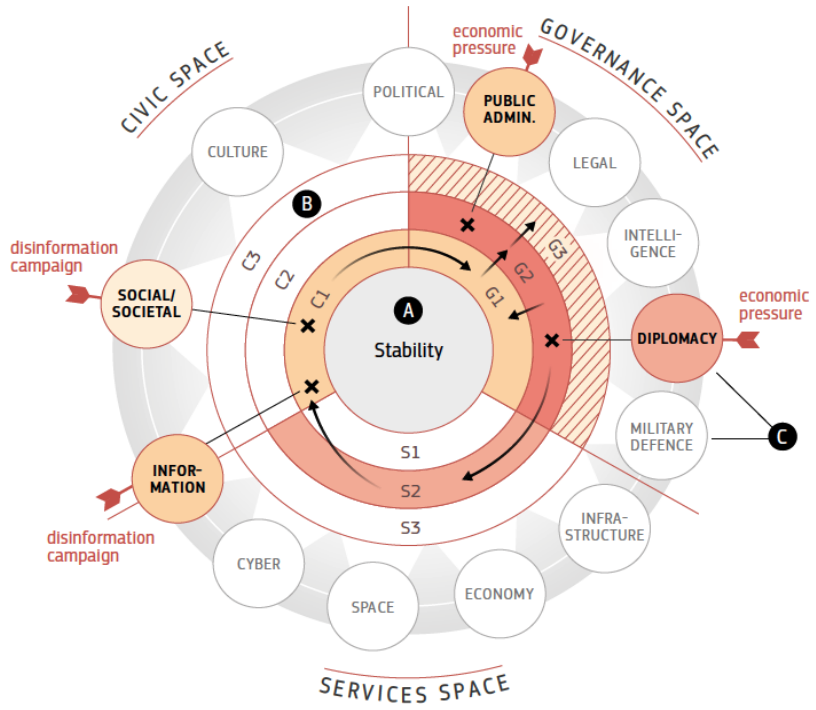


Figura 25. Fuente: https://www.hybridcoe.fi/wp-content/uploads/2023/04/CORE_comprehensive_resilience_ecosystem.pdf

En la figura 25 pueden verse los **elementos básicos de este modelo,** en el que se destacan tres áreas: **espacio cívico**, **espacio de servicios** y **espacio de gobernanza**, con tres capas de actuación en cada uno de ellos: local, nacional e internacional. En los espacios citados se identifican unos **"dominios" de actuación** que actúan de puntos de entrada o "escudos" contra las herramientas híbridas con el fin de **incrementar la resiliencia de la sociedad.**

Las **interconexiones entre dominios son cruciales para desarrollar la resiliencia frente a las amenazas híbridas**. A medida que los efectos de las amenazas híbridas puedan extenderse a través de dominios, es necesario interrumpir su propagación y contener los efectos. Esto no significa que los dominios deben estar desconectados, sino que las propias conexiones entre ellos deben ser resilientes.

El modelo permite **analizar determinadas acciones de guerra híbrida,** como se indica en la figura 26, focalizado para el caso de campañas de desinformación. En la parte inferior de la figura 26 puede verse la interpretación del **modelo de resiliencia**. Obsérvese que, para los autores, solo algunos de los "dominios" identificados están afectados en el caso concreto de las campañas de desinformación.



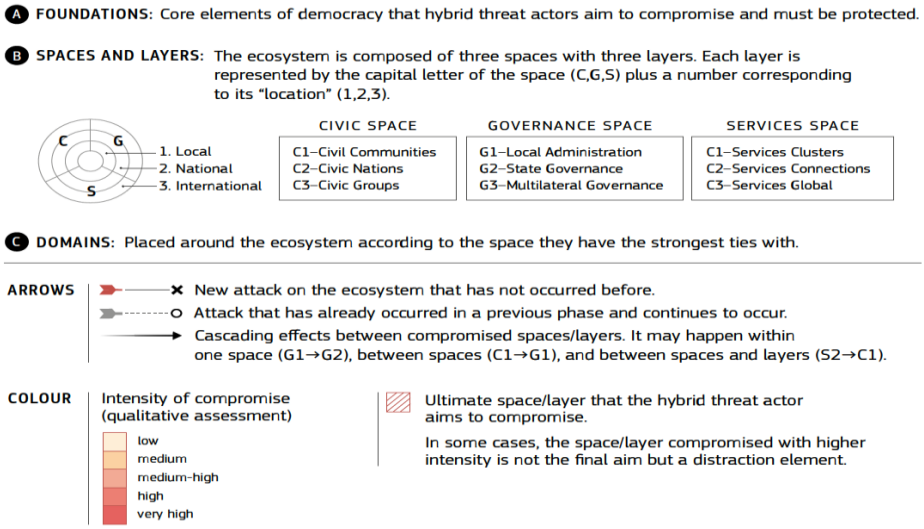


Figura 26. Modelo de resiliencia propuesto por la UE ante la guerra híbrida. Fuente: https://www.hybridcoe.fi/wp-content/uploads/2023/04/CORE_comprehensive_resilience_ecosystem.pdf

El objetivo pretendido con el desarrollo de este modelo es analizar cómo diferentes actores, desde sus diversas perspectivas, deben intervenir de manera coordinada para tener éxito en la guerra híbrida con la idea de proteger un elemento clave de la democracia como es la estabilidad interviniendo desde diversos ámbitos diferentes. La idea subyacente es que es necesario crear un "***ecosistema de lucha contra la guerra híbrida***" que implique la coordinación y el intercambio de información entre diferentes tipos de actores.

Desde luego, como se indica en el ejemplo de la figura 26, una de las acciones más relevantes (empleando cualquier modelo de resiliencia) es defenderse de la que, posiblemente, sea una de las armas más empleadas en las guerras híbridas, que es la **guerra informativa** de la que las **campañas de desinformación** constituyen el elemento fundamental; generalmente, ligadas a ciberataques y a otros elementos de guerra hibrida de baja o mediana intensidad como atentados a infraestructuras. La siguiente sección aborda explícitamente el concepto de desinformación y sus elementos relacionados.

2.3. Guerra informativa: Desinformación

Una herramienta básica para implementar la guerra cognitiva es la generación, difusión y control de la **desinformación.** No es un fenómeno nuevo. Boswinkel et al. (2022) recuerdan que ya *durante la guerra del Peloponeso entre Esparta y Atenas, hace unos 2400 años, el espacio de la información se explotó para engañar a las fuerzas enemigas, mantener la confianza de las tropas e influir en la opinión pública*. Obviamente, las tecnologías de la información han cambiado profundamente la escala, el alcance, la velocidad y la capacidad de control de las operaciones de información.

Se trata de un **fenómeno social complejo** en el que, si bien las tecnologías juegan un papel relevante, no pueden limitar una concepción limitada de la desinformación. Una forma de evitar este tipo de simplificación es considerar el **entrelazamiento de las dimensiones social, económica y estratégica como una característica definitoria de los fenómenos de desinformación**. Esas tres dimensiones de la desinformación convergen en torno a la naturaleza intencional de la desinformación: se trata de procesos planificados de carácter

estratégico que buscan unos efectos determinados en el marco de las dinámicas de influencia social (Aguado et al., 2024).

Algunos conceptos y términos relevantes relacionados para analizar el **contexto de la desinformación** que debemos tener presentes, y demuestran la amplitud de enfoques, son los siguientes:

- **Desinformación e información errónea**: Difusión deliberada de información falsa o tendenciosa para engañar, confundir o manipular la opinión pública.
    - Mientras que la **información errónea** se refiere a la difusión accidental de información inexacta, la **desinformación** no solo es inexacta, sino que tiene por objetivo engañar y se difunde con el fin de causar graves prejuicios.
    - Es verdad que la alteración intencionada de una fuente de información puede contribuir a difundir información errónea por quien la utilice como base de su propia información. Con ello, se produce un efecto de **multiplicación de la desinformación,** que es un factor especialmente buscado en el diseño de las campañas.
- **Operaciones Psicológicas (PSYOP)**: Operaciones destinadas a influir en las emociones, los motivos, el razonamiento objetivo,
    - Con ello se pretende, en última instancia, influir en el comportamiento de gobiernos, organizaciones, grupos e individuos extranjeros
- **Ingeniería social**: Actuaciones diseñadas para manipular a las personas con el fin de que realicen acciones o divulguen información errónea o confidencial, a menudo, mediante la apariencia de comunicación legítima.
- **Posverdad.** La posverdad se entiende como una tendencia de las audiencias a optar por la congruencia de opiniones y la facilidad de acceso/cognición en lugar de la veracidad como criterios principales para la selección de la información (Kalpokas, 2024)[54].
- **Operaciones de influencia**. Término utilizado en el ámbito geopolítico para definir el uso de técnicas de desinformación por parte de los Estados para inmiscuirse en los asuntos de otros (Bradshaw, 2020),
- **Relaciones públicas oscuras**. Término utilizado para expresar dinámicas similares entre empresas o marcas competidoras (Rodríguez-Fernández, 2023).
- **Presuasión.** Influencia previa a la definición del concepto o la generación de la voluntad por parte del sujeto. Se moldea esa voluntad antes de que existan en lugar de tener que cambiarla después, algo mucho más costoso (Cialdini, 2016).

Bernal et al. (2020) diferencian la guerra cognitiva de la guerra informativa haciendo hincapié en sus distintos objetivos: mientras que la **guerra informativa** se centra en **controlar la difusión de la información,** la **guerra cognitiva** tiene como objetivo estratégico **controlar o alterar la forma en que las personas reaccionan a la información**. Estos autores definen la guerra cognitiva como la "*militarización de la opinión pública, por parte de una entidad externa, con el propósito de (1) influir en la política pública y gubernamental y (2) desestabilizar las instituciones públicas"*.

[54] Recientemente, se ha empezado a emplear otro concepto relacionado denominado ***posrealidad***, más sutil, seductor y difícil de detectar, centrado en la aparición de realidades fabricadas, hechas a la medida de nuestros deseos. https://contextomedia.com/de-la-posverdad-a-la-posrealidad-la-mentira-es-tu-mas-profundo-deseo/

En el contexto del presente informe con énfasis en aspectos tecnológicos, se define la guerra cognitiva como "*la construcción estratégica y la difusión de narrativas para dar forma a las percepciones, crear ideologías cohesivas o desmantelar las narrativas sociales y políticas de los adversarios, debilitando sus capacidades en un contexto de confrontación híbrida aprovechando el uso de las tecnologías de información y comunicaciones"*.

En los últimos años, el problema derivado del **incremento de la desinformación**, apoyado y potenciado por el desarrollo de tecnologías como la IA, ha penetrado en gobiernos, instituciones y ciudadanos, incluso llegando hasta las Naciones Unidas, por su **efecto desestabilizador en las sociedades democráticas**. En este sentido, la ***Asamblea General de las Naciones Unidas*** ha expresado su preocupación por la proliferación de la desinformación y requiere promover la **cooperación internacional en la lucha contra la desinformación**.

El informe del *Secretario General de la ONU* de agosto de 2022, titulado *"Contrarrestar la desinformación para promover y proteger los derechos humanos y las libertades fundamentales"* (ONU, 2022), reclama esa cooperación, al mismo tiempo que requiere que en esa lucha se preserven derechos y libertades esenciales del ser humano; equilibrio difícil de conseguir en la práctica. En todo caso, supone un cambio de actitud de las Naciones Unidas frente a una interpretación habitual de prohibir el uso de la "fuerza" solo cuando se trata de acciones con "efectos cinéticos" (Bernal, et al., 2020)[55].

Por parte de la UE, también se han generado múltiples resoluciones para combatir la desinformación. Una de las últimas, referida a Rusia, es la *Resolución del Parlamento Europeo,* de 23 de enero de 2025, sobre la **desinformación y la falsificación de la historia por parte de Rusia para justificar su guerra de agresión contra Ucrania**[56].

En la Resolución mencionada del Parlamento Europeo se dice: "*el régimen ruso ha estado haciendo un uso generalizado de la desinformación (incluida la basada en argumentos históricos distorsionados) y la manipulación de la información y las injerencias extranjeras en un intento de justificar su crimen de agresión para incitar a la población rusa a apoyar su régimen ilegal y su guerra de agresión ilegal contra Ucrania, para interferir en los procesos democráticos de otros países y para reducir el apoyo entre sus poblaciones a la asistencia internacional continuada y al respaldo a Ucrania contra la guerra de agresión de Rusia; que el régimen ruso niega una identidad nacional diferenciada a Ucrania, y afirma falsamente que forma parte del mundo ruso («Russki mir»), una narrativa arraigada en la ideología imperialista"*.

Como una de las **medidas propuestas por el Parlamento Europeo,** se dice expresamente: *"Pide encarecidamente a la Unión y a sus Estados miembros que redoblen y coordinen sus esfuerzos, conjuntamente con socios afines, para contrarrestar con rapidez y firmeza la desinformación y la manipulación de la información y las injerencias extranjeras por parte de Rusia a fin de proteger la integridad de sus procesos democráticos y reforzar la resiliencia de las sociedades europeas, entre otras cosas promoviendo*

[55] El párrafo 4 del Artículo 2 de la Carta de las Naciones Unidas, dispone que *"todos los Miembros se abstendrán, en sus relaciones internacionales, de recurrir a la amenaza o al uso de la fuerza contra la integridad territorial o la independencia política de cualquier Estado, o en cualquier otra forma incompatible con los propósitos de las Naciones Unidas"*.

[56] https://www.europarl.europa.eu/doceo/document/TA-10-2025-0006_ES.html

activamente la alfabetización mediática y apoyando a los medios de comunicación y al periodismo profesional de calidad, en particular al periodismo de investigación que destapa la propaganda rusa, sus métodos y sus redes, así como apoyando la investigación sobre las nuevas tecnologías de influencia híbrida".

La última de las frases se refiere a la necesidad de apoyar el **desarrollo de nuevas tecnologías específicas** contra la desinformación (p.ej. para la detección automática, control inteligente de la difusión, formación del ciudadano, y reducción del impacto), como serían, sin citarlas, las derivadas de la inteligencia artificial, o la neurotecnología.

El éxito de las **operaciones de influencia basadas en la desinformación** depende del cumplimiento de tres condiciones principales. En primer lugar, deben producir **narrativas** que transmitan de manera efectiva y persuasiva el mensaje diseñado para un objetivo concreto; en segundo lugar, deben ser capaces de **difundir ese mensaje** a un público objetivo lo suficientemente amplio como para influir en las decisiones políticas a nivel nacional o institucional; y, en tercer lugar, deben **persuadir** a una proporción suficiente de personas de ese mensaje para que sus comportamientos generen una cadena de reacciones que logren el objetivo pretendido.

Puede hablarse de un "**ecosistema de la desinformación**" que, siguiendo a Aguado et al (2024), agrupa a los siguientes actores con complejas interacciones entre ellos: Estados, empresas y corporaciones, partidos políticos, actores sociales, actores particulares, facilitadores e intermediarios, medios de comunicación, medios partidistas y plataformas de redes sociales. Todos ellos intervienen con diferentes papeles[57] en función del país, momento y temática en la creación, difusión y aprovechamiento de las campañas de desinformación.

Los análisis realizados en diferentes casos de estudio apoyan la suposición de que **las operaciones de influencia pueden moldear las preferencias de política exterior** de los ciudadanos, pero, al mismo tiempo, han revelado su **impacto limitado**. La experiencia de las campañas de desinformación de Rusia en Ucrania indica que no es sencillo conseguir influencias apreciables con movimientos relevantes en las opiniones de la población empleando simplemente información basada en sitios web y su difusión en redes sociales[58].

No siempre el impacto ha sido tan limitado; su influencia fue determinante en 2003 en la campaña para convencer a la población de muchos países de la **existencia real de armas de destrucción masiva en Irak,** lo que justificó la decisión de comenzar una guerra convencional[59]; también, en cierta medida, el "*éxito*" en predisponer a parte de la

[57] Los roles identificados en aguado et al. (2023) son: activadores/lanzadores, impulsores, legitimadores, difusores, relacionadores y condicionadores.

[58] En https://verifica.efe.com/desinformacion-y-mentiras-ucrania-y-rusia/ pueden encontrarse múltiples ejemplos. Estas campañas de desinformación también son llevadas a cabo por Ucrania sobre Rusia. No se dispone de información fiable sobre cómo han influido, pero, de igual manera que en el sentido de Rusia sobre Ucrania, puede suponerse que su efecto tampoco parece ser muy elevado.

[59] El 5 de febrero de 2003, el secretario de Estado de Estados Unidos, Colin Powell, en el Consejo de Seguridad de la ONU argumentó: *"Mis colegas, cada declaración que hago hoy está respaldada por fuentes, fuentes sólidas. No son afirmaciones. Lo que les estamos dando son hechos y*

población del Reino Unido en contra de la UE durante la campaña del BREXIT basado en tres supuestos "incorrectos" con fuerte apoyo de la *prensa euroescéptica*[60] (Parnell, 2023).

Vale la pena mencionar que los **medios de difusión global centralizados**, y específicamente la **televisión**, conservan importantes ventajas sobre las redes sociales descentralizadas como **vectores para las operaciones de influencia** (Maschmeyer et al., 2023). Como reacción, puede mencionarse la prohibición impuesta por la UE y el gobierno ucraniano a los tres canales de televisión partidistas identificados como importantes medios de desinformación (entre ellos RT).

En marzo de 2022, tras la invasión de Ucrania, la UE suspendió urgentemente las actividades de radiodifusión de ***Sputnik*** y ***RT-Russia Today*** (bajo el control permanente, directo o indirecto, de las autoridades de la Federación de Rusia) en la UE, o dirigidas a esta, *"hasta que cese la agresión contra Ucrania y hasta que la Federación de Rusia y sus medios de comunicación asociados dejen de llevar a cabo acciones de desinformación y manipulación de la información contra la UE y sus Estados miembros"*[61].

Posteriormente, en mayo de 2024, la Unión Europea anunció la **prohibición de cuatro medios de comunicación adicionales,** acusados de difundir propaganda pro-Kremlin y "desestabilizar" a los países vecinos de Ucrania. Los cuatro medios son *"Voz de Europa", "RIA Novost", "Izvestia" y "Rossiyskaya Gazeta",* dado que, para la UE, están "*bajo el control permanente, directo o indirecto de Rusia y han sido "decisivos" para fomentar el apoyo a su invasión ilegal de Ucrania"*.[62]

La oportunidad de lanzamiento de estas campañas de desinformación puede ser múltiple; sin embargo, es muy habitual que surjan o se intensifiquen en periodos electorales. En la figura 27 pueden verse algunos ejemplos de actuación de **desinformación** relacionados con la guerra cognitiva ocurridos en varios países del mundo con ocasión de diversas **contiendas electorales**. El objetivo es el de **influir en los resultados de la votación** mediante el empleo de herramientas de control del acceso a la información y su difusión.

---

conclusiones basadas en inteligencia sólida". Powell utilizó información que los funcionarios de inteligencia le aseguraron que era creíble. Había fotos de reconocimiento, mapas y gráficos elaborados, e incluso conversaciones telefónicas grabadas entre miembros de alto rango del ejército iraquí. https://www.opb.org/article/2023/12/11/20-years-ago-the-u-s-warned-of-iraq-s-alleged-weapons-of-mass-destruction/?utm_source=chatgpt.com

[60] Los tres elementos fueron que el Reino Unido enviaba 350 millones de libras a la semana a la UE; que Turquía se uniría a la UE de forma inminente; y que los migrantes estaban desbordando al Reino Unido debido a las políticas migratorias de la UE.

[61] https://www.consilium.europa.eu/es/press/press-releases/2022/03/02/eu-imposes-sanctions-on-state-owned-outlets-rtrussia-today-and-sputnik-s-broadcasting-in-the-eu/

[62] https://es.euronews.com/my-europe/2024/05/17/la-union-europea-prohibe-la-difusion-de-propaganda-rusa-a-cuatro-medios-de-comunicacion

Figura 27. Desinformación en contiendas electorales en diferentes países. Fuente: https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence

Por razones políticas a largo plazo no necesariamente ligadas a contiendas electorales, el **gobierno chino** ha concretado sus campañas de desinformación con ataques sobre *Estados Unidos* o *Taiwán,* con episodios que han sido especialmente intensos, aunque tampoco parece que hayan sido muy efectivos.

En el caso de los **ciberataques** procedentes de China, es Taiwán el país (no reconocido por la propia China continental) que concentra gran parte de los ciberataques procedentes de China. En un reciente informe de la *Oficina de Seguridad Nacional de Taiwán* (*National Security Bureau, NSB*), se indicó que China efectuó en 2024 más de 2 millones de *actos de «desinformación» contra Taiwán en el marco de las «estrategias híbridas» llevadas a cabo por el Partido Comunista (PCCh) para «socavar la confianza pública en el Gobierno» y «exacerbar las divisiones sociales» en la isla*.[63] La figura 28, extraída del informe citado de NSB, da buena idea del volumen alcanzado y de su crecimiento (alrededor de un 40% interanual), así como de los medios de difusión empleados por China.

[63] La NSB detectó un total de 2,16 millones de piezas de desinformación procedentes de China el año pasado, superando las 1,33 millones registradas en 2023. https://www.swissinfo.ch/spa/taiw%c3%a1n-denuncia-m%c3%a1s-de-2-millones-de-actos-de-%22desinformaci%c3%b3n%22-de-china-en-2024/88673531

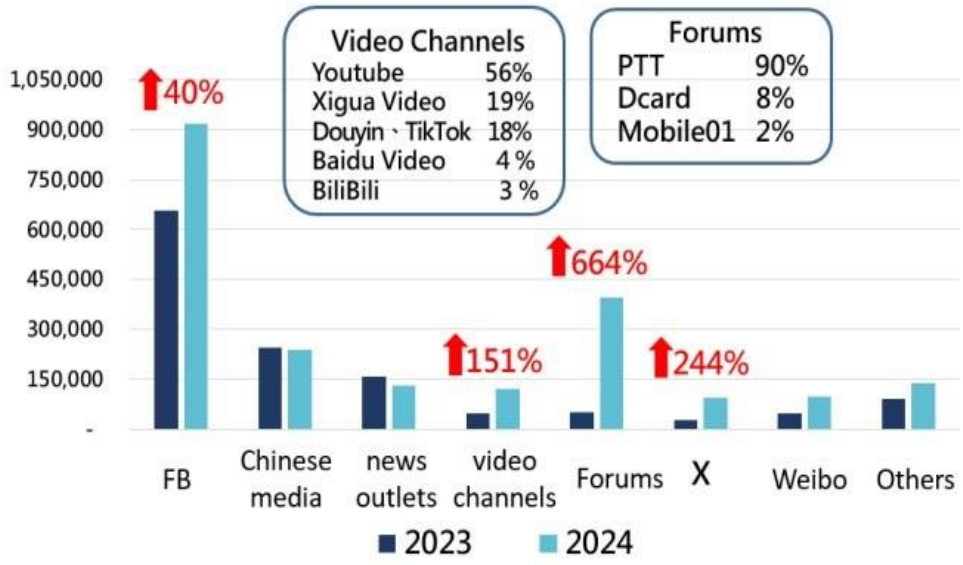


Figura 28. Desinformación de China sobre Taiwán NSB. Fuente: https://www.ocac.gov.tw/OCAC/Eng/Pages/Detail.aspx?nodeid=329&pid=71573982

China aprovecha la inteligencia artificial para producir desinformación, como el uso de tecnología de imagen falsa (‘*deepfake*’) para hacerse pasar por figuras políticas en vídeos falsos, intentando engañar y manipular la opinión pública. Para ello, emplea diferentes **grupos chinos especializados en la guerra cognitiva**. Dos de ellos son:

- **Flax Typhoon** (Storm 09-19) es el grupo de amenazas más destacado que tiene como objetivo la isla de *Taiwán*. Este grupo se dirige principalmente a las telecomunicaciones, la educación, la tecnología de la información y la infraestructura energética, por lo general aprovechando un dispositivo VPN (*Virtual Private Network*) personalizado para establecer directamente una presencia dentro de la red de destino.

- **Charcoal Typhoon** (CHROMIUM) tiene como objetivo las instituciones educativas, la infraestructura energética y la alta tecnología de fabricación en Taiwán.

Ningún país con capacidades tecnológicas será ajeno a estas prácticas. Taiwán, según China, realiza ciberataques y campañas de “*sabotaje anti-propaganda*” dirigidas hacia Pekín. El *Ministerio de Seguridad* de China acusa directamente a *Anonymous 64*, presuntamente vinculado a las fuerzas militares de Taiwán. Además, creen que su propósito es “*difundir contenido que denigra el sistema político de la China continental y sus principales políticas*”[64], Similar a lo que sucede en el sentido contrario.

También se ha indicado, sobre todo en Estados Unidos, que el uso de aplicaciones como *Deepseek* (véase figura 29 izquierda) se utiliza para reflejar la visión del mundo del Partido Comunista Chino en las respuestas ofrecidas a preguntas sobre temas como Taiwán, los uigures o la guerra de Ucrania. Parecida visión a lo que ocurrió en 2024 con otra aplicación como *TikTok*. Eso se acompaña de imágenes generadas por IA (véase figura 29 derecha),

[64] https://www.lisanews.org/internacional/china-denuncia-a-taiwan-presuntos-ciberataques-y-sabotajes/

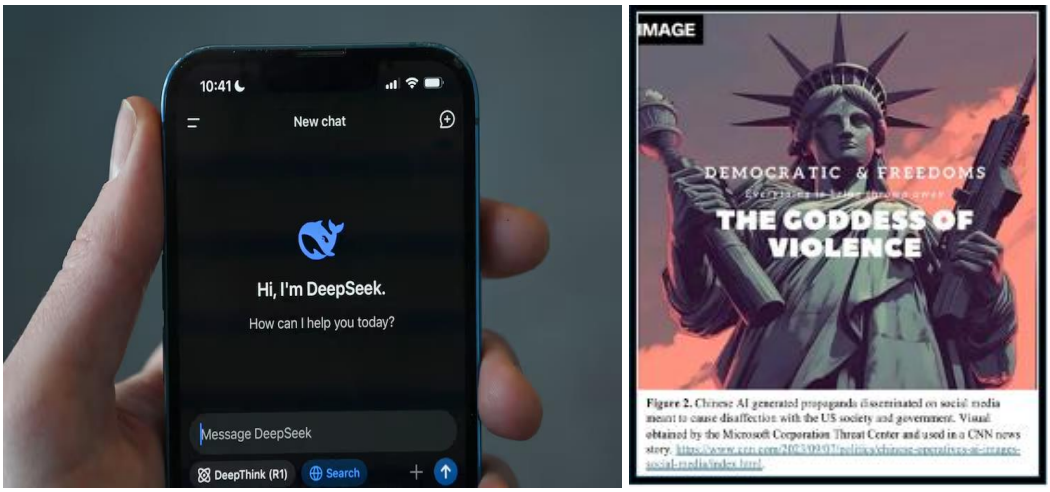


Figura 29. Izquierda: Aplicación de IA conversacional de Deepseek. Fuente: https://www.lanacion.com.ar/el-mundo/deepseek-bajo-la-lupa-investigadores-advierten-sobre-sesgo-ideologico-y-la-propaganda-china-nid02022025/ Derecha: imagen de propaganda china generada por IA. Fuente: https://smallwarsjournal.com/2025/01/22/the-challenge-of-ai-enhanced-cognitive-warfare-a-call-to-arms-for-a-cognitive-defense/

Otro ejemplo de campañas de desinformación es la **estrategia de desestabilización seguida por Irán sobre Israel** (de hecho, no solo sobre Israel sino también sobre sus aliados, como Estados Unidos). La estrategia de Irán se **concentra sobre un determinado tema en un corto periodo de tiempo,** empleando medios propios y de *proxies* (como Hamas, Hezbollah o los hutíes), como es el caso de la guerra de Gaza, en los que se combinan intervenciones públicas, amenazas y ataques alentados desde los líderes políticos iraníes o de su proxies. La estrategia iraní se desarrolla en **tres fases,** como se indica en la figura 30. Una primera fase, de carácter reactivo y de engaño, seguida de una segunda fase con ataques coordinados con un extenso número de grupos contra Israel, y una tercera fase, en la que se expande el ámbito geográfico de actuación.

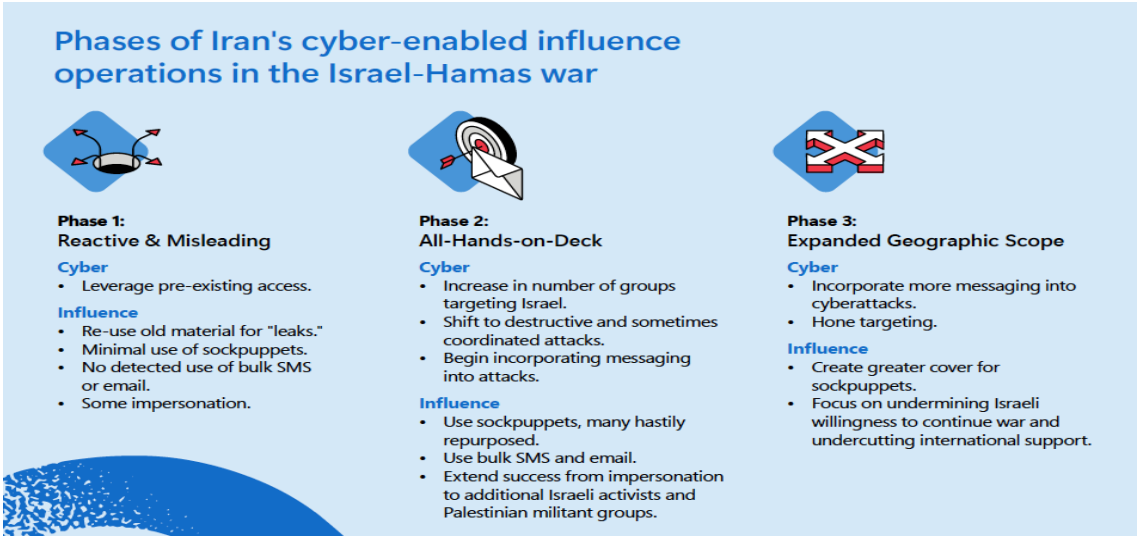


Figura 30. Fases de las operaciones de ciber influencia de Irán en el contexto de la guerra de Israel en Gaza. Fuente: https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/iran-surges-cyber-enabled-influence-operations-in-support-of-hamas

Otro **ejemplo concreto de desinformación** en otra zona conflictiva es el que se indica en la figura 31, en el que la foto empleada relativa a una operación de ciberataque (en este caso del tipo *ransomware*) se refiere a otra ciudad de Israel. El objetivo es difundir una narrativa sobre la alta precisión y grandes efectos de los ciberataques dirigidos no solo a la población de Israel (desmotivación y miedo) sino para elevar la moral de la población de Gaza e Irán.

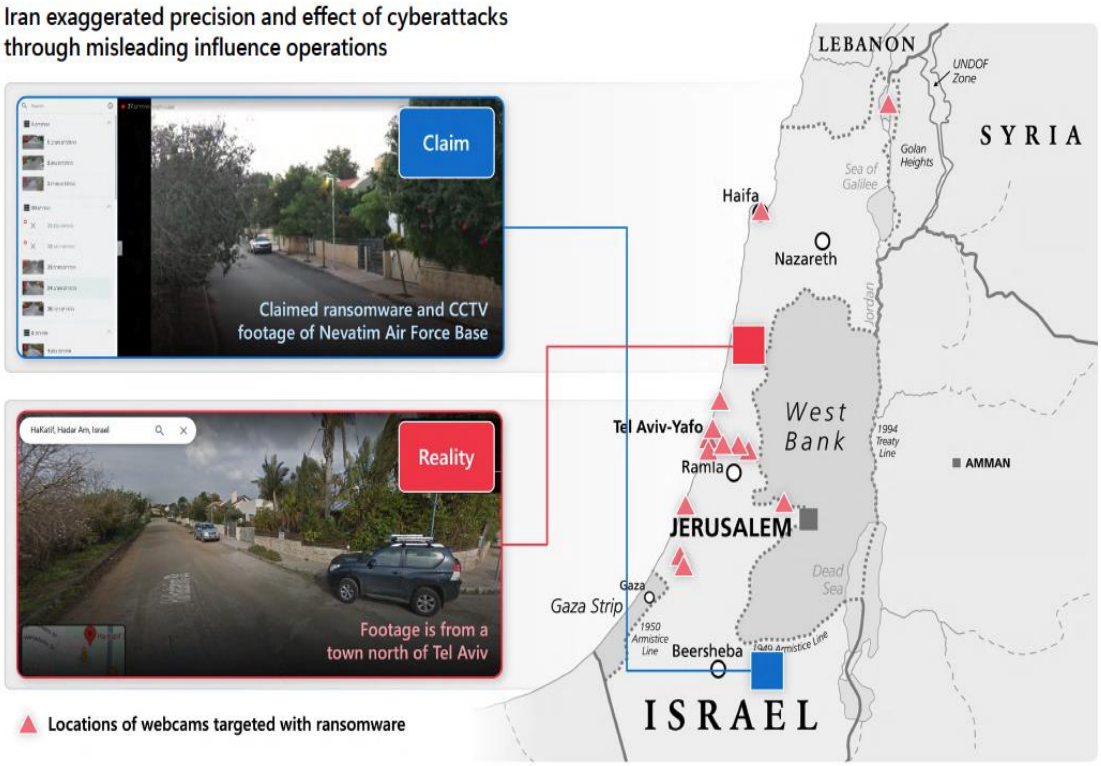


Figura 31. Ejemplo de desinformación sobre Israel. Fuente: https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/iran-surges-cyber-enabled-influence-operations-in-support-of-hamas

La actual guerra en la franja de Gaza es un ejemplo de cómo la **desinformación se utiliza como arma de guerra más allá de los límites geográficos y de los contendientes**. *Hamás* explota las redes sociales para dar difusión a imágenes impactantes de los ataques de Israel sobre la población civil como parte de narrativas que han sido documentadas y analizadas por organizaciones como *Human Rights Watch*, *Amnistía Internacional* y desmentidas por las Fuerzas de Defensa de Israel como parte de su campaña de propaganda (López-Garay, 2025).

Figura 32. Imágenes manipuladas artificialmente. Fuente: WALTER, Jan D. Fact check: AI-generated images of children in Gaza. DW, 02/02/2024 (en López-Garay, 2025).

2.5. Manipulación e Interferencia de Información Extranjera en la UE

Existe una considerable superposición conceptual entre la guerra cognitiva y conceptos próximos como la **Manipulación e Interferencia de Información Extranjera** (*Foreign Information Manipulation and Interference, FIMI*), las amenazas híbridas, el conflicto híbrido y, finalmente, la guerra híbrida.

El ***Servicio Exterior de la UE*** (EEAS) define el **FIMI**[65] como "*un patrón de comportamiento que amenaza o tiene el potencial de afectar negativamente a los valores, procedimientos y procesos políticos*". Dicha actividad es de carácter manipulador, llevada a cabo de manera intencional y coordinada. Los actores de dicha actividad pueden ser actores estatales o no estatales, incluidos sus representantes dentro y fuera de su propio territorio.

Anualmente, el EEAS genera un **informe sobre actuaciones en relación con incidentes FIMI**. Su primer informe data de febrero de 2023 (EEAS, 2023), el segundo de enero de 2024 (EEAS, 2024), y el tercero ha sido publicado en marzo de 2025 (EEAS, 2205).

Baste un ejemplo. El departamento de cultura y comunicación del Ministerio Federal de Asuntos Exteriores de Alemania había identificado en 2024 una red de más de 50.000 cuentas falsas que generaban hasta 200.000 mensajes diarios. Su objetivo era persuadir a los alemanes de que el apoyo del gobierno a Ucrania pone en peligro la prosperidad alemana y aumenta el riesgo de guerra nuclear "*identificando dudas y sentimientos de malestar existentes e intentando agrandarlas*" (Pujol, 2024).

Como Estrategia de la UE para defenderse contra la desinformación, se han creado unas estructuras especializadas como el **Sistema de Alerta Rápida (SAR)** para permitir actividades conjuntas con otras instituciones de la UE y los Estados miembros. El objetivo es desarrollar un marco y una metodología integral para la recopilación sistemática de

[65] https://www.eeas.europa.eu/eeas/tackling-disinformation-foreign-information-manipulation-interference_en

pruebas de incidentes FIMI, facilitada por un *Centro de Análisis e Intercambio de Información* (FIMI ISAC). El EEAS, en estrecha cooperación con la Comisión Europea y los Estados miembros, está reforzando continuamente el conjunto de instrumentos de la UE para hacer frente a la FIMI (*FIMI Toolbox*), con el fin de imponer sanciones a los autores.

La edición de 2024 identificó 750 incidentes FIMI investigados entre diciembre de 2022 y noviembre de 2023. (EEAS, 2024). Estos incidentes FIMI analizados se distribuyen entre muchos países, tal y como se indica en la figura 33. Obsérvese que se han concentrado en Ucrania (160 incidentes) y Estados Unidos. España presenta un número relativamente bajo (entre 6 y 15 incidentes).

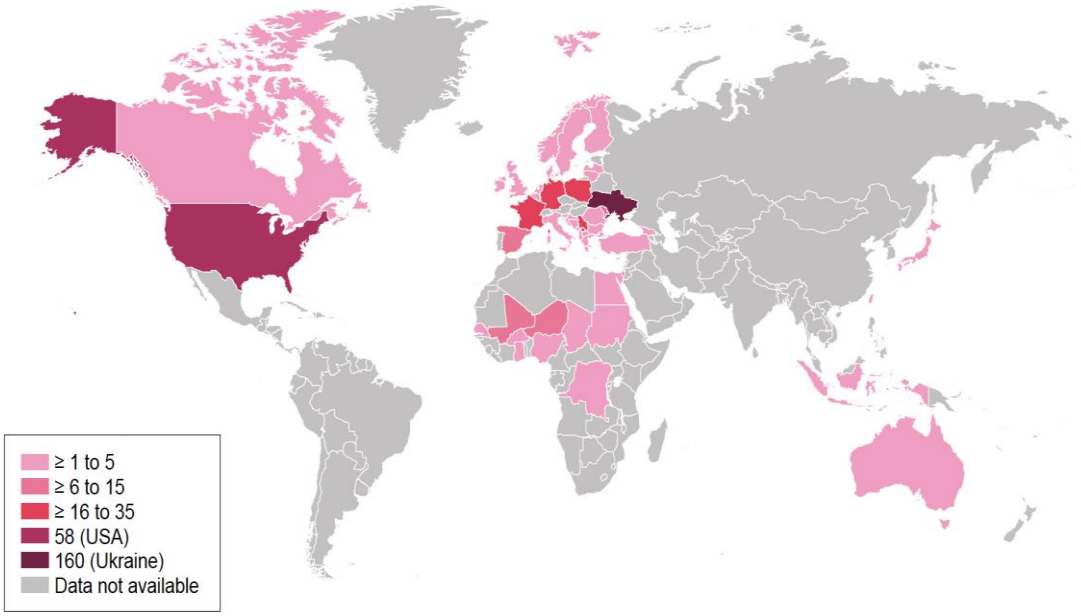


Figura 33. Incidentes FIMI en 2023. Fuente: EEAS (2024)

En el **tercer informe de incidentes FIMI** (EEAS, 2025) se detectaron y analizaron 505 incidentes entre el 4 de noviembre de 2023 y el 4 de noviembre de 2024. Dentro de esta muestra, se involucraron 38.000 canales únicos en 25 plataformas diferentes y se registraron un total de más de 68.000 observables (piezas de contenido) **afectando a 90 países**.

Las **conclusiones más relevantes** extraídas del informe indican:

- Ucrania sigue siendo el país más atacado en los incidentes investigados, con 257. Después de Ucrania, Francia también fue uno de los principales objetivos de actores hostiles, con 152 casos detectados por el SEAE, que se originaron en el ecosistema FIMI ruso y chino. Estos datos muestran que se necesitan esfuerzos continuos y adicionales para abordar el uso continuo de FIMI para socavar la estabilidad y la seguridad del país, así como para erosionar el apoyo a Ucrania.

- La mayor parte de los incidentes analizados fueron dirigidos contra organizaciones (85% de los incidentes), lo que implicó a 322 organizaciones diferentes. Entre ellas se encuentran entidades internacionales como la UE y la OTAN, y las fuerzas armadas de ciertos países occidentales, como Alemania, Francia y Estados Unidos. También fueron atacados varios medios de comunicación de diversos países como *Le Parisien*, *BBC, France 24, Der Spiegel* y *La Stampa*.

- Los eventos relevantes son catalizadores importantes para la actividad de FIMI. En el 21,3% de los incidentes analizados, la actividad de FIMI aprovechó la atención ya existente en torno a eventos como cumbres políticas, elecciones, emergencias, visitas oficiales y otros. Las elecciones han sido objetivos principales de FIMI durante 2024. Estas operaciones comienzan mucho antes del día de las elecciones y continúan después. Entre los eventos electorales con más incidentes registrados se encuentran las elecciones europeas de 2024[66], las elecciones presidenciales de Taiwán, las elecciones presidenciales de Estados Unidos, las elecciones presidenciales de Moldavia y el referéndum de adhesión a la UE, las elecciones parlamentarias de Georgia y las elecciones legislativas francesas.

- Aunque *X/Twitter* fue la plataforma más relevante, la gran mayoría de los incidentes no tienen lugar sobre una única plataforma, sino que tienden a estar activos en múltiples frentes, con contenido publicado de forma cruzada por diferentes cuentas en varias plataformas. La elección de la plataforma depende de los canales preferidos del público objetivo (p.ej. muchos incidentes dirigidos a países africanos se encuentran predominantemente en *Facebook*).

- El uso de la IA en las operaciones de FIMI constituyó una evolución más que una revolución en la realización de ataques FIMI, ya que los enfoques de respuesta existentes siguen siendo aplicables. El uso de IA en incidentes FIMI se ha vuelto más frecuente y los avances en IA generativa se han reflejado gradualmente en incidentes FIMI. Su uso facilita a los actores de amenazas la realización o automatización de ciertas actividades como, por ejemplo, la creación de contenido, además de hacerlas más rentables.

Desde un punto de vista de **impacto político**, la necesidad de mejorar la **protección de las elecciones contra acciones FIMI** ha adquirido mucha mayor relevancia. Los actores de FIMI comienzan a preparar sus operaciones para atacar las elecciones con mucha antelación e intensifican gradualmente sus ataques. Preparan un entorno de información alternativo, dirigido a los votantes, los partidos políticos y los candidatos, así como a la confianza en la democracia.

Las entidades relacionadas con incidentes FIMI pueden ser de muchos tipos. En la figura 34 (EEAS, 2205) pueden verse diferentes tipos clasificados en cuatro grupos: canales estatales oficiales, canales controlados por el Estado, canales “ligados” a un Estado, y canales “alineados” con el Estado.

[66] En relación con las elecciones europeas, el EEAS detectó 42 casos de actividad FIMI rusa, que se intensificó en las semanas previas a la votación, alcanzando su punto máximo entre el 6 y el 9 de junio, y continuando mucho más allá de esa fecha.

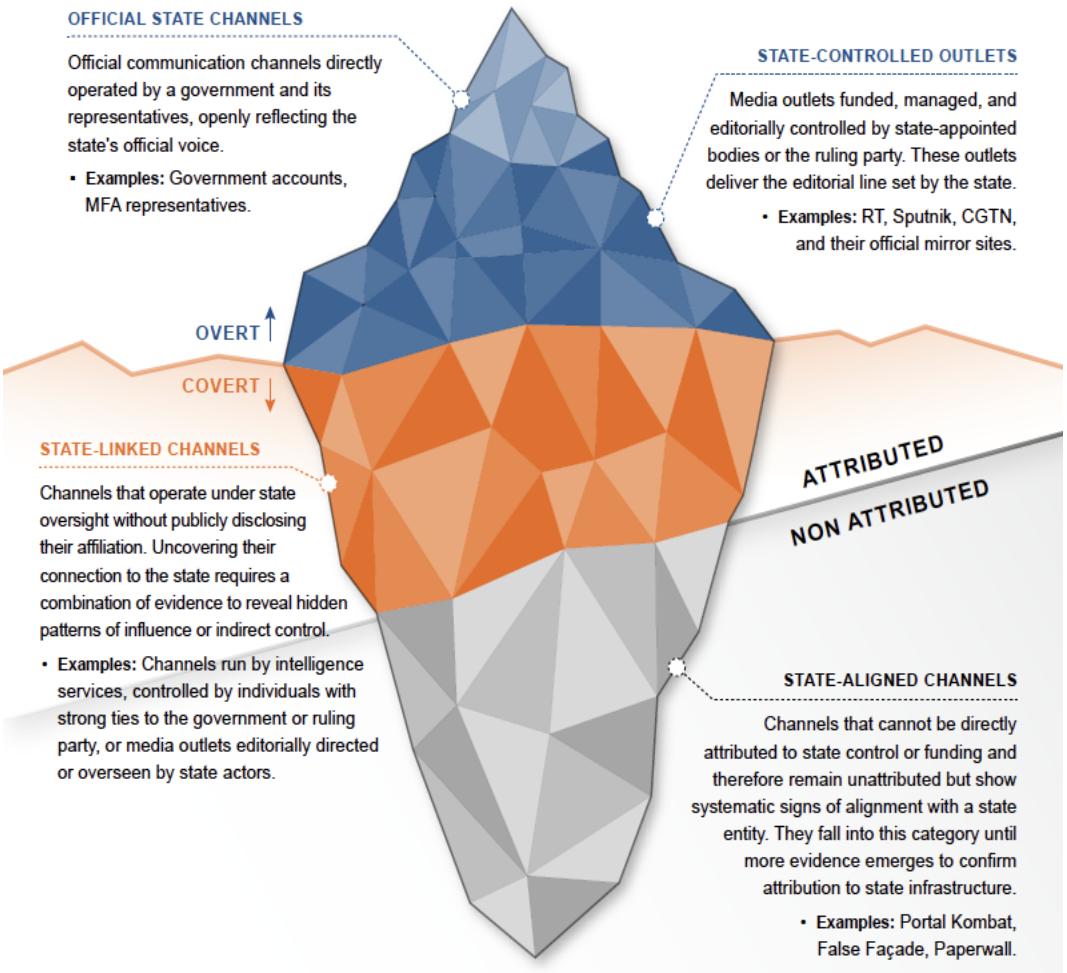


Figura 34. Tipos de entidades relacionados con incidentes FIMI. Fuente: EEAS (2025)

Obsérvese en el figura 34 que en unos casos se trata de **actividades no encubiertas** (caso de canales oficiales o controlados por los Estados) mientras que en otros casos pueden ser **actividades encubiertas**, aunque **no siempre atribuibles** de forma rotunda. Este es el caso de los canales alineados con el Estado, que no pueden atribuirse directamente al control o la financiación del Estado y, por lo tanto, no se atribuyen, pero que muestran signos sistemáticos de alineación con una entidad estatal. Entran en esta categoría hasta que surjan más pruebas que confirmen la atribución a la infraestructura estatal.

La figura 35, obtenida de EEAS (2025), indica las **entidades relacionadas con los gobiernos de China y Rusia** detectadas en 2024 **con actividad relacionada con FIMI**.

Figura 35. Entidades de China y Rusia relacionados con incidentes FIMI detectados en 2024. Fuente: EEAS (2025)

La **caja de herramientas FIMI** (*FIMI toolbox*), es decir, el conjunto de medidas puestas a disposición de los Estados miembros de la UE para defenderse de este tipo de incidentes de interferencia externa es muy amplia, aunque la experiencia de uso es aún limitada[67], y su impacto es muy variable. El EEAS ha elaborado también una **metodología de uso de la caja de herramientas,** que puede encontrarse en el informe de 2024 aplicada al caso de incidentes relacionados con elecciones. En conjunto, el objetivo es incrementar la resiliencia de la sociedad europea frente a la interferencia externa procedente de otros gobiernos.

En la figura 36 pueden verse las **herramientas existentes distribuidas en cuatro dominios:** 1) Acción exterior de la UE, 2) Conciencia situacional, 3) Construcción de la resiliencia, y 4) Disrupción y regulación.

[67] Focalizada en incidentes relacionados con elecciones en 2023.

Figura 36. Caja de herramientas FIMI. Fuente: EEAS (2024)

De la experiencia obtenida en este caso, puede verse en la figura 37 el tipo de **incidentes esperables** distribuido en cinco tipos, y su evolución en el tiempo desde mucho antes del día de las elecciones hasta después de ella. Obsérvese que muchos de los tipos de incidentes se prolongan durante todo el proceso.

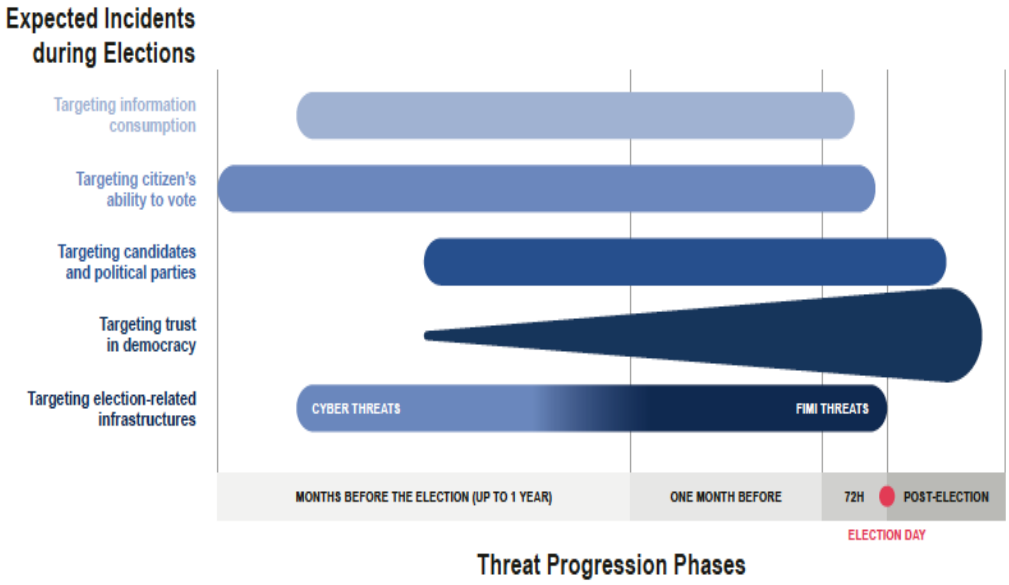


Figura 37. Fases de progresión de la amenaza. Fuente: EEAS (2024).

En definitiva, poco a poco todos los países se han ido dotando de estructuras, metodologías, herramientas tecnológicas y procedimientos para poder enfrentarse al incremento de tensiones en la guerra cognitiva. En la siguiente sección se analiza expresamente el **uso de la IA**.

3. Relevancia de la IA en la generación y lucha contra la desinformación en la guerra cognitiva

3.1. Evolución tecnológica de la guerra cognitiva

El **proceso de digitalización acelerado** desde finales del siglo pasado ha permitido disponer de herramientas y servicios digitales con las que la creación, difusión y procesamiento de información multimedia dirigida y compartida entre millones de personas sea factible a una velocidad, distancia, eficiencia, y personalización muy superior a lo que la sociedad podía imaginar hace una década.

Aunque la relevancia intencionada de la **"desinformación"** como herramienta conceptual ha sido empleada desde hace mucho tiempo en múltiples conflictos sociales, políticos y militares, es en las últimas dos décadas, con el proceso de digitalización de la sociedad en el que miles de millones de personas están interconectadas, cuando han surgido **diversas herramientas potenciadas por la IA** que permiten facilitar y sofisticar la generación de información falsa o tendenciosa, y, asimismo, la detección y eliminación, hasta cierto punto, de esta información.

Concretamente, el uso de herramientas de IA ha permitido **incrementar y automatizar la generación y difusión de campañas de manipulación y desinformación,** alcanzando un nivel más elevado de sofisticación e impacto potencial sobre la población objetivo para condicionar sus decisiones e influencia[68]. La **irrupción de la IA generativa**, integrada en multitud de productos y servicios, no ha hecho más que empoderar aún más estas posibilidades; aunque su uso haya permitido **difuminar la realidad y alterar la veracidad de la información con consecuencias sociales muy relevantes**.

No es extraño que, desde el punto de vista de su uso en defensa, las capacidades tecnológicas potenciadas por la IA se hayan aprovechado como factores "habilitadores" en la **evolución de la guerra cognitiva**; más aún cuando se han integrado en otros sistemas tecnológicos como las redes sociales y los sistemas de comunicaciones móviles de banda ancha.

Actualmente, el **dominio de las plataformas tecnológicas digitales potenciadas por la IA** forma parte de las estrategias de las grandes potencias tecnológicas como Estados Unidos, Rusia o China, y también de la UE, para **influir en la opinión pública** en todo el mundo a favor de objetivos estatales o de determinadas organizaciones públicas y privadas. A ello se destinan recursos crecientes, personal especializado y las herramientas más avanzadas disponibles, entre ellas las de IA.

El uso de la IA ha permitido la **manipulación creciente de la "sociedad"** (o de grupos concretos de la misma), sobre todo, aprovechando la enorme difusión y uso de las **redes sociales** y el impacto de la opinión de determinadas entidades y personas conocidas como

[68] https://www.ie.edu/university/news-events/news/deceptive-use-ai-technology-is-turning-mind-main-battlefield-21st-century-according-ie-university/

"influencers", reales o virtuales[69], con millones de *seguidores* en las mismas. Con el uso de herramientas de IA generativa multimodal, capaces de generar videos y subirlos a plataformas ampliamente utilizadas como *YouTube*, el impacto conseguido es mayor (Pelevina y Salojärvi, 2025).

Esta tendencia del **uso de "Influencers" especializados** también se ha manifestado en zonas de conflictos militares como en la guerra de Ucrania, en la que ***"Influencers de guerra"*** han usado *TikTok* como herramienta para difundir determinadas narrativas con sofisticación creciente. Divon y Kutrok (2025) han identificado tres **estilos principales** empleados por los influencers. *(1) narración POV, utilizando plantillas de punto de vista para conectar a las audiencias con testimonios personales de guerra; (2) contenido memético, yuxtaponiendo miras de guerra con memes de tendencia para mejorar el compromiso y la capacidad de compartir; y (3) política lúdica, que combina entretenimiento y comentarios políticos a través de funciones como LIVE para recaudar fondos y movilizar audiencias globales.*

Los videos generados por IA están socavando la realidad y facilitando que, hoy en día, prácticamente cualquier persona pueda generar información multimedia que puede parecer real. Concretamente, los **videos generados por herramientas de IA generativa parecen reales y son difíciles de distinguir**. La imagen de la figura 38 es un video generado artificialmente en el que una supuesta "*corresponsal de guerra*" (virtual) aparece comentando los efectos del ataque de Irán sobre Israel.

Figura 38. Video generado por herramientas de IA. Fuente: https://www.dw.com/en/fake-news-reporters-whats-actually-true/video-74060837

En todo caso, **aunque el video sea generado por una herramienta de IA, no significa necesariamente que el contenido de la información que contiene sea falso** (*fake news*). Esta valoración responde al análisis del contenido del mensaje, incluido que las imágenes que le acompañan respondan a la noticia transmitida y no al medio empleado para su difusión. En estos casos, sí debe existir un análisis de verificación de la información[70].

[69] Los "*influencers virtuales*" son personajes ficticios generados por ordenador a los que se les dota de determinada "personalidad" y mensajes que les hacen más atractivos a determinados grupos de personas. https://www.cyberclick.es/que-es/influencers-virtuales

[70] En el ejemplo de la figura 38, la periodista, su voz y el vídeo pueden ser generados por una herramienta de IA, pero las imágenes del fondo sí deben corresponderse realmente con el ataque de Irán sobre Israel al que se refiere la noticia, y no a un evento ocurrido en otro tiempo y lugar.

Desde un punto de vista complementario, **también el uso de herramientas de IA permite luchar contra la desinformación**. *RAND Corporation* ha identificado 82 herramientas de IA para luchar contra la desinformación[71]. El **análisis por personal especializado**, apoyado por herramientas tecnológicas de "verificación" de contenidos, permite, por ejemplo, determinar, en muchos casos, si una imagen o un vídeo es falso o no. En otros casos, no es sencillo llegar a una conclusión. Aunque es probable que las tasas de detección sean similares a las de otros medios, el impacto del **engaño mediante *deepfakes* de video** puede ser mucho mayor que el conseguido con el engaño verbal o textual, aunque el mensaje sea el mismo[72].

Un factor esencial para lograr el objetivo de una campaña determinada es el acceso y procesamiento de la **creciente cantidad de datos personales** generados al navegar por las redes sociales, comprar en línea o utilizar chatbots de IA como *ChatGPT*, *Perplexity,* u otros muchos. Ya sea de forma consciente o no, el uso de servicios digitales proporciona **información** detallada sobre nuestras preferencias de consumo, comportamientos sociales y estados emocionales, lo que permite a las empresas proveedoras de servicios digitales elaborar **perfiles precisos y facilitar la microsegmentación** de los usuarios **para diversos fines.**

En el caso de la UE, la Ley de IA de la UE incide profusamente en preservar los derechos de los ciudadanos, tanto en la defensa de sus datos como en la limitación a las prácticas que se pueden desarrollar a partir de ellos, que pueden afectar al interés general o los principios fundacionales de los ciudadanos de la UE.

Desde el punto de vista de **interés político**, el procesamiento de esta información personal también permite **analizar o introducir sesgos** hacia determinadas ideologías, partidos políticos, miembros del gobierno, representantes o candidatos en elecciones; problema cuya relevancia ha crecido como resultado de la evolución de la tecnología.

La irrupción de la **IA generativa**, combinada con el aprendizaje automático y su capacidad de "*aprender*" de la interacción con sus interlocutores, permite **crear contenido altamente realista**, con el que llegar de manera personalizada a individuos dentro de un público objetivo de manera eficiente y efectiva mediante el uso de indicaciones en el momento adecuado, con **poca intervención humana**, logrando una **mayor verosimilitud del mensaje**. Específicamente, ha permitido crear automáticamente **campañas de desinformación con realidades sintéticas** difíciles de contrarrestar. De hecho, con las nuevas herramientas de **IA generativa multimodal** es difícil conocer si una información (texto, datos, video, imágenes, etc.) es real o ha sido creada de forma sintética, a no ser que se empleen herramientas especializadas.

Hay muchos tipos de uso de **imágenes falsas** creadas con diversos objetivos más allá de intereses en los conflictos bélicos. Veamos algunos ejemplos:

- Intereses económicos. El 22 de mayo de 2023, se difundió rápidamente la noticia de un **ataque al Pentágono,** junto con una imagen a través de un Twitter verificado (asociado con Bloomberg News), con reacción en mercados bursátiles y de

[71] https://www.rand.org/research/projects/truth-decay/fighting-disinformation/search.html

[72] El dominio de las señales visuales en la percepción humana se conoce como el efecto de dominancia visual de Colavita. https://en.wikipedia.org/wiki/Colavita_visual_dominance_effect

inversión en los Estados Unidos. Sin embargo, el usuario que compartió la historia era falso, la historia fue inventada y es muy probable que la imagen sea generada por una herramienta de imágenes de IA[73].

- Ciberdelincuencia. Una tendencia observada recientemente es el uso de "*aplicaciones nudificadoras*" que emplean la IA para **crear pornografía no consentida a partir de imágenes de las víctimas**. Las imágenes y videos generados son utilizados para la extorsión, abuso y acoso de las víctimas y, así, obtener un beneficio económico[74].

Cuando estas imágenes se refieren o se relacionan con **conflictos bélicos**, las falsificaciones se diseñan expresamente para provocar una **fuerte reacción emocional** en su receptor (p.ej. al incluir los cuerpos de bebés, niños o familias). En los sangrientos primeros días de la guerra, los partidarios de *Israel* y *Hamas* alegaron que el otro bando había victimizado a niños y bebés, como se indicó en la figura 32 anterior de 2024, pero hay muchas más imágenes, como las incluidas en la figura 39, que han generado **polémica respecto a su autenticidad**.

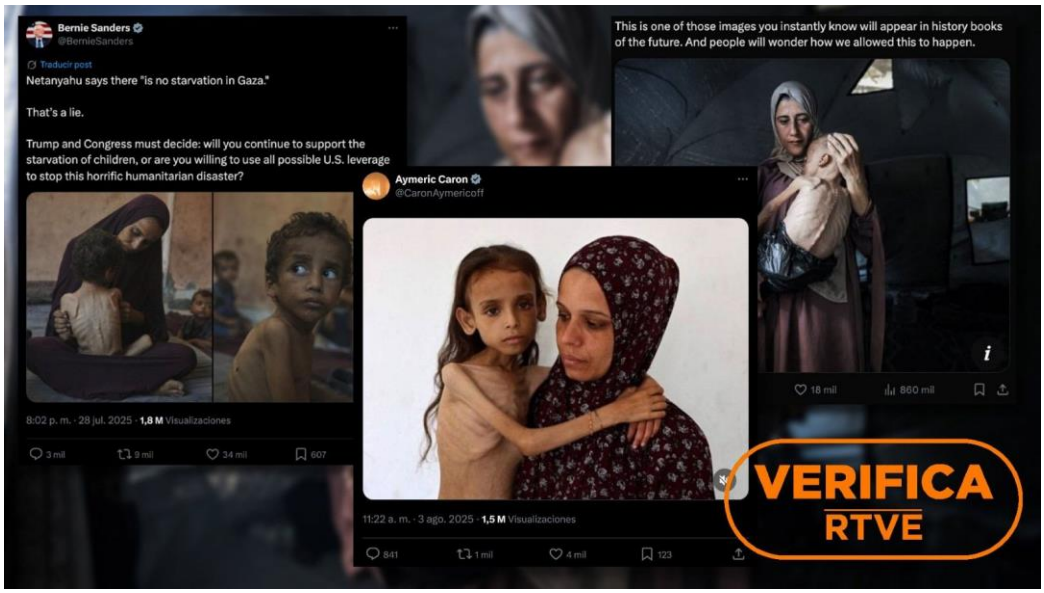


Figura 39. ¿Imágenes manipuladas por IA? Fuente: https://www.rtve.es/noticias/20250808/mientes-grok-imagenes-ninos-desnutridos-gaza/16688487.shtml

El elemento relevante y perturbador es que la herramienta de IA ***Grok,*** de la red social *X,* consideraba que se trataba de imágenes falsas[75], mientras que la empleada por RTVE (*VerificaRTVE*), indicaba que eran ciertas. Lo que sí ocurre es que algunas de ellas pueden ser "*escenificadas*" en su toma por parte de los medios de comunicación. En definitiva,

[73] https://www.riskintelligence.eu/background-and-guides/ai-tools-in-hybrid-warfare-a-double-edged-sword

[74] https://www.trendmicro.com/en_us/research/24/f/behind-the-great-wall-void-arachne-targets-chinese-speaking-user.html

[75] *Grok* indicaba que correspondían a la hambruna de Yemen en 2017. https://www.rtve.es/noticias/20250808/mientes-grok-imagenes-ninos-desnutridos-gaza/16688487.shtml

con estas controversias sobre la realidad de una imagen se pretende crear dudas en los usuarios y alterar o contrarrestar las narrativas oficiales.

La **IA generativa multimodal** (clonando no solo la imagen, sino también la voz en videos hiperrealistas) se combina con otro tipo de herramientas como la **realidad virtual** (con la generación de modelos realistas 3D) para permitir **ataques cognitivos** basados en un área creciente como es la ***manipulación emocional***. El desarrollo tecnológico de estas herramientas es muy rápido y la frontera entre ficción y realidad se hace más borrosa.

Los **entornos de realidad virtual** (*Virtual Reality Environment, VRE*), como el *Metaverso* de la empresa *Meta* pueden usarse con ese fin; experimentos y estudios de casos de psicoterapia han demostrado la capacidad de los VRE para inducir emociones negativas como el miedo o la ansiedad, haciendo que una determinada persona esté más predispuesta o susceptible a aceptar falsos rumores o teorías de conspiración (Fenstermacher et al., 2023). Además, la **combinación de datos biométricos recopilados de los VRE** podría proporcionar una imagen completa del estado emocional de un usuario y permitir el seguimiento de cómo actúan, interactúan y reaccionan los usuarios a un nivel detallado, lo que respalda la posible manipulación de los usuarios[76]. El posible uso militar para interpretar estos datos, si se *hackeasen*, en relación con la alteración del nivel de moral de combate, por ejemplo, es posible.

En la edición de la feria CES de 2025 se ha dado un paso más con la presentación por parte de la empresa coreana *LG Electronics* de su visión de **"Inteligencia Afectuosa"** *(Affectionate Intelligence)*, con la que las tecnologías de IA se integran de manera más humana y personalizada en la vida de los usuarios mediante sistemas de IA capaces de generar conexiones más humanas, **adaptándose a las emociones, necesidades y preferencias de los usuarios**[77]. La figura 40 permite ver esta visión en un contexto doméstico[78].

Su uso en **aplicaciones duales** podría ser relevante en la medida en la que dispositivos inteligentes (no solo electrodomésticos) aprendan de los hábitos del usuario para ofrecer una **rutina más eficiente y personalizada adaptada a entornos conflictivos**.

[76] Entre estos datos biométricos puede citarse la respuesta galvánica de la piel (intensidad de las emociones del usuario), la información de los músculos faciales (revela los sentimientos), la tensión muscular mide (detecta microexpresiones extremadamente rápidas, casi imposibles de falsificar), el seguimiento ocular detecta lo que el sujeto destaca visualmente y la electroencefalografía mide el nivel de atención de un usuario (Fenstermacher et al., 2023).

[77] Con la plataforma *LG ThinQ*, la empresa ha desarrollado soluciones que aprenden y evolucionan en función del estilo de vida de cada persona, generando una conexión más natural entre el usuario y la tecnología. Con ello, LG pretende redefinir la interacción entre los humanos y los dispositivos inteligentes https://global.techradar.com/es-mx/computing/artificial-intelligence/affectionate-intelligence-la-tecnologia-que-entiende-y-se-adapta-a-ti

[78] Un ejemplo destacado fue la integración de la tecnología de *"inteligencia afectuosa"* en los televisores OLED de próxima generación de LG, que ofrecen recomendaciones de contenido basadas en el estado de ánimo detectado por sensores biométricos y cámaras especializadas.

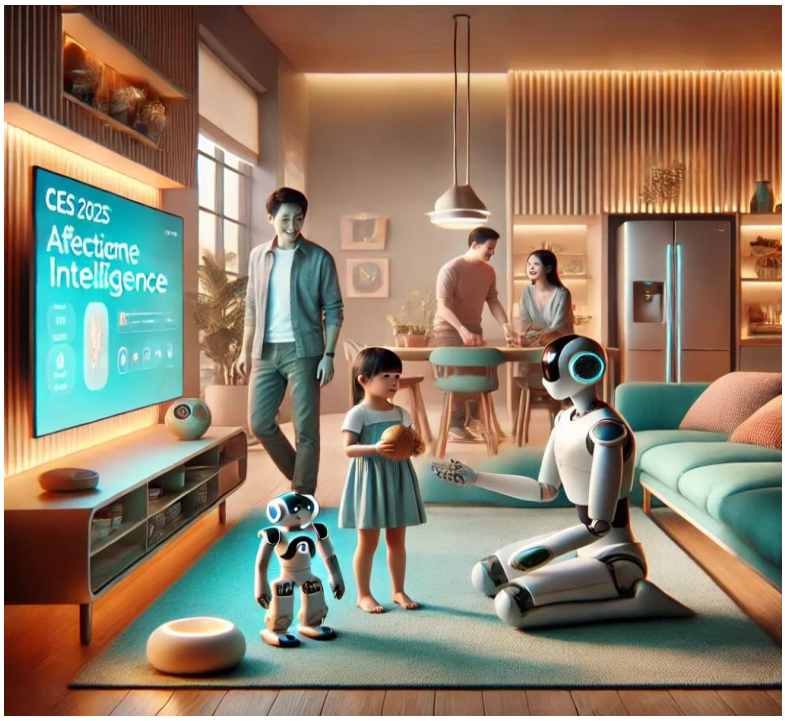


Figura 40. Visión de la "inteligencia afectuosa" de LG. Fuente: https://www.altaidigital.co/post/lg-nos-muestra-su-affectionate-intelligence

Un caso reciente de la *agencia de noticias EFE, en el que se ha* empleado una herramienta basada en IA, denominada ***EFE Verifica*****,** se refiere a rebatir la noticia de que el helicóptero militar siniestrado en Washington, el 29 de enero de 2025, en un choque con un avión de pasajeros, no iba pilotado por una **militar transgénero** a la que se acusaba en redes sociales (identificada con nombre y foto)[79]. La noticia era falsa y la mujer cuyas imágenes circulan en redes está viva y no tiene relación alguna con el accidente, tal y como lo ha confirmado ella misma en un vídeo colgado en *Facebook* (véase figura 41).

El trasfondo político de la noticia falsa reside en el apoyo a la decisión del Presidente de Estados Unidos, Donald Trump, que firmó el 28 de enero de 2025, un día antes del accidente, una orden ejecutiva que busca **restringir fuertemente el servicio militar de personas transgénero**[80]. Forma parte, por tanto, de una campaña con claras resonancias políticas.

[79] Perfiles en X y Facebook afirman que la tercera ocupante del helicóptero accidentado en Washington era Jo Ellis, una mujer transgénero con más de 15 años de experiencia como piloto de aeronaves militares, quien, según los mensajes, habría fallecido https://verifica.efe.com/mujer-trans-piloto-helicoptero-militar-accidente-washington-falso/

[80] La medida prohíbe a quienes expresen una identidad de género distinta a su sexo asignado al nacer alistarse en las Fuerzas Armadas, justificando que no cumplen con los "estrictos estándares" físicos y mentales requeridos para el servicio. https://es.euronews.com/2025/01/28/trump-firma-una-orden-para-revisar-la-politica-del-pentagono-sobre-las-tropas-transgenero?utm_source=perplexity

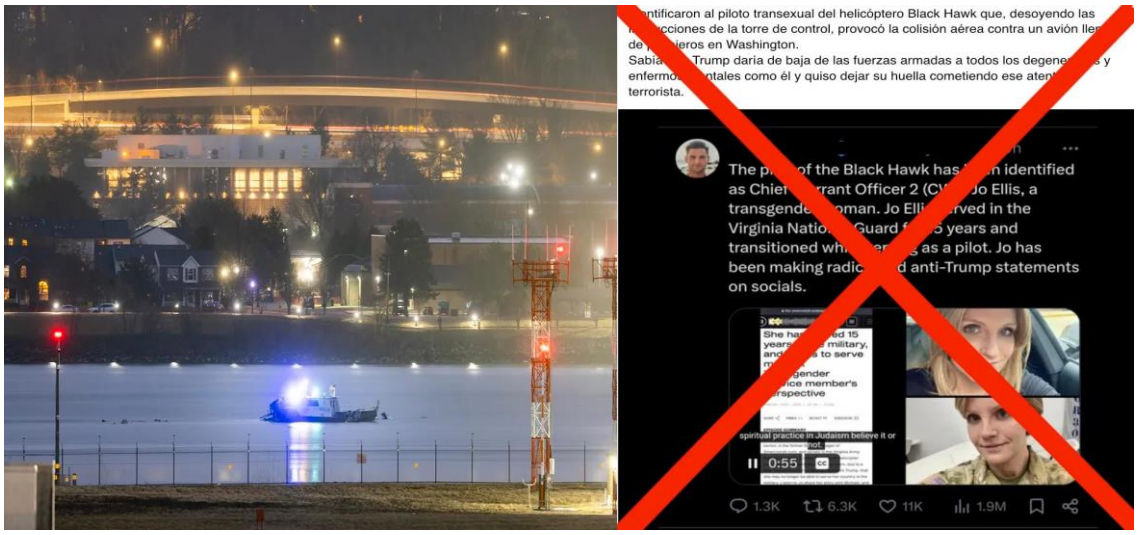


Figura 41. Noticia falsa relacionada con el accidente de helicóptero en enero de 2025 en Estados Unidos. Fuente: https://verifica.efe.com/mujer-trans-piloto-helicoptero-militar-accidente-washington-falso/

Existen diferencias significativas en la forma y el volumen de casos en los que noticias falsas (*fake news*) afectan a diferentes países y a los sectores más proclives. La figura 42, procedente de un estudio de *Sopra Steria*[81] (Courtois, 2024), resume la situación con datos de **crecimiento de noticias falsas de 2022 a 2023.**

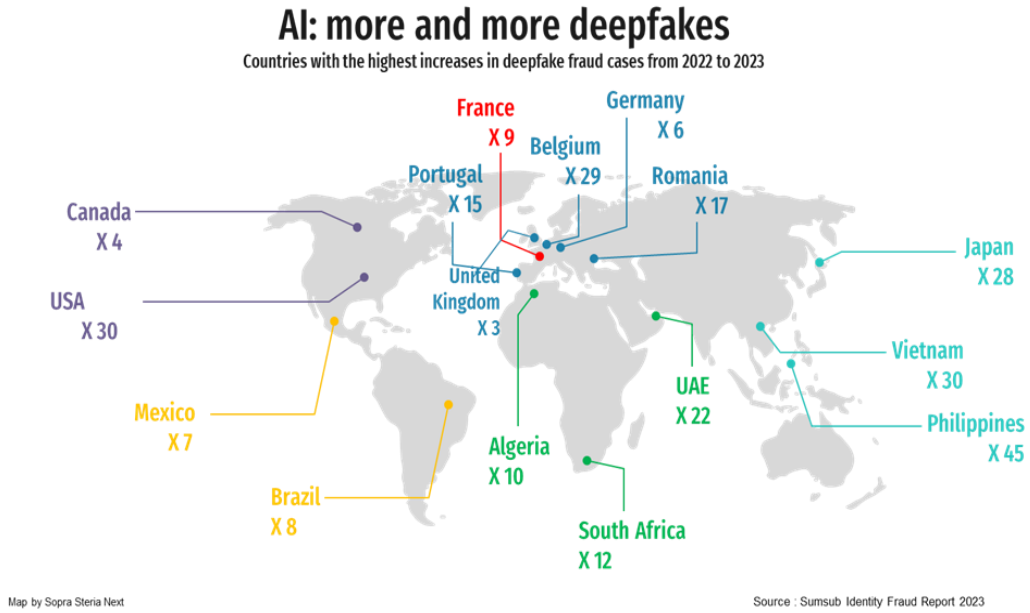


Figura 42. Países con un crecimiento mayor de noticias falsas (2022-2023). Fuente: https://www.soprasteria.com/insights/details/combatting-disinformation-the-ai-war

El crecimiento espectacular desde 2022 coincide con la difusión del uso de herramientas de IA generativa y la creciente implicación de unidades gubernamentales especializadas en la generación y difusión de narrativas intencionadas. También sus efectos varían en función de las decisiones tomadas en los países para reaccionar a ellas. Concretamente,

[81] https://www.soprasteria.com/insights/details/combatting-disinformation-the-ai-war

las campañas de desinformación a gran escala constituyen un reto importante para Europa y requieren una respuesta coordinada de los países de la UE, las instituciones de la UE, las plataformas en línea, los medios de comunicación y los ciudadanos de la UE[82].

El sector de los **medios de comunicaciones digitales** es el que recibe un porcentaje más elevado de noticias falsas (el 4,27%). Con más detalle, en la figura 43 (Entrust, 2025) se pueden ver los casos de fraude de identidad más relevantes por sectores. Con una clara diferencia, es el sector de ***criptomonedas*** el más afectado.

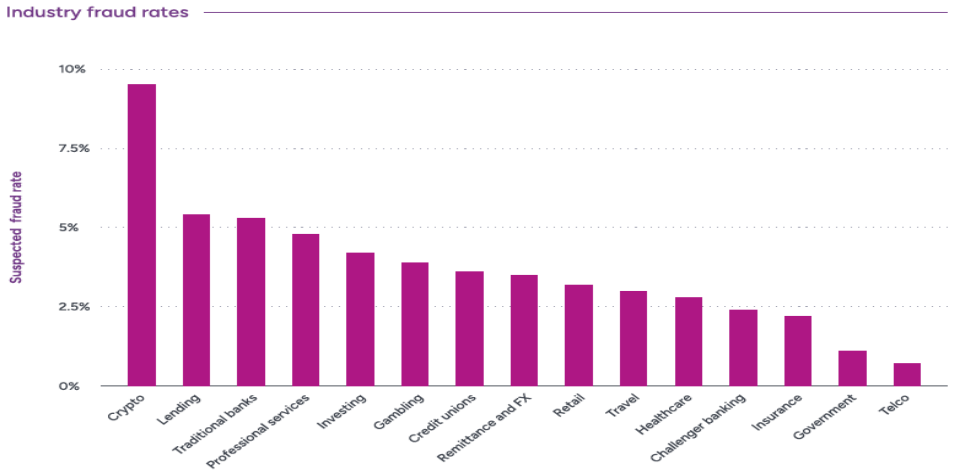


Figura 43. Sectores industriales más afectados por fraudes de identidad. Fuente: Entrust (2025)

El potencial efecto de las noticias falsas para crear situaciones de caos en la sociedad ha sido pronosticado muchas veces como posible herramienta de la guerra cognitiva en diferentes supuestos, un ejemplo típico es la creación de **colapso económico injustificado** como la señalada por Marma (2025).

> *Situación en la que, en unas pocas semanas, millones de personas en un país en desarrollo pierden repentinamente la fe en su sistema bancario. Todo comienza con noticias falsas, publicaciones virales en las redes sociales e informes financieros generados por IA que son tan convincentes que parecen reales. Los informes predicen un desastre económico inminente, lo que provoca pánico. La gente comienza a apresurarse a retirar su dinero, temiendo lo peor. Pero el colapso del sistema bancario no fue causado por ningún problema económico real, sino el resultado de una operación psicológica cuidadosamente planificada (PSYOP), destinada a socavar la confianza en las instituciones del país.*

Las consecuencias de las "*fake news*" no se limitan ataques a países o instituciones, como parte de una guerra cognitiva a gran escala, también afectan individualmente a personas. Wilbor (2025) detalla un **caso ficticio** relevante por la persona afectada, por la tecnología empleada, por el país implicado, y por la finalidad: la neutralización de un objetivo militar.

> *Un militar recibe un video a través de Facebook en el que aparece su esposa manteniendo relaciones sexuales con otro militar. Al recibir el video, se enfurece, va a su casa, y golpea a su esposa hasta matarla, y se suicida. Éxito del sistema de IA empleado (en la historia por China) al neutralizar a un costo insignificante a un militar formado. El sistema de IA analizó el historial de redes sociales del militar y su esposa y determinó que era susceptible de responder con*

[82] En https://digital-strategy.ec.europa.eu/es/policies/online-disinformation se pueden ver las distintas iniciativas lanzadas por la Comisión Europea.

violencia ante un video falso de infidelidad matrimonial. La IA recopiló fotos de las redes sociales de la esposa y otra persona y generó un video de carácter sexual.

El problema es que, desde el punto de vista tecnológico, la situación descrita podría haber sido real. El ejemplo demuestra cómo **métodos emergentes de amenazas cognitivas potenciados por la IA pueden emplearse para explotar las vulnerabilidades de personal militar**. Si estas técnicas se aplicasen de forma sincronizada a varias personas de una misma unidad, y no a individuos aislados, se correría el riesgo de interferir y degradar cognitivamente las capacidades de una unidad militar: esa evolución nos conduce al dominio de las **neuroarmas** que se tratarán posteriormente.

La tentación de **actuar sobre la regulación** para evitar estos casos de abuso de la buena intención es evidente. Sin la existencia de **robustas salvaguardas legislativas** los sistemas de IA pueden hacer que la censura, vigilancia y expansión de la desinformación sea más rápida, barata y efectiva. La forma de llevarlo a cabo varía ampliamente de unos países a otros. Sin embargo, no existe un consenso en su aplicación y, en la práctica, se aborda de forma muy diferente en la legislación de los países en función de la ideología del gobierno, la extensión y conocimiento de la población sobre estas técnicas, y la posición de la sociedad en la aceptación de medidas coercitivas.

La respuesta legislativa cambia significativamente de un país a otro. Los casos de la UE limitando el uso de cámaras para reconocimiento facial en sus regulaciones y el de China, permitiéndolos en gran medida, son respuestas legislativas muy diferentes en las que asegurar la privacidad es un valor básico, por encima de la seguridad (caso de la UE) o al revés (caso de China).

En la figura 44 se indican las **áreas y medios técnicos empleados actualmente** para controlar la información al/del ciudadano.

- **Privacidad**. Sistemas de vigilancia basados en *big data* agregan y analizan cantidades masivas de datos personales sensibles que alimentan el entrenamiento y uso de modelos de IA.
- **Libre expresión**. Despliegue de sistemas automatizados para censurar información política, social y religiosa de acuerdo con diversos parámetros definidos por los gobiernos lo que, a su vez, incentiva la autocensura[83].
- **Acceso a información**. Plataformas digitales que promueven contenido incendiario sobre información fiable empleando IA generativa
- **Proceso**. Herramientas de vigilancia potenciadas por la IA monitorizan los medios de comunicación buscando causas de infracción de la ley.
- **No discriminación**. Los datos de entrenamiento de algoritmos de IA pueden perpetuar sesgos y exacerbar la discriminación frente a minorías de forma intencionada.
- **Asociación y asamblea**. Sistemas de IA de reconocimiento facial pueden identificar y trazar manifestantes permitiendo actuar a las fuerzas policiales contra ellos.

[83] El proceso coercitivo que promueve la *autocensura* (limitación o censura que se impone uno a sí mismo según la RAE) ha sido bien estudiado en el siglo XX sin necesidad de emplear herramientas informatizadas ni IA (Casamadrid, 2025).

Figura 44. Áreas de actuación para el control de la información. Fuente: https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence#generative-ai-supercharges-disinformation

El uso de **herramientas de IA** ofrece mejoras a diferentes actores para conseguir el efecto deseado. Con un poco más de detalle:

- El contenido que se publica y comparte se puede analizar mediante técnicas de IA como el **procesamiento del lenguaje natural (NLP)**.
- Varias herramientas de clasificación, como los bosques de decisión y las redes neuronales LSTM, facilitan la **traducción automática**.
- La **tecnología de grafos** aprovecha el potencial de la IA para analizar las relaciones entre puntos de datos.
- Los sistemas de aprendizaje automático pueden crear herramientas para **detectar imágenes que han sido manipuladas o manipuladas**, por ejemplo, buscando rastros dejados por los sistemas utilizados para capturar imágenes alteradas por algoritmos de retoque.
- Las aplicaciones de IA también pueden ser entrenadas para **detectar desinformación en las redes sociales** y emitir advertencias. Al analizar bloques de datos procedentes de intercambios en redes como *Twitter* y *Telegram*, las IA pueden reconocer elementos estilísticos típicos de las noticias falsas.
- También pueden ser entrenados para **identificar contenido potencialmente problemático** y ayudar a los operadores a comprender por qué se ha marcado.

El desarrollo de la tecnología de IA y relacionada con ella va a hacer que tanto las herramientas de desinformación como las que luchan contra ella evolucionen muy rápidamente incrementando su sofisticación. En el futuro, la combinación de estas herramientas con el uso de la **realidad aumentada** (RA), la **realidad virtual** (RV), las **interfaces cerebro ordenador** (BCI), y el análisis de los datos neuronales que genera su uso, también pueden permitir a actores maliciosos **hackear la "realidad" que nos rodea**,

nuestros estados de ánimo y reacciones, y con ello dar forma a nuestro comportamiento en las direcciones buscadas.

Del informe del *Centro Criptológico Nacional* (CCN) de 2024 se ha adaptado la figura 45 en la que se indican diferentes **factores de mejora derivadas del uso de herramientas de IA.**

Creciente interacción

El uso de herramientas de IA ofrece mejoras a:	Actores estatales altamente capacitados	Actores estatales capaces, empresas comerciales que venden a Estados, grupos de ciberdelincuencia organizada	Hackers a sueldo menos cualificados, ciberdelincuentes oportunistas, hacktivistas
Intención	Alta (estrategias ligadas a posiciones geopolíticas)	Alta (estrategias ligadas a intereses estatales y corporativos)	Media (estrategia oportunista)
Capacidad	Altamente cualificados en IA, ciberseguridad y redes, bien dotados de recursos	Disponen de personal cualificado en IA, ciberseguridad y redes, con limitaciones de recursos	Disponen de personal con conocimientos reducidos individuales, con recursos limitados
Reconocimiento	Mejora mínima (objetivo secundario)	Mejora moderada (variabilidad dependiendo del tipo de actor)	Mejora progresiva (variabilidad dependiendo del tipo de actor)
Ingeniería social, phishing,	Mejora continua con el desarrollo de herramientas	Mejora continua con el desarrollo y adaptación de herramientas de IA	Mejora significativa en el uso de herramientas de IA
Herramientas (malware, exploits)	Mejora constatada y en rápido desarrollo	Mejora continua con el desarrollo y uso de herramientas de IA	Mejora continua con el uso de herramientas de IA
Movimiento lateral	Mejora potencial mínima	Mejora potencial mínima	Ninguna mejora apreciable
Implicaciones	Son los actores mejor situados para aprovechar el potencial de la IA en operaciones cibernéticas contra redes con el uso de malware avanzado	Mejora de las capacidades de reconocimiento, ingeniería social y exfiltración. Proliferarán herramientas de IA entre los ciber agentes principiantes	Reducción de la barrera de entrada para operaciones de acceso efectivas y escalables: aumento del volumen de dispositivos y cuentas comprometidas con éxito

Figura 45. Fuente: adaptada de https://www.ccn-cert.cni.es/es/informes/informes-ccn-cert-publicos/7274-ccn-cert-ia-04-24-ciberamenazas-y-tendencias-edicion-2024/file.html

Es interesante analizar cómo su uso y capacidades para obtener beneficios de la IA **depende del actor implicado**:

- Actores estatales altamente capacitados como unidades especializadas de los gobiernos.
- Actores estatales capaces, empresas comerciales que venden a Estados, grupos de ciberdelincuencia organizada.
- *Hackers* a sueldo menos cualificados, ciberdelincuentes oportunistas, *hacktivistas*.

Por último, se debe mencionar la **Ley de IA de la Unión Europea** que también contempla restringir la actividad de creación de noticias falsas al exigir a quienes utilizan sistemas de IA para generar o manipular contenidos (ya sean imágenes, audio o vídeo) la *obligatoriedad de indicar si esos contenidos han sido creados o modificados artificialmente*.

Sin embargo, **confiar únicamente en la legislación no permite protegerse completamente contra las campañas de desinformación llevadas a cabo por otros países competidores**, por lo que la lucha contra la desinformación debe librarse utilizando las mismas "armas" que los adversarios: 1) tecnologías avanzadas impulsadas

por la IA, y 2) valores sólidos que refuercen la soberanía nacional; todo ello, acompañado de una formación del ciudadano en herramientas digitales muy superior a la actual.

Los marcos legales actuales, si bien son esenciales, están **fragmentados** y pueden impedir una coordinación rápida en el ámbito de la guerra cognitiva[84]. El marco legal debe evolucionar para que estas operaciones puedan ser más ágiles, permitiendo esfuerzos conjuntos en los que la experiencia de inteligencia se complemente con la infraestructura y las capacidades operativas de Defensa, especialmente en escenarios de alto riesgo.

La guerra cognitiva establece un **campo de batalla compartido entre diferentes unidades gubernamentales**. A medida que se difuminan las líneas entre el conflicto militar tradicional y las operaciones de información, es esencial una colaboración más estrecha entre los ministerios implicados (p.ej. en España, al menos Presidencia, y los ministerios de Defensa, Interior y Exteriores), los diversos órganos de Inteligencia dependientes de ellos, y órganos similares en la cooperación internacional (p.ej. en la UE y en la OTAN).

3.2. Deliberación y desinformación generada por la IA

3.2.1. Impacto de la IA en la generación de desinformación

En una sociedad en la que la construcción y difusión de narrativas ha adquirido gran importancia para imponer posiciones interesadas de grupos de presión y de gobiernos, la forma en la que **la narrativa alimenta el proceso de deliberación en la sociedad y realimenta las posiciones de partida** ha adquirido gran importancia.

En un **modelo ideal** en el que la información empleada para la construcción de la narrativa sobre un determinado tema es **veraz**, el objetivo es llegar a un **consenso** en base a la integración de argumentos individuales destilados y enriquecidos con **debates** transparentes que permitan llegar a **compromisos entre diferentes posiciones**, y que los compromisos alcanzados sean **aceptados mayoritariamente por la población** como base para la toma de decisiones de los responsables políticos (véase figura 46).

El problema surge cuando este proceso deliberativo se **contamina** porque a los argumentos individuales de las personas que participan en el debate se unen **acciones de desinformación que influyen en la discusión** y la conformación del consenso. El problema, además, es que la "veracidad" de la información no es fácil de determinar *a priori*, y se corre el riesgo de que el debate trate la desinformación de la misma manera que la información veraz.

[84] Como ejemplo de fragmentación competencial, en Estados Unidos la división entre el Título 10, que rige las operaciones militares, y el Título 50, que rige las acciones de inteligencia encubiertas, significa que el Departamento de Defensa y la Comunidad de Inteligencia a menudo operan en esferas separadas. La *Comunidad de Inteligencia* (basada en la experiencia de agencias como la CIA y la NSA) ya es el líder de las operaciones de influencia encubiertas. El Departamento de Defensa, por otro lado, desempeña un papel fundamental en la guerra convencional y en la seguridad de sus propios sistemas. https://www.linkedin.com/pulse/ai-define-cognitive-warfare-enrique-de-la-torre-nz3ne/

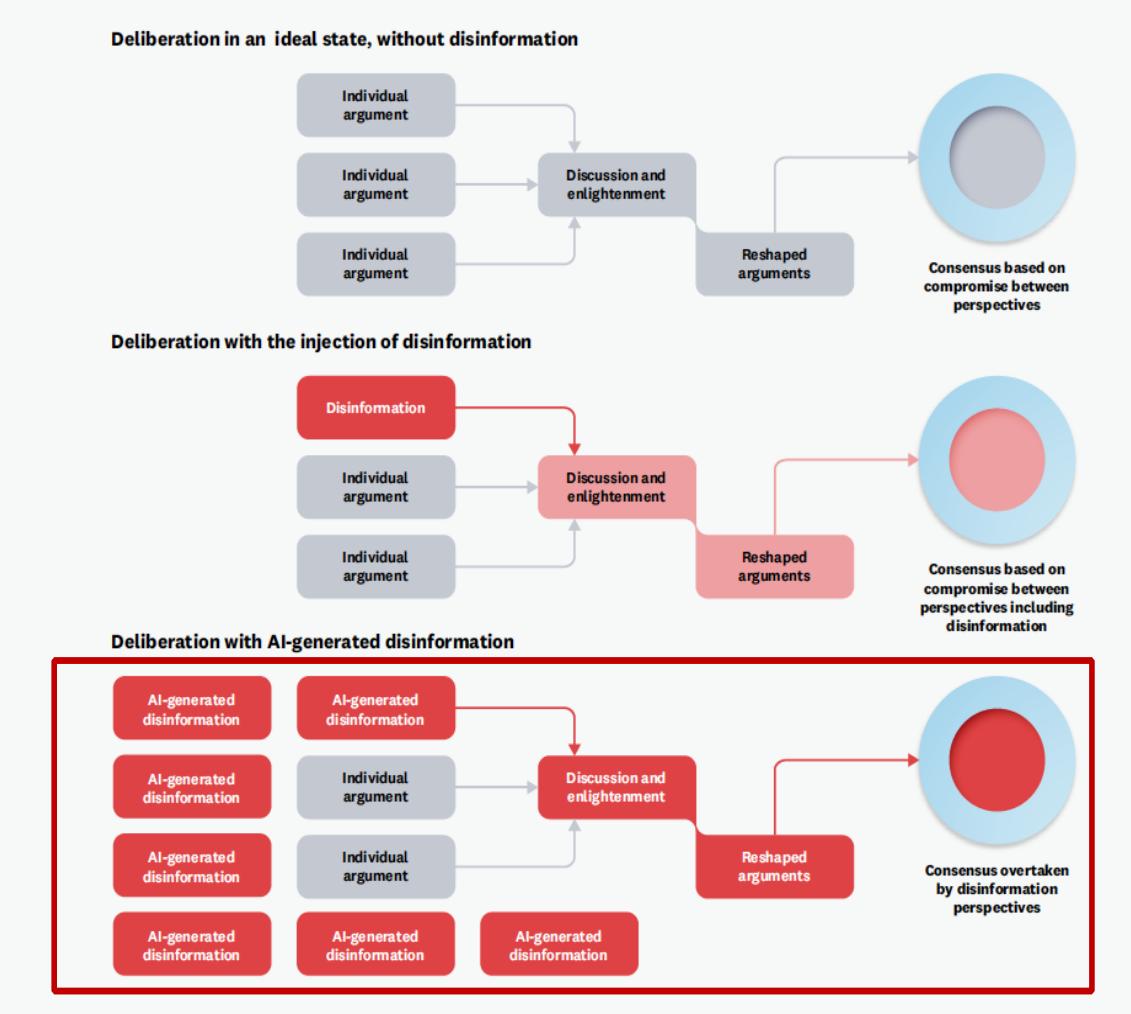


Figura 46. Deliberación y desinformación con IA. Fuente: https://collections.unu.edu/eserv/UNU:9217/artificial_Intelligence_powered_disinformation.pdf

Al final, casi de forma imperceptible, el consenso final no solo está basado en los compromisos alcanzados entre los participantes en el debate, sino que incluyen elementos de desinformación (p.ej. propaganda) sin que se sepa discernir la realidad. Se trata de un **problema que se incrementa si la formación o la predisposición de los individuos receptores potenciales de esa (des)información no es la adecuada para poder tener una base sólida de contraste**. De hecho, algunos usuarios son capaces de resistirse a ser "persuadidos" debido a sus habilidades racionales, pero otros son más susceptibles[85].

Esta situación no es nueva, y se ha producido en la historia desde el siglo XIX, primero con la prensa, para un porcentaje reducido de la población, y luego con los medios de comunicaciones de masas electrónicos no interactivos (analógicos) como la radio y la TV. Posteriormente, como proceso ligado a la transición hacia la digitalización de la sociedad, ampliado con cierta interacción (asimétrica) aprovechando servicios de internet, como

[85] Las vulnerabilidades de las personas pueden provenir de factores fisiológicos inherentes, como la edad o la demografía, factores psicológicos, como emociones o prejuicios, o factores ambientales, como el acceso limitado a una diversidad de información validada, que impiden la capacidad de una persona para tomar decisiones informadas (Impiombato et al., 2024).

sucede en la información almacenada en sitios web, *blogs* o redes sociales (Ramírez, 2023). En algunos casos, **generando información incorrecta de forma deliberada o no**[86].

Hasta muy recientemente, la generación de la información manipulada o incorrecta (desinformación) y su combinación en el debate social se realizaba **manualmente**, con lo que la elaboración de una narrativa compleja implicaba tiempo y recursos humanos especializados en su construcción y alineamiento con intereses políticos y socioeconómicos.

A medida en que la tecnología digital se desarrollaba, la tendencia era orientar el proceso de difusión hacia colectivos predeterminados, inicialmente muy amplios y, progresivamente, más focalizados, hasta permitir, en base al análisis complejo de enormes volúmenes de información, **"personalizar" el mensaje y el debate para parecer exprofeso para una persona concreta**. En esta nueva fase, el **factor habilitador** ha procedido de la IA en la que los argumentos individuales se ven enmascarados por elementos de **desinformación generados automáticamente por herramientas de IA**.

El **consenso final**, como se indica en la figura 46, está mediatizado por un factor de desinformación con relevancia variable. Esa condición puede tener consecuencias en los procesos finales de toma de decisión.

El uso de las herramientas de IA en el proceso de construcción y difusión de una determinada narrativa puede enmarcarse en los siguientes pasos:

- **Fase 1. Elección de un tema de interés** (político o social) en una sociedad que justifique el desarrollo de una campaña de desinformación basada en una "narrativa" diseñada expresamente para ello. En este paso, la elección del tema puede estar basado en el uso de técnicas de *big data* combinadas con datos históricos y encuestas de opinión, aprovechando eventos recientes de interés en un país, región o grupo de opinión que haya alcanzado suficiente notoriedad y se alineen con objetivos políticos.
- **Fase 2. Generación de información susceptible de crear opinión en redes sociales**. Existen diversas estrategias y herramientas de IA para facilitarlo o lograrlo. Algunas de ellas son: creación automática de cuentas falsas en números significativos, generación de contenido multimedia "creíble" multilingüe, difusión del contenido y generación de respuestas en cadena, uso de cuentas falsas, "*influencers*", repercusión en medios convencionales, y monitorización de los temas relevantes en un momento dado para la audiencia ("*trend topics*") así como su difusión geográfica y persistencia en el tiempo; muchos de ellos son muy volátiles y persisten únicamente durante pocas horas.
- **Fase 3. Provocación de respuesta política** que refuerce o contradiga la narrativa creada. Esta respuesta puede ir dirigida a gobiernos (amigos y enemigos), organizaciones multilaterales con posiciones oficiales, y grupos de oposición política en determinados territorios.
- **Fase 4. Impacto físico en el terreno** por parte de la población con manifestaciones de repulsa o atentados terroristas.

Es en la fase 2 en la que la IA, sobre todo la *IA generativa*, adquiere su mayor importancia. En la figura 47 puede verse esta relevancia en términos de **argumentación para el uso de**

[86] Este fenómeno ha implicado la emergencia de personas con conocimientos especializados en el debate como son los "articulistas", "comentaristas" e "influencers", en redes sociales.

la IA y su efecto: 1) Incremento del volumen de desinformación, 2) Incremento de la calidad de la desinformación, 3) Incremento de la personalización de la desinformación, y 4) Generación involuntaria de información falsa, pero plausible.

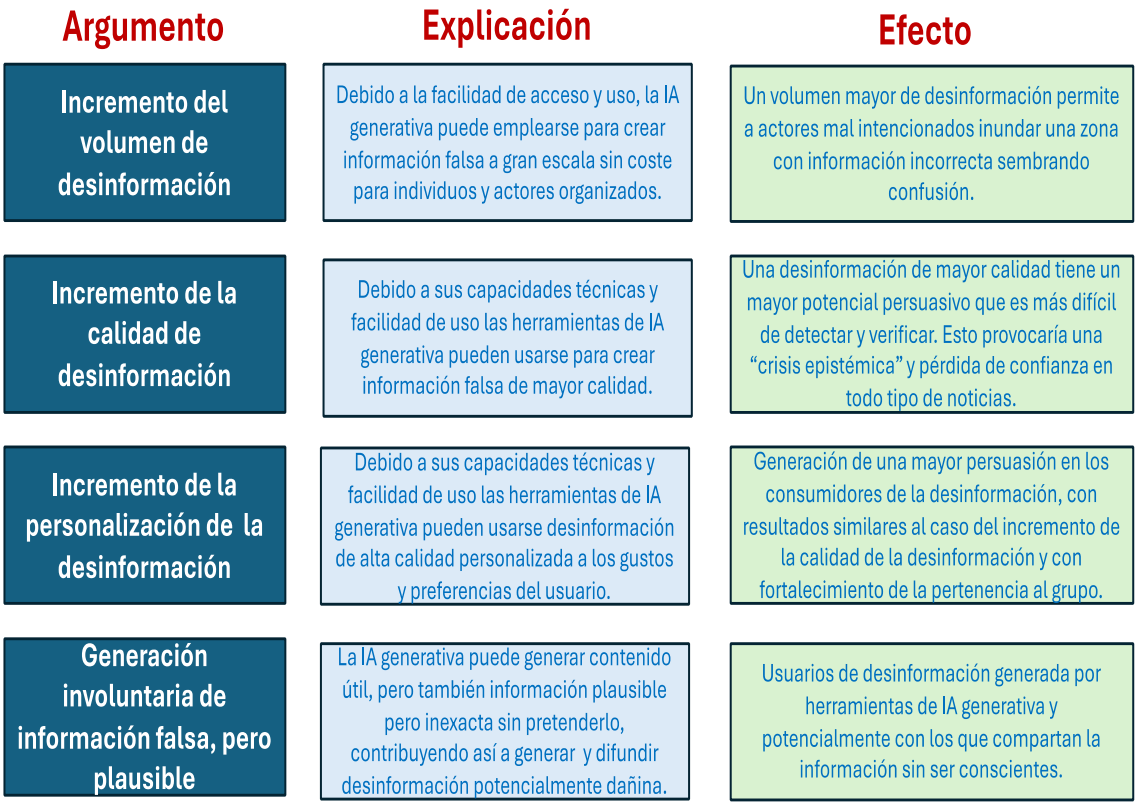


Figura 47. Argumentos y efectos del uso de herramientas para la desinformación. Fuente: elaboración propia.

Desde nuestro punto de vista, la evolución del **contenido generado por IA** (*Artificial Intelligence Generated Content, AIGC*) va a desarrollarse muy rápidamente en el futuro, contribuyendo a **incrementar el volumen e impacto de las noticias falsas como parte de la guerra cognitiva** (Ramírez, 2023) (Cao et al., 2025)

Para Ergo Sum Team et al., (2024) *"Es casi seguro que actores maliciosos explotarán AIGC para atacar el dominio psicocognitivo. Los efectos AIGC consistieron en dos vectores principales: Vectores Psico-Cognitivos, que están diseñados para atacar el Dominio Psico-Cognitivo y la agencia epistémica, y Vectores a Nivel de Máquina, que están diseñados para atacar la tecnología e influir indirectamente en el Dominio Psico-Cognitivo"*. Esta visión se representa esquemáticamente en la figura 48.

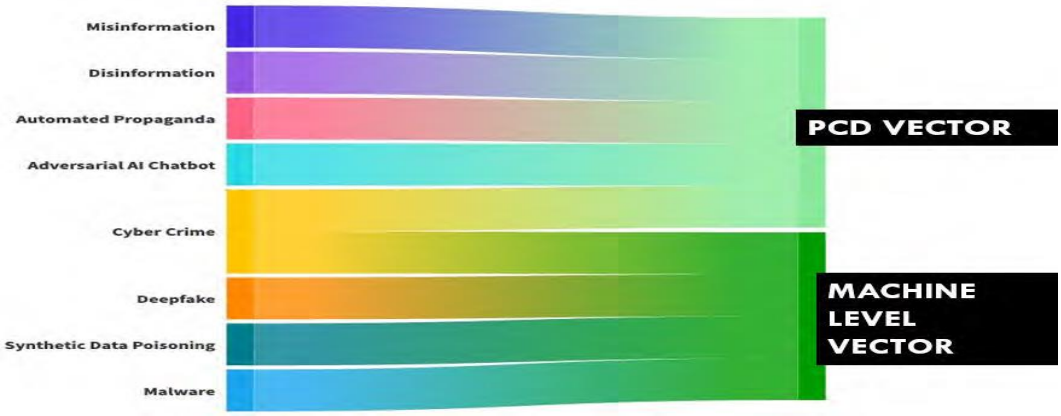


Figura 48. Efectos de la AICG. Fuente: Ergo Sum Team et al. 2024

China, en base a su desarrollo tecnológico en IA emplea, según la visión de Ergo Sum et al., (2024), mucho más en el uso de la tecnología (vector a nivel de máquina) como se representa en la figura 49.

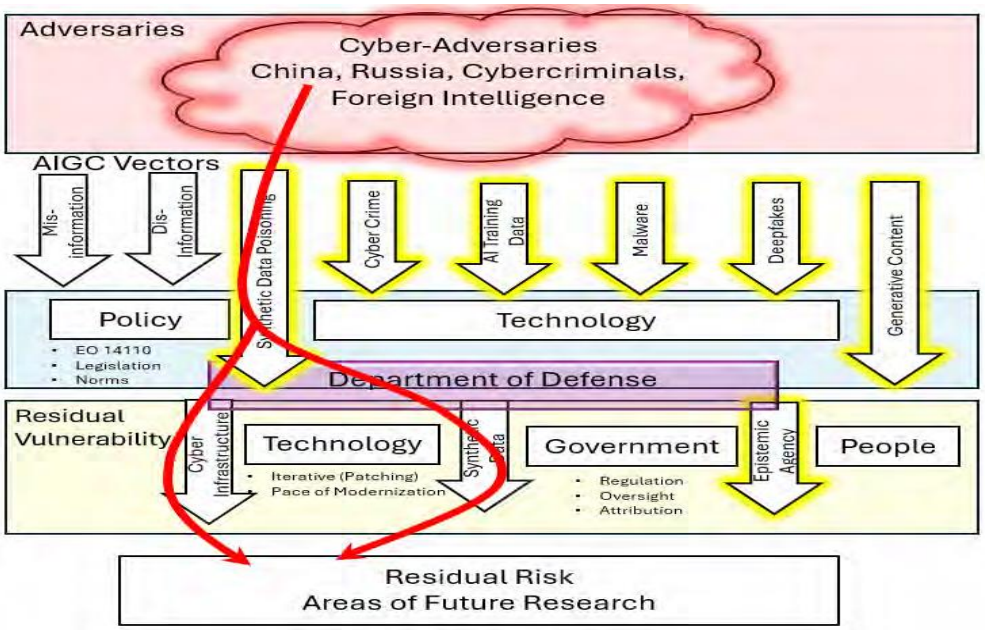


Figura 49. Empleo por China del vector de nivel de máquina. Fuente: Ergo Sum et al. 2024

China confía en sus capacidades tecnológicas en IA apoyadas por un uso intensivo y avanzado de técnicas de *big data* para generar contenido específico de sus campañas de desinformación e incrementar su impacto.

3.2.2. Herramientas de IA de lucha contra la desinformación

De igual manera que la IA está sirviendo para alimentar la generación de desinformación, también puede emplearse para luchar contra ella basado en la **comprobación automática de la veracidad de determinada información.** Se puede decir que **los modelos de IA generativa son un intento innovador de hacer frente a las noticias falsas,** ya que presentan herramientas complementarias a las disponibles actualmente (Nayuni, 2024) aprovechando la capacidad de los modelos de procesamiento de datos en

lenguaje natural, aprendizaje automático y generación de datos sintéticos, **los modelos mejoran la identificación de información falsa y fraudulenta**. Concretamente, estas tecnologías analizan la estructura, semántica y fuentes de las noticias, comparándolas con bases de datos verificadas para identificar patrones de desinformación.

El mapa de agrupación de esfuerzos realizados que se indica en la figura 50 (Pilati y Venturini, 2025) permite ver la multiplicidad de esfuerzos en marcha agrupados por las relaciones entre ellos. Muchos de ellos han generado herramientas de IA, pero su uso no está generalizado[87].

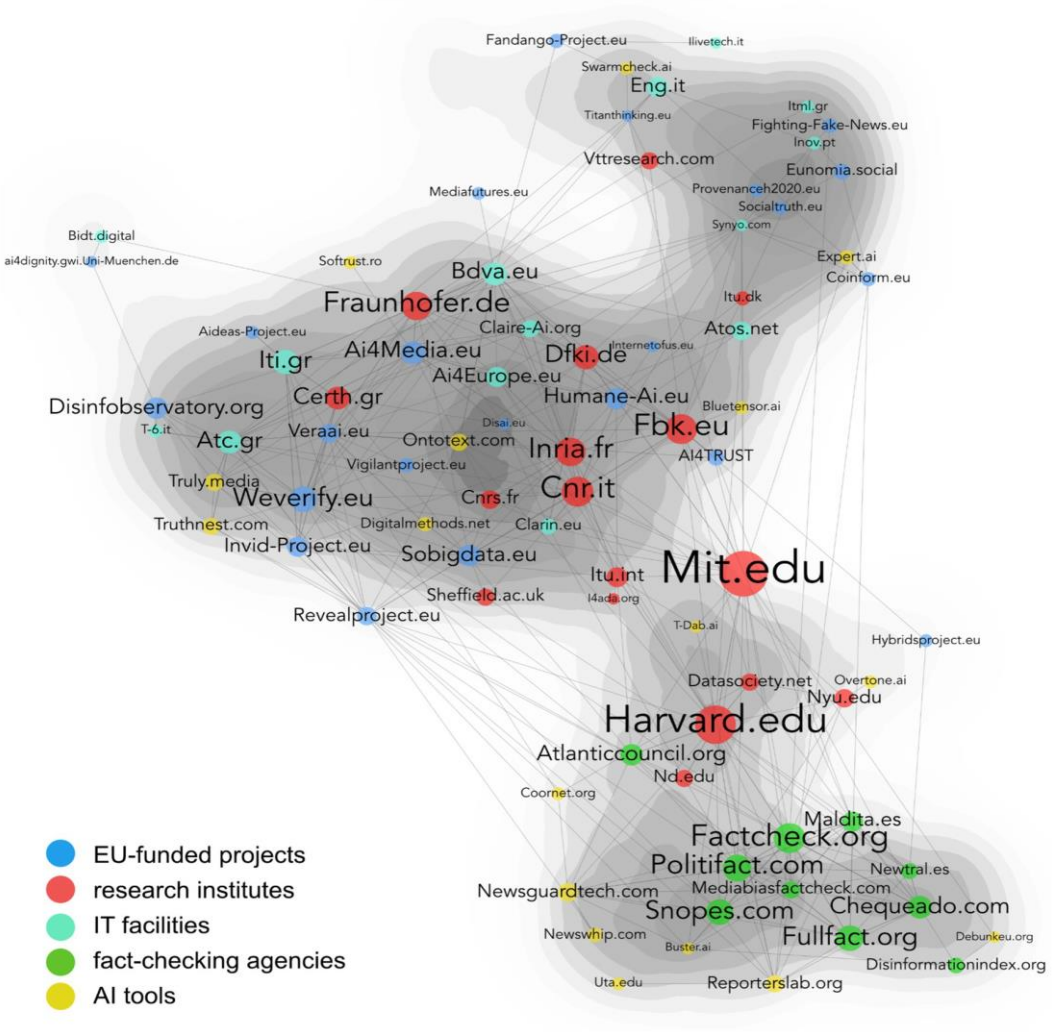


Figura 50. Proyectos públicos, corporativos y herramientas de IA contra la desinformación. Fuente: Pilati y Venturini, 2025.

Existen actualmente diversas herramientas de IA cuya idea básica es **determinar si una *"frase"*** (mensaje en los medios) **está respaldada por hechos para considerarla veraz** (un hecho es una sentencia que puede ser verificada, una opinión no es un hecho). Se presentan a modo de ejemplo tres herramientas.

[87] La posición de los nodos en el espacio está determinada por el algoritmo *Force Atlas 2*, que considera la fuerza y el tipo de conexiones entre nodos. Cuanto más cerca están dos nodos en la visualización, más fuertes y numerosas son sus conexiones directas o indirectas,

Logically: *https://logically.ai/*
Proporciona una aplicación móvil y una extensión de navegador que ofrece servicios de verificación de hechos e imágenes, combinando IA y verificadores humanos. La IA analiza afirmaciones, opiniones y eventos, monitorizando en tiempo real más de un millón de dominios web y plataformas de redes sociales para evaluar la veracidad de la información (véase figura 51).

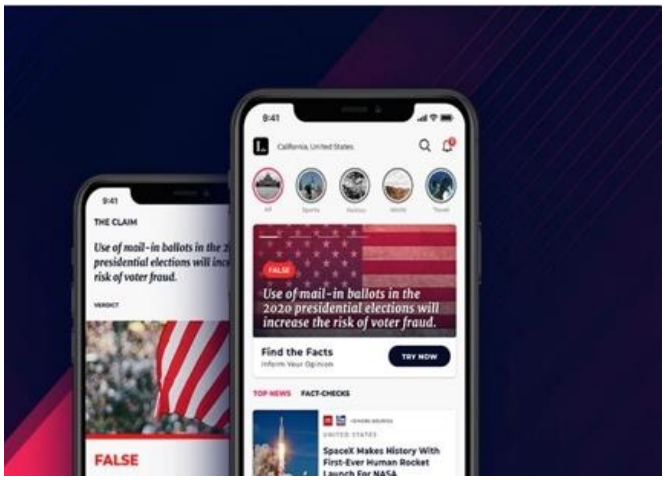


Figura 51. Logically. Fuente: https://logically.ai/

Stand4Israel: *https://stand4israel.net/*
Plataforma creada por el gobierno de Israel para combatir las noticias falsas y la desinformación (véase figura 52). La misión es recopilar la verdad simple sobre los eventos en curso en Israel, obtenida directamente de medios de comunicación oficiales y confiables en Israel y en todo el mundo.

Figura 52. Stand4Israel. Fuente: https://stand4israel.net/

Full Fact: https://fullfact.org/ai/
Ofrece varias herramientas de verificación de hechos, incluidas las que están automatizadas mediante el uso de IA (véase figura 53). Incluye herramientas de IA para ayudar a los verificadores de datos a comprender cuál es la información más importante y digna de verificación del día. Se diseñó un algoritmo para identificar cuándo alguien repite

a sabiendas algo que sabe que es falso. La información de noticias detectadas como *fake news* se actualiza muy frecuentemente (https://fullfact.org/latest/).

Figura 53. Herramienta Full Fact. https://fullfact.org/ai/

Todas las grandes empresas digitales que proporcionan servicios digitales de información a través de plataformas digitales se han provisto de herramientas de verificación. Como ejemplo, ***Meta***, que incluye *Facebook* e *Instagram*, ha estado invirtiendo en I+D para detectar y combatir los deepfakes y la desinformación. Desde marzo de 2024, *Meta* comenzó a **etiquetar el contenido generado por IA (AIGC) en sus sitios web**. Para ello, *Meta* utiliza una combinación de algoritmos de IA, que incluyen aprendizaje automático y visión por computadora, para analizar el contenido e identificar posibles manipulaciones. *Meta* también colabora con organizaciones de verificación de datos de terceros para verificar la exactitud de la información. Actuaciones similares se llevan a cabo por *Google*, *Microsoft*, *Intel* y muchas otras empresas.

En todo caso, los procedimientos empleados, aunque se apoyan en IA, no están totalmente automatizados y **requieren el uso de operadores humanos** de manera intensiva. Es interesante indicar que todos los medios de comunicación actuales de cierta entidad emplean este tipo de herramientas, acuciados por su propia responsabilidad y reputación frente a sus lectores, y también por la progresiva introducción de normativas y legislaciones que les obligan a ello.

Otra herramienta más sofisticada, en realidad, se trata de una *plataforma de servicio contra la desinformación* es **ClaimBuster (**https://idir.uta.edu/claimbuster/). Se trata de una **herramienta de verificación de datos en vivo automatizada** basada en la web desarrollada por la Universidad de Texas en Arlington (*ClaimPortal* es utilizado como la plataforma de servicios a usuarios registrados).

La herramienta se basa en el **procesamiento del lenguaje natural y el aprendizaje supervisado** (basado en un conjunto de datos codificado por humanos) para identificar información fáctica y falsa. En la figura 54 puede verse la arquitectura general de la plataforma. En el esquema se representa el componente de "detección de reclamaciones", resaltado en el cuadro delineado en azul claro. La API proporciona un fácil acceso a los modelos y se puede acceder gratuitamente.

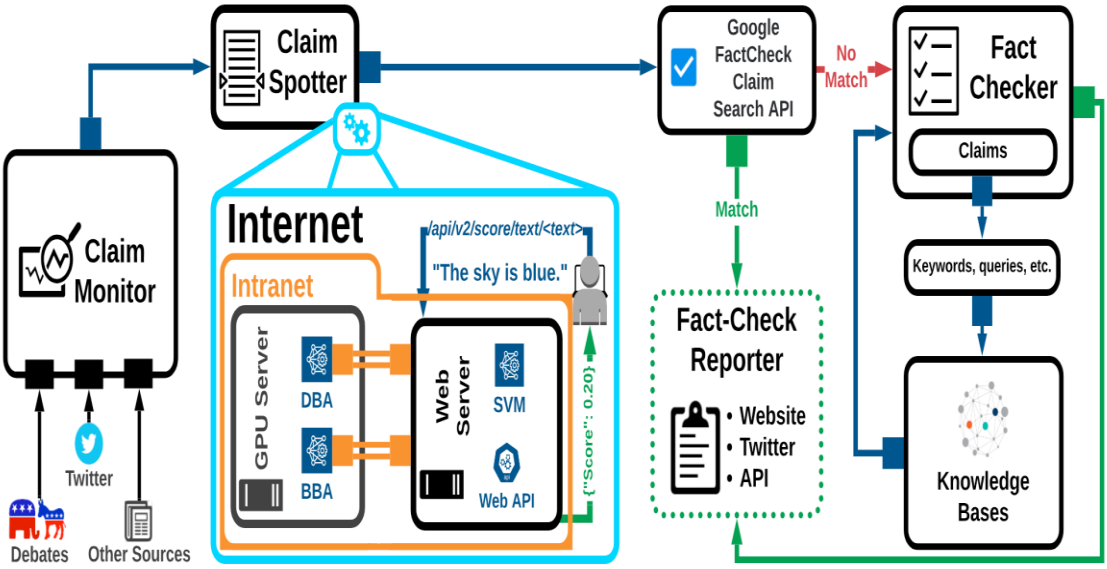


Figura 54. Arquitectura de ClaimPortal. Fuente: https://github.com/idirlab/idirpubs/blob/main/publications/2024/behavioraltendencies-asonam24-zhang.pdf

Las herramientas denominadas ***"verificadores de datos"*** leen un documento e identifican y extraen todo el contenido que necesita verificación de hechos. Suelen basarse en interfaces tipo GPT, especializado para ofrecer capacidades precisas de verificación de hechos analizando contenido sospechoso en interacción con el usuario (Gutiérrez-Caneda, 2024). La combinación de la IA con la experiencia humana mejora la precisión. Por ejemplo, ***FactCheck.org***[88] integra el aprendizaje automático con el análisis de redes sociales para rastrear las tendencias de desinformación.

Una herramienta de este tipo es la proporcionada por ***Google*** (*Google Fact Check Explorer*)[89] y sobre las que empresas de distribución de noticias (agencias) han elaborado sus propios verificadores automáticos. Agencias en España como *Newtral*, *EFE Verifica*[90] y *Maldita.es* son miembros de la *International Fact Checking Network*[91] y siguen estándares internacionales para garantizar el máximo rigor en sus verificaciones[92].

3.2.3. Herramientas experimentales de IA generativa para detectar noticias falsas

No siempre es sencillo determinar la veracidad de la noticia y para ello es necesario emplear **herramientas de IA más especializadas**; algunas de ellas son aún prototipos generados en proyectos de investigación. Dos ejemplos son:

[88] https://www.factcheck.org/

[89] https://newsinitiative.withgoogle.com/es-es/resources/trainings/google-fact-check-tools/

[90] https://verifica.efe.com/?t&utm_source=perplexity

[91] Lanzada en 2025 por el Instituto Poynter agrupa a 170 organizaciones comprobación de hechos de todo el mundo https://www.poynter.org/ifcn/

[92] https://www.ttandem.com/blog/contenidos-asi-funcionan-los-verificadores-de-noticias/?utm_source=perplexity

- **Mejora del MIT** (2023) de un modelo basado en transformadores (BERT) con una unidad recurrente cerrada bidireccional, logrando una puntuación F1 del 98 % en la detección de noticias falsas, lo que indica la capacidad excepcional del modelo para distinguir el contenido falso del real en el conjunto de datos. La puntuación F1 es una métrica que equilibra la precisión (cuántas noticias falsas identificadas son realmente falsas) y la memoria (cuántas noticias falsas reales se identificaron correctamente)[93].

- **FANDANGO** (*FAke News discovery and propagation from big Data ANalysis and artificial intelliGence Operations*), financiado por la UE en el programa H2020[94]. Detección de noticias falsas como "servicio" realizado para RTVE y la fundación CIVIC. Se basa en la recogida de datos multimedia junto a la creación de un modelo de "*lago de datos*" sobre el que se desarrollan herramientas de aprendizaje automático y análisis de datos para detectar falsificación de imágenes o vídeo y detección de la fuente en actividades de noticias falsas[95].

- **Keele** (Asowo el al., 2025). Una herramienta desarrollada por investigadores de la Universidad de Keele para detectar noticias falsas ha obtenido un impresionante nivel de precisión del 99%. El método desarrollado por los investigadores utiliza una técnica de "*votación por conjunto*", que combina las predicciones de múltiples modelos de aprendizaje automático diferentes para dar una puntuación general.

Un requisito previo es disponer de conjuntos de datos como ***LIAR***[96], que proporcionan **datos etiquetados para entrenar modelos de IA, lo que permite un mejor rendimiento en la identificación de noticias falsas en diversas plataformas**. LIAR es un conjunto de datos disponible públicamente para la detección de noticias falsas. Se recopilaron 12.8K declaraciones cortas etiquetadas manualmente en varios contextos de POLITIFACT.COM, que proporciona un informe de análisis detallado y enlaces a documentos de origen para cada caso. Este conjunto de datos también se puede utilizar para la investigación de verificación de hechos. En particular, este nuevo conjunto de datos es un orden de magnitud mayor que los conjuntos de datos públicos de noticias falsas más grandes de tipo similar.[97]

Herramientas como el proyecto ***IVERES*** *("Identificación, Verificación y Respuesta. El estado democrático ante el reto de la desinformación interesada"*)[98], liderado por la Universidad Autónoma de Barcelona e implementado por RTVE, utiliza la IA para analizar

[93] https://builtin.com/artificial-intelligence/fight-ai-generated-fake-news

[94] https://cordis.europa.eu/project/id/780355/reporting/es

[95] https://www.upm.es/sfs/Rectorado/Vicerrectorado%20de%20Investigacion/AIR4S/BP_Society_Artificialfakenews.pdf

[96] https://www.kaggle.com/code/hendrixwilsonj/liar-data-analysis

[97] El conjunto de datos LIAR4 incluye 12,8 K declaraciones cortas etiquetadas por humanos de la API de *politifact.com*, y cada declaración es evaluada por un editor de *politifact.com* para determinar su veracidad https://www.politifact.com/

[98] https://iveres.es/ (terminado en noviembre de 2024).

textos, imágenes y audios, evaluando su autenticidad en tiempo real. Este sistema combina flexibilidad y escalabilidad para adaptarse a diferentes contextos informativos. Una de las innovaciones clave del Proyecto IVERES es la **creación de ontologías personalizadas** que mejoran la capacidad de monitorización y análisis de contenido en redes sociales. Estas ontologías han sido desarrolladas en colaboración con la Universidad Carlos III de Madrid y se aplican en áreas como en la lucha contra la desinformación, el racismo, la xenofobia y la violencia de género[99].

Asimismo, empresas startups como ***AdVerif.ai*** emplean algoritmos que detectan inconsistencias entre titulares y contenido, generando **informes detallados sobre la probabilidad de falsedad**[100]. Su algoritmo actúa en varias fases:

- Detección de irregularidades en el contenido: Escanea titulares y cuerpos de texto para identificar inconsistencias, como titulares que no coinciden con el contenido, uso excesivo de mayúsculas o lenguaje emocional.
- Verificación en bases de datos: Contrasta las noticias con una base de datos que incluye miles de publicaciones legítimas y falsas, actualizada semanalmente, para identificar patrones o coincidencias.
- Generación de informes: Proporciona a los clientes un informe con una puntuación que indica la probabilidad de que el contenido sea falso o problemático.
- Reconocimiento avanzado: Puede identificar sátiras, analizar tendencias políticas del medio y detectar incongruencias en perfiles sociales, como cuentas automatizadas o bots.

Aunque la IA generativa parece prometedora para la detección de noticias falsas, su eficiencia está ligada a la mejora gradual de las tecnologías desde una visión multimodal (con imágenes, voz, audios, etc.), a la forma de abordar cuestiones éticas y a la evolución de la regulación para hacerlas más útiles y facilitar su adopción, en lugar de ser un sustituto de las prácticas tradicionales de verificación manual.

La tecnología relacionada con las noticias falsas y la IA generativa para crearlas han hecho posible que los denominados informalmente ***Manipuladores Persistentes Avanzados (APM)*** creen y difundan desinformación de una nueva manera. Su uso en campañas de desinformación y sociales (Ramírez, 2023), así como operaciones psicológicas requiere un enfoque igualmente avanzado y multidimensional de la ciberseguridad y la defensa cognitiva.

3.2.4. Análisis de sentimientos con herramientas de IA

Valorar el impacto de un determinado documento como factor de influencia en la sociedad, más allá de su veracidad, es un factor cuyo análisis está adquiriendo una creciente relevancia. Un enfoque basado en IA empleado para ello es el denominado **análisis de sentimientos**, también conocido como *minería de opiniones*[101].

[99] https://rubik-audiovisual.com/tecnologia-para-la-lucha-contra-la-desinformacion-y-la-discriminacion/

[100] https://clubinterprensa.org/una-ia-capaz-de-detectar-fake-news/?utm_source=perplexity

[101] https://www.ibm.com/es-es/topics/sentiment-analysis

El análisis de sentimientos es el proceso de analizar grandes volúmenes de texto para determinar si expresa un sentimiento positivo, negativo o neutro. La clasificación de sentimientos evolucionó para proporcionar **análisis de sintaxis multilingüe o análisis de entidades**; fundamentalmente, se ha empleado en el marketing digital.

Se trata de una aplicación de la IA en el campo del **procesamiento del lenguaje natural (PLN)** empleada para clasificar un texto en varios tipos de "***sentimientos***", como positivos o negativos, felices, tristes o neutros. Este proceso se realiza en varias fases[102]:

- División de cada documento de texto en sus partes componentes (oraciones, frases, fichas y partes de la oración)
- Identificación de cada frase y componente portador de sentimiento
- Asignación de una puntuación de opinión a cada frase y componente (de -1 a +1)
- Opcional: Combinación de puntuaciones para análisis de sentimientos de varias capas.

Desde un punto de vista más técnico, las soluciones de análisis de sentimiento que se han empleado en el dominio civil para conseguir sistemas de atención al cliente más efectivas de la actualidad **combinan técnicas de aprendizaje automático, procesamiento de lenguaje natural y estadísticas avanzadas** integrados en sistemas multiagente inteligentes para mejorar las experiencias de los usuarios. El desarrollo de técnicas de IA generativa multimodal permite extender este uso a la voz (teniendo en cuenta parámetros vocales como el tono y la modulación) y el lenguaje corporal en el caso de vídeos (movimiento de las manos, ojos, sudoración, etc.) que no son posibles de obtener con técnicas de procesamiento de lenguaje natural.

Actualmente, **existen en el mercado varios paquetes de software de análisis de sentimientos** con interfaces de aplicación para el usuario (*Application Interface, API*), como *Google Cloud Natural Language API*, *Microsoft Azure Text Analytics API* o *Lexalytics Salience*[103] que se pueden personalizar para la investigación de mercado, la gestión de marcas y el análisis de productos/servicios. Otro ejemplo es *IBM watsonx Assistant*[104], plataforma de inteligencia artificial conversacional impulsada por modelos de lenguaje de gran tamaño (LLM), que permite a las organizaciones usuarias crear agentes de voz y chatbots con IA con el que ofrecer soporte de autoservicio automatizado en una interfaz conversacional. En Sheikh (2025) se evalúan 16 herramientas de análisis de sentimiento con IA.

Desde hace algunos años (Kahn et. Al., 2021) se ha empezado a analizar si es posible **obtener información de los sentimientos de una persona con técnicas inalámbricas no invasivas** a partir de datos biomédicos tales como la frecuencia cardíaca, la frecuencia respiratoria y la electroencefalografía (EEG) relacionados con varias manifestaciones físicas de las emociones. La idea, como se indica en la figura 55, consiste en obtener información como la mencionada de un sujeto monitorizado en respuesta a estímulos preparados individualmente (recuerdos, fotos, música y vídeos) que **evocan una determinada emoción** durante el experimento.

[102] https://www.lexalytics.com/technology/sentiment-analysis/

[103] https://www.lexalytics.com/salience/

[104] https://www.ibm.com/es-es/products/watsonx-assistant

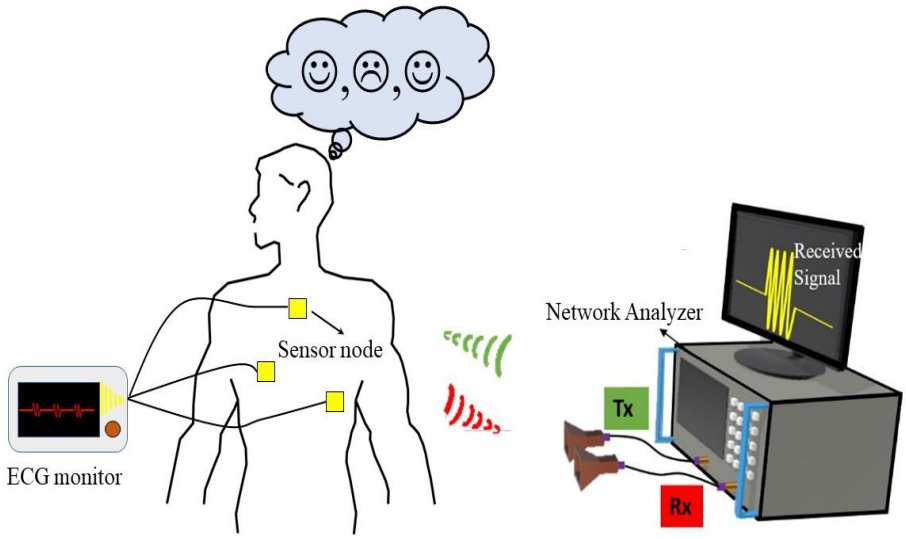


Figura 55. Análisis de sentimientos mediante señales de radiofrecuencia. Fuente: Khan et al., 2021)

La señal de RF reflejada en el cuerpo del sujeto se preprocesa e introduce en algoritmos de aprendizaje automático (ML) para clasificar cuatro tipos básicos de emociones, como la ira, la tristeza, la alegría y el placer. Un paso más realizado por Khan y sus colegas fue utilizar redes neuronales profundas en vez de algoritmos "clásicos" de aprendizaje automático[105].

El campo del análisis de sentimientos en el ámbito de la defensa y, sobre todo, en el de la seguridad se originó por la necesidad de **examinar la subjetividad de los textos** difundidos en medios de comunicación o documentos internos de organizaciones en áreas como la lucha antiterrorista y determinar los grados de radicalización, o conocer el nivel de moral hacia el combate de soldados en el frente de batalla.

Como ejemplo de "**análisis de sentimientos**" es interesante comentar los resultados de un estudio bibliométrico de 2376 artículos del *Web of Science (WoS*) que relacionaban los conceptos de **Guerra** y **Ucrania** (junio de 2023)[106]. En la figura 56, a la izquierda, puede verse el mapa de correlaciones, y, a la derecha, el análisis de las correlaciones más relevantes. La **clasificación de los mensajes** sobre los conceptos o términos relacionados con la guerra de Ucrania que se indican en la gráfica de la derecha de la figura responde a una **tipificación del mensaje como positivo, neutro o negativo**. Obsérvese que la mayoría de las opiniones analizadas se tipificaron como "neutras".

[105] Para la clasificación se utilizó una arquitectura de red neuronal convolucional (CNN) integrada con celdas de aprendizaje de secuencias de memoria a corto y largo plazo (LSTM) que aprovechan tanto la señal de RF procesada como la reflexión de RF sin procesar.

[106] https://tdhj.org/blog/post/artificial-intelligence-cognitive-warfare-twitter/

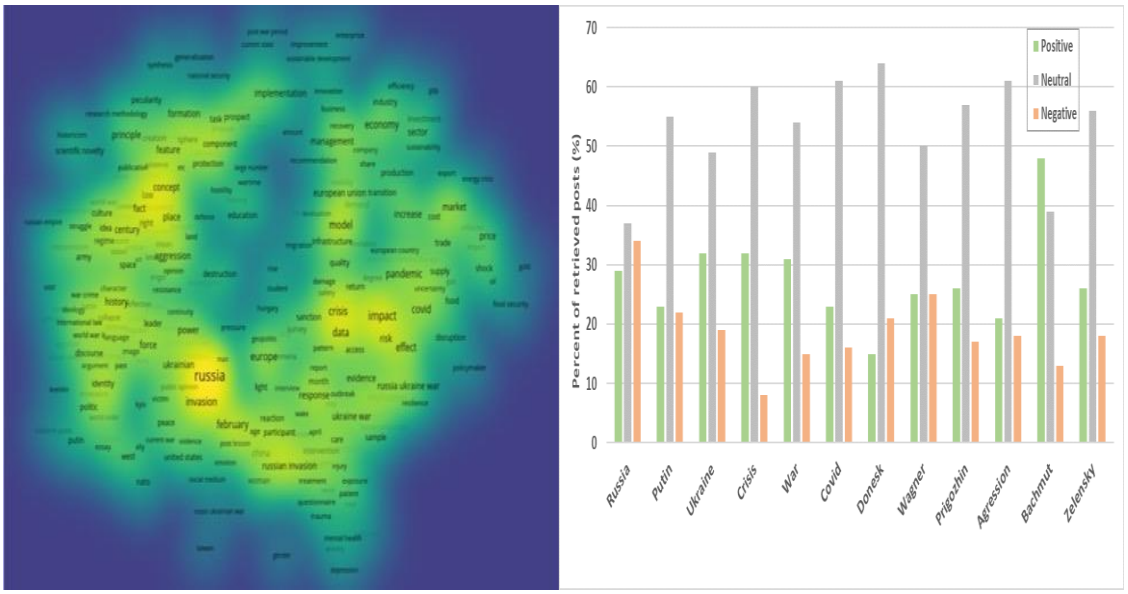


Figura 56. Análisis de sentimientos. Fuente: https://tdhj.org/blog/post/artificial-intelligence-cognitive-warfare-twitter/

Más recientemente, Kertcher y Zwilling (2025) aplicaron dos técnicas de procesamiento de lenguaje natural (VADER y BERT) para realizar un "análisis de sentimientos" focalizado en las intervenciones en la ONU sobre la guerra de Ucrania. El trabajo exploró cómo se puede **interpretar el sentimiento global a través de un enfoque de método mixto que combina técnicas innovadoras de PNL con análisis narrativo**. Los discursos diplomáticos presentan retos significativos para los métodos de procesamiento de lenguaje natural debido a su complejidad, renuencia a adoptar una postura definitiva y uso extensivo del contexto histórico y cultural.

Entre los resultados se destaca que los discursos con sentimiento positivo generalmente enfatizan los valores, normas, instituciones y actores internacionales que apoyan a Ucrania, junto con varias soluciones al conflicto. Por el contrario, los discursos con sentimiento negativo se centran más en relatos detallados del sufrimiento ucraniano y la agresión rusa, incluida la ocupación, los bombardeos civiles y las amenazas nucleares.

El uso de las técnicas de análisis de sentimientos en la guerra cognitiva plantea serias **preocupaciones éticas si se combina con técnicas de manipulación de los sentimientos**, de difusión de desinformación o el uso de tácticas psicológicas para influir en las creencias y comportamientos de las personas, y, específicamente, personal militar. Conceptualmente, estas prácticas pueden afectar negativamente a las personas, las sociedades y los procesos democráticos, pero también a la **capacidad de resistencia** frente a adversarios[107].

3.2.5. Uso de la IA en la valoración y negociación de conflictos militares

La emergencia de la IA generativa y el uso de agentes inteligentes permite a los gobiernos su **integración en la toma de decisiones relativas a conflictos militares y de política exterior** de alto riesgo, favoreciendo el análisis de escenarios en los que se analizan

[107] https://tdhj.org/blog/post/artificial-intelligence-cognitive-warfare-twitter/

opciones entre diferentes "bandos". Especialmente con la aparición de modelos avanzados de IA generativa como GPT-4, el uso de los "*juegos de guerra*" (escenarios simulados de conflictos) permite analizar la dinámica de los procesos de escalada.

Dado que muchas empresas en Estados Unidos, entre ellas *Palantir Technologies Inc*. y *Anduril Industries Inc*. están desarrollando **plataformas de decisión basadas en IA** para el ejército de Estados Unidos es muy relevante conocer cómo sistemas de IA generativa responden a diversas situaciones y si sus respuestas están polarizadas.

En una experiencia desarrollada en 2023 sobre un escenario en el Pacífico[108], basada en alimentar el modelo con 60.000 páginas de datos de código abierto, incluidos documentos militares estadounidenses y chinos, se preguntaba al sistema si Estados Unidos podría disuadir un conflicto en Taiwán y quién ganaría si estallara la guerra. El sistema respondió: "Probablemente sería necesaria la intervención directa de EE.UU. con fuerzas terrestres, aéreas y navales", advertía en otra que "EE.UU. tendría dificultades para paralizar rápidamente al ejército de China". La nota final del sistema era: "Hay poco consenso en los círculos militares sobre el resultado de un posible conflicto militar entre Estados Unidos y China sobre Taiwán".

Un elemento clave es conocer si las respuestas obtenidas por un sistema de IA generativa son similares ¿Contestarían igual otros sistemas de IA? ¿Se pueden producir alucinaciones? ¿Hasta qué punto depende de los datos de entrenamiento? Dadas las diferencias existentes entre diferentes sistemas de IA generativa basados en LLM, es necesario analizar si el LLM elegido condiciona las respuestas y, por tanto, las decisiones. A raíz de estas experiencias el Departamento de Defensa de Estados Unidos trabaja con empresas de seguridad tecnológica para **probar y evaluar en qué medida pueden confiar en los sistemas de toma de decisión potenciados por IA**.

En una experiencia llevada a cabo por Rivera et al. (2024) se proporciona información cualitativa y cuantitativa centrada en el uso de grandes modelos de lenguaje (LLM). En resumen, la interacción con los cinco LLM estudiados muestran **formas y patrones de escalada difíciles de predecir**. Concretamente, los investigadores observaron que los modelos tienden a desarrollar una dinámica de "carrera armamentista", lo que conduce a un mayor conflicto y, en casos excepcionales, incluso al despliegue de armas nucleares.

En el experimento realizado, ocho *agentes nacionales autónomos*, todos utilizando el mismo modelo de lenguaje (GPT-4, GPT-3.5, Claude 2, Llama-2 (70B) Chat o GPT-4-Base) interactuaron entre sí en simulaciones por turnos sobre un escenario simulado (modelo del mundo)[109]. Se consideraron tres escenarios posibles:

- **Escenario neutro**, sin eventos iniciales
- **Escenario de invasión** donde un agente de una nación invadió a otro antes del inicio de la simulación.
- **Escenario de ciberataque** donde un agente nación llevó a cabo un ataque cibernético a otro antes del inicio de la simulación.

[108] https://www.csis.org/analysis/first-battle-next-war-wargaming-chinese-invasion-taiwan.

[109] En cada turno, 1) los agentes realizan acciones predefinidas que van desde visitas diplomáticas hasta ataques nucleares y envían mensajes privados a otras naciones. 2) Un modelo de mundo separado LLM resume las consecuencias de las acciones en los agentes y el mundo simulado. 3) Las acciones, los mensajes y las consecuencias se revelan simultáneamente después de cada día y se incorporan a las indicaciones para los días siguientes.

Para cada uno de esos escenarios, los *agentes-nación* podían tomar diferentes acciones en cada ronda de simulación con pesos diferentes[110]. Después de las simulaciones, se calcularon las puntuaciones en un rango de escalada (ES) en función de puntuaciones. En la figura 57 se puede ver el esquema empleado para analizar el uso de LLMs basado en la interacción conversacional de operadores humanos con agentes inteligentes.

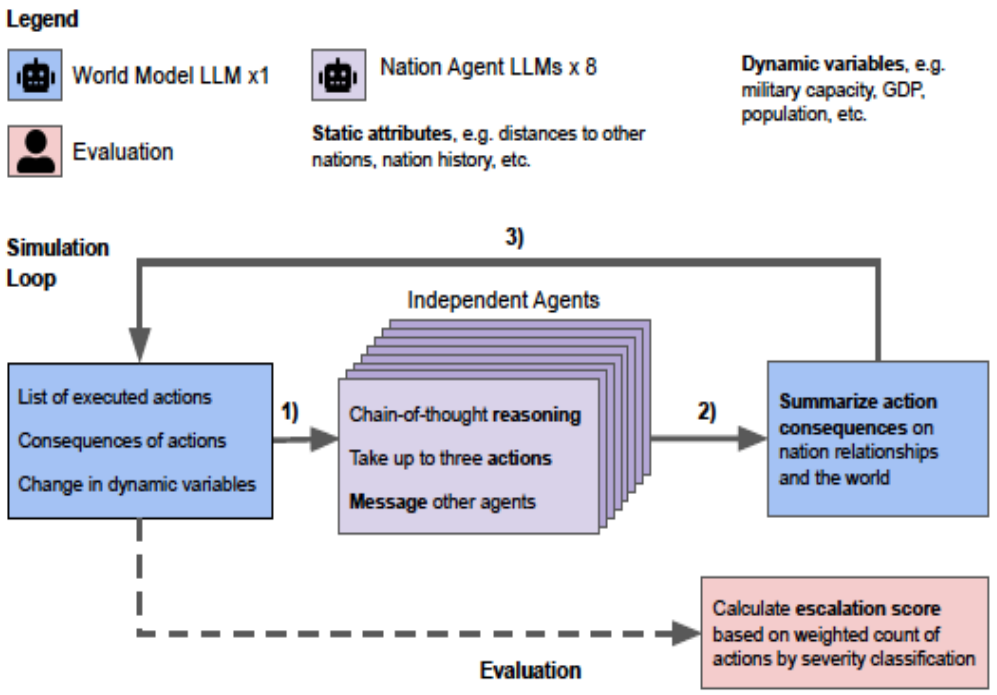


Figura 57. Uso y valoración de LLMs en el proceso de toma de decisión para analizar la escalada de conflictos. Fuente: Rivera et al., 2024

Los resultados descritos en el trabajo (Rivera et al., 2024) muestran una **escalada inicial estadísticamente significativa para todos los modelos**. Además, ninguno de los cinco modelos en los tres escenarios muestra una desescalada estadísticamente significativa a lo largo de las simulaciones.

3.2.6. Responsabilidades y límites en la detección de noticias falsas para los proveedores de plataformas digitales

Una discusión que ha emergido con fuerza en los dos últimos años es la de determinar la **responsabilidad legal**, no solo ética, que tienen o deben tener las grandes empresas, ya sean proveedoras o usuarias de servicios digitales en la detección de noticias falsas impulsada por IA. Sus responsabilidades dependen el pise y marco legal aplicable, pero pueden incluir:

- **Monitorización activa**: Implementar herramientas de escucha en la sociedad (*social listening*) para identificar menciones y posibles bulos en redes sociales y medios digitales en tiempo real.

[110] Las acciones y pesos eran: desescalada (-2 puntos), status-quo (0 puntos), postura (4 puntos), escalada no violenta (12 puntos), escalada violenta (28) y escalada nuclear (60 puntos).

- **Protocolos de crisis**: Diseñar manuales específicos para gestionar crisis relacionadas con noticias falsas, evaluando las acciones tomadas para mejorar continuamente.
- **Colaboración y transparencia**: Asociarse con organizaciones de verificación de hechos y mantener canales oficiales como fuentes confiables para contrarrestar la desinformación.
- **Uso de tecnología avanzada**: Adoptar tecnologías como Inteligencia Artificial Explicable (XAI) para detectar y verificar información falsa de manera eficiente.
- **Educación interna y externa**: Capacitar a empleados y otras partes interesadas en la identificación de noticias falsas y fomentar la alfabetización mediática entre el público.
- **Acciones preventivas**: Suspender ingresos publicitarios en plataformas que promuevan desinformación, cerrar cuentas falsas y fomentar la transparencia en la publicidad programática.

Utilizando un enfoque de "auto-regulación" promovido por diversos países, las grandes empresas proveedoras de servicios digitales han desarrollado o adquirido herramientas de IA específicas como *Google* con su "*Google News Initiative*" que utiliza la IA para priorizar el contenido confiable y de alta calidad en su plataforma; o *Meta* (en su red social *Facebook*) que ha implementado una herramienta de IA para detectar patrones de desinformación y colabora con verificadores de datos humanos para proporcionar contexto adicional para el contenido marcado.

En todo caso, **el marco regulatorio es confuso**[111] y *Meta* ha informado en enero 2025 que va a detener estos procesos de verificación de datos por terceros en Estados Unidos (no, por ahora, en el resto del mundo), reemplazándolo con un modelo de "*Community Notes*", que permitirá a los usuarios añadir contexto a publicaciones potencialmente engañosas, similar al sistema de la red *X* (anteriormente *Twitter*)[112]. *Meta* continuará moderando contenido relacionado con drogas, terrorismo y explotación infantil. No es muy distinto a lo que hace *Wikipedia* para valorar sus artículos.

Thomas Regnier, portavoz de la Comisión Europea de soberanía tecnológica, ha avisado a *Meta* de que está obligada a cumplir con las obligaciones de la *Ley de Servicios Digitales (Digital Services Act, DSA*) de la UE[113]. Con base en la normativa comunitaria, *"antes de implementar funcionalidades que puedan tener un impacto crítico en los riesgos sistémicos, las plataformas en línea de gran tamaño deben realizar una evaluación de riesgos y presentar un informe a los servicios de la Comisión*".

[111] La sección 230 de la *Communications Decency Ac*t exime a las plataformas digitales de su papel como medios de comunicación y las hace, por lo tanto, irresponsables.

[112] Las razones aducidas eran que los sistemas automatizados de Meta eliminaron publicaciones que cumplían con las normas debido a infracciones menores, afectando a millones de usuarios y comprometiendo el derecho a la información. https://www.rtve.es/noticias/20250107/meta-abandona-sistema-verificacion-datos-por-terceros-estados-unidos/16397475.shtml?utm_source=perplexity

[113] https://www.europapress.es/sociedad/noticia-bruselas-replica-zuckerberg-absolutamente-nada-ley-ue-impone-censura-meta-otras-plataformas-20250108152038.html

En todo caso, la Comisión Europea ha rechazado que la DSA que se aplica a las grandes plataformas ejerza presión sobre las compañías para forzar la retirada de contenidos legales. Además, Meta ha firmado el "*Código de buenas prácticas sobre desinformación*", en vigor desde 2022 que contiene **compromisos específicos sobre verificación de datos** junto a otros muchos actores [114].

Investigadores del **Foro Económico Mundial (WEF)**, indican **que la *"información falsa"* se encuentra en la cima de las preocupaciones globales actuales**, y anticipan que estas herramientas de IA gratuitas y ampliamente disponibles "*automatizarán y expandirán las campañas de desinformación"* con mayor velocidad y precisión; especialmente la supuesta capacidad de la tecnología para personalizar las campañas de desinformación adaptando un mensaje a las preferencias y convicciones del público objetivo[115].

Más allá de aprovechar las mejoras tecnológicas, se requiere un **marco legislativo integrado** que abarque leyes consensuadas y una aplicación con el establecimiento de prácticas óptimas para la integridad y la seguridad de los datos. Todo ello, junto a procesos educativos que aseguren no solo disponer de personal especializado, sino de una **población bien formada** para defenderse. Cómo hacer eso, al mismo tiempo que se mantengan los valores y libertades democráticas de las sociedades europeas, sigue siendo un capítulo abierto y no resuelto.

3.3. IA y cibercrimen

La frontera entre las áreas de defensa y seguridad en el contexto de guerra hibrida y cognitiva es tenue, y desde el punto de vista tecnológico, muy borrosa. Debe recordarse que los ciberataques se combinan con campañas de desinformación con un propósito conjunto de desestabilización previo o simultáneo al desencadenamiento de un conflicto militar abierto.

Un ámbito en el que la tecnología de IA influye directamente es en su **relación con la ciberseguridad**. No se va a abordar esta relación en toda su amplitud, sino únicamente en relación con el **cibercrimen** y, específicamente, con aquellas actuaciones procedentes de grupos ligados a actores gubernamentales cuando su objetivo no es la población civil (objetivo habitual de los ciberdelincuentes) sino entidades públicas o personas relevantes en los gobiernos si con ello se consigue un **efecto desestabilizador**. Su evolución en términos de volumen, relevancia económica, impacto social y sofisticación tecnológica ha obligado a prestarle una mayor atención desde el punto de vista político.

No se trata de un fenómeno novedoso. Las entidades responsables de seguridad han usado la IA para la ciberseguridad desde finales de los años 1980 con avances tecnológicos clave. Al principio, los equipos de seguridad usaban sistemas basados en reglas que desencadenaban **alertas** basadas en los parámetros que definían. A partir de 2000, los avances en el aprendizaje automático, han permitido a los equipos de

[114] Las principales plataformas en línea, las plataformas emergentes y especializadas, los agentes del sector de la publicidad, los verificadores de datos, las organizaciones de investigación y de la sociedad civil presentaron un Código de buenas prácticas en materia de desinformación reforzado tras las orientaciones de la Comisión de mayo de 2021. https://digital-strategy.ec.europa.eu/es/policies/code-practice-disinformation

[115] https://verfassungsblog.de/european-security-and-the-threat-of-cognitive-warfare/

operaciones comprender los **patrones de tráfico típicos** y las acciones del usuario en toda una organización para identificar y responder cuando ocurre algo inusual.

La mejora más reciente es el uso de la IA generativa que permite a los profesionales de seguridad profundizar en preguntas muy específicas utilizando agentes inteligentes. Un ejemplo es el de ***Microsoft Security Copilot***[116]. En la figura 58 puede verse una pantalla de su uso en el **análisis de un incidente de ciberseguridad**.

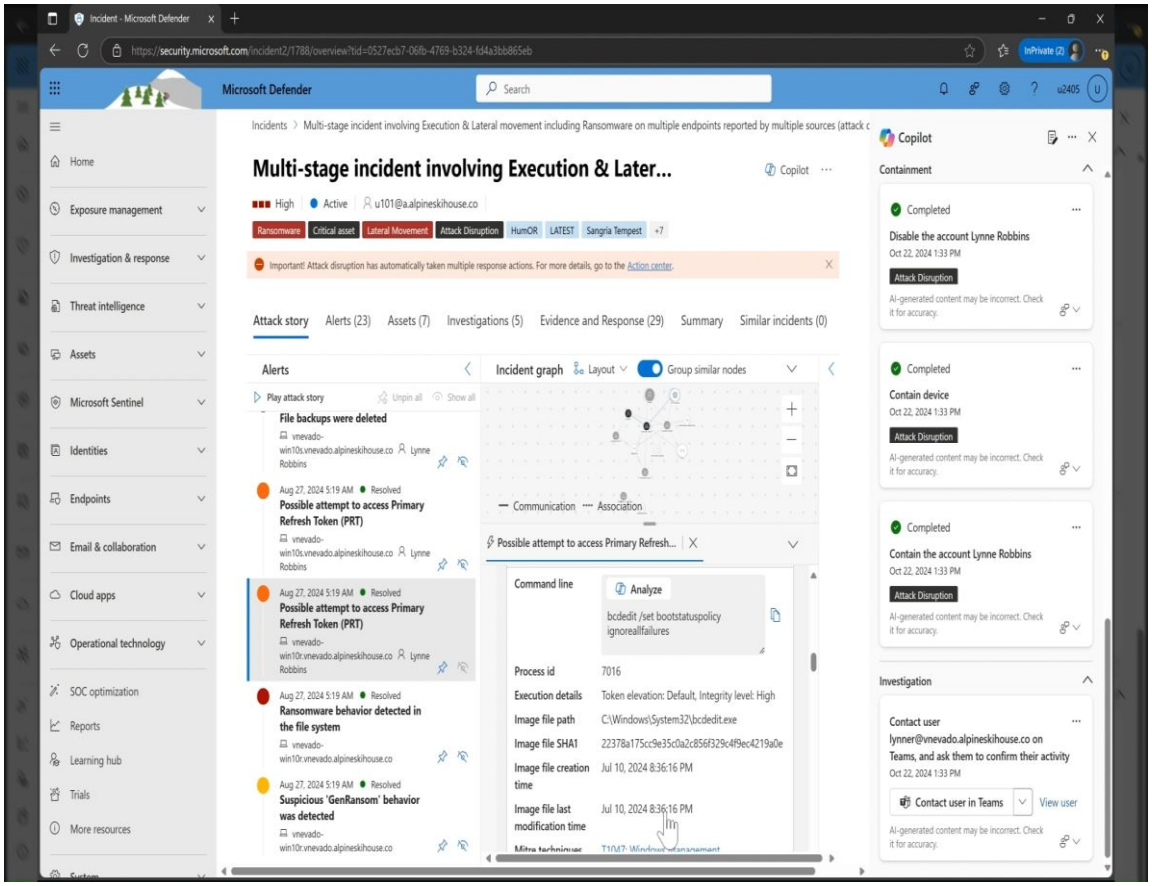


Figura 58. Uso de Microsoft Security Copilot. Fuente: https://www.microsoft.com/es-es/security/business/ai-machine-learning/microsoft-security-copilot

Este asistente pretende **simplificar las operaciones de seguridad** para analistas y otros profesionales de seguridad al **sintetizar los datos en recomendaciones y conclusiones procesables** con el contexto adecuado para ayudar a guiar las investigaciones de incidentes. También permite **crear informes y presentaciones legibles** que los analistas pueden usar para ayudar a otros usuarios de la organización a comprender lo que está sucediendo, y **responder a preguntas sobre un incidente** o una vulnerabilidad en lenguaje natural o mediante gráficos.

Obviamente, de igual manera que la IA se emplea por las fuerzas y cuerpos de seguridad, también lo hacen los ciberdelincuentes. De acuerdo con el informe de 2024 del Centro Criptológico Nacional (CCN, 2024), el **malware como servicio (MaaS)** ha proliferado a lo

[116] https://www.microsoft.com/es-es/security/business/ai-machine-learning/microsoft-security-copilot

largo de 2023 apoyado por el uso de herramientas más sencillas de usar por los delincuentes lo que permite que individuos con conocimientos tecnológicos limitados puedan obtener beneficios económicos a través de acciones de ciberdelincuencia. La falta de conocimientos tecnológicos para llevar a cabo acciones ha derivado en lo que se denomina **Hacking as a Service (HaaS)**, donde los técnicos proporcionan sus productos online a todo aquél que esté dispuesto a pagarlos[117].

Tanto la sofisticación como la oferta del *malware* ofrecido en estos servicios es cada vez mayor, dificultando su detección e incrementando el impacto que tiene en sus víctimas. Además, **grupos cibercriminales han comenzado a integrar el uso de la IA generativa en sus ciberataques** para aumentar sus probabilidades de éxito y hacer más rentable su modelo de negocio.

Como ejemplo de esta evolución, una nueva tendencia observada desde 2023 es la aparición de campañas de **publicidad maliciosa** (*malvertising)*[118] donde los atacantes se hacen pasar por empresas y herramientas de IA publicitando sus servicios, aprovechando la creciente demanda de este tipo de software para captar nuevas víctimas y obtener recursos económicos de ellas[119].

Otro ámbito en el que la IA ha tenido una difusión e impacto creciente es en el denominado "***fraude de identidad***" o también ***"fraude biométrico"***. El problema surge porque el acceso a multitud de servicios digitales esenciales para el ciudadano se realiza mediante una información de carácter biométrico como es la *cara del usuario,* que es comparada con otra almacenada previamente en el sistema. Con la extensión de la IA generativa, generar una imagen falsa es relativamente sencillo.

Un ejemplo de herramienta de IA para luchar contra ello es la proporcionada por *SumSub*. La empresa ha desarrollado una tecnología de detección avanzada con IA diseñada para realizar comprobaciones biométricas faciales que ha dado origen a la herramienta de "***verificación de vida***" (*Liveness verification*)[120]. La idea es confirmar que el verdadero propietario de un documento está presente con una comprobación de vida fácil de usar. La herramienta añade una **capa adicional de protección contra el robo de identificaciones y los ataques de suplantación de identidad**. En la figura 59 puede verse una imagen del uso de la herramienta.

---

[117] En su forma más simple, es un tipo de externalización de servicios de ciberseguridad. En lugar de que una empresa contrate a un empleado o equipo a tiempo completo para que se encargue de sus necesidades de ciberseguridad, puede subcontratar estas tareas a un hacker profesional. Estos piratas informáticos pueden ser contratados por contrato y pueden realizar una amplia gama de tareas, incluidas pruebas de penetración, evaluaciones de vulnerabilidades e incluso ataques cibernéticos a gran escala. https://www.logpoint.com/en/blog/hacker-as-a-service-what-is-haas/

[118] El *malvertising* (publicidad maliciosa) es una técnica de ciberataque que utiliza publicidad en línea para distribuir malware. Los ciberdelincuentes introducen código malicioso en anuncios aparentemente legítimos, que se muestran en sitios web confiables. Cuando un usuario interactúa con estos anuncios o simplemente carga la página, puede infectar su dispositivo sin saberlo. https://www.incibe.es/aprendeciberseguridad/malvertising

[119] https://www.bitdefender.com/en-us/blog/labs/ai-meets-next-gen-info-stealers-in-social-media-malvertising-campaigns

[120] https://docs.sumsub.com/docs/liveness

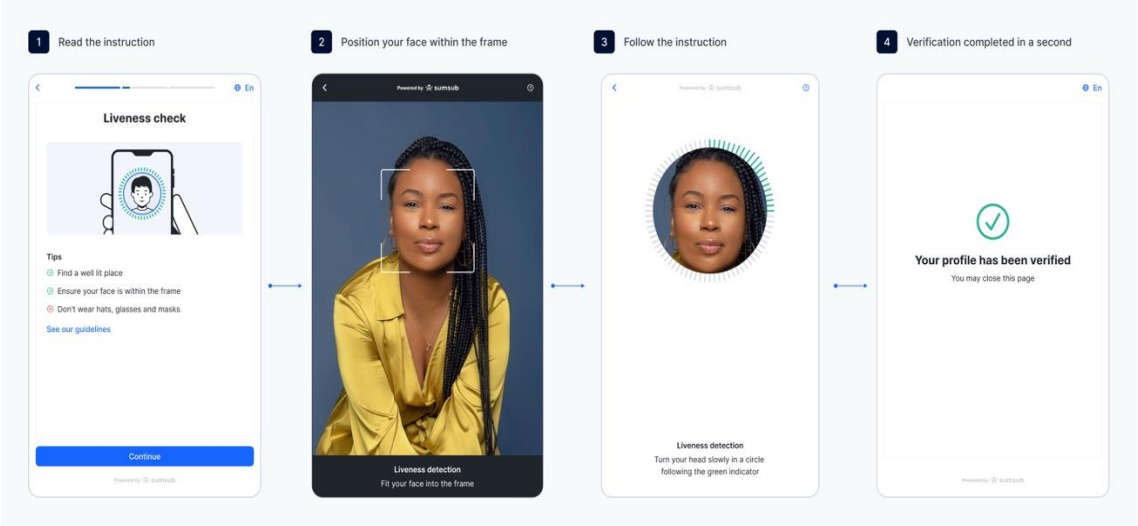


Figura 59. Herramienta de IA para luchar contra el fraude biométrico. Fuente: https://docs.sumsub.com/docs/liveness

La herramienta diseñada está basada en una red neuronal que escanea un rostro y crea un mapa 3D para analizar la imagen. Tarda un segundo en reconocer los rasgos faciales e informar si el solicitante es una persona real (no un *deepfake*, una imagen facial generada por IA, una máscara de papel, una foto de una pantalla, una muñeca o algo similar), no un duplicado y el titular de la cuenta y los documentos.

En los casos en los que las regulaciones requieren un video corto para la verificación en vez de una única foto, el problema de identificación no se resuelve totalmente **porque la IA generativa de tipo multimodal es capaz de generar videos difíciles de diferenciar de uno real**. Algunos sistemas de verificación de identidad biométrica piden al usuario que diga una frase al azar para determinar si es quien dice ser, además de aportar un corto vídeo.

El documento generado por el **NIST** (*National Institute of Standards and Technology*) de Estados Unidos en enero de 2025 (NIST, 2025) pretende dar un paso más proponiendo un conjunto de **directrices voluntarias para mejorar la seguridad, la protección, y la confiabilidad de los modelos de cimentación de doble uso.** En concreto, se centra en la gestión del riesgo de que dichos modelos se utilicen de forma deliberada para causar daños a la seguridad pública o a la seguridad nacional[121]. Será necesario esperar a conocer el **recorrido real de unas directrices voluntarias**, pero, al menos, reflejan la necesidad de actuar de forma consensuada

3.4. Riesgos éticos en la guerra cognitiva impulsada por la IA

El análisis realizado en las páginas anteriores ha presentado **el fenómeno de la desinformación como parte de la guerra cognitiva**, su impacto y las posibilidades de

[121] Los escenarios contemplados incluyen el uso de un modelo para facilitar el desarrollo de armas químicas, biológicas, radiológicas o nucleares; permitir ciberataques ofensivos; y generar contenido dañino o peligroso, como material de abuso sexual infantil (e imágenes íntimas no consentidas de personas reales

reducir el problema (su difusión e impacto en la sociedad) empleando diversas herramientas de IA.

El problema de **la "gobernanza de la IA en el ámbito de la defensa" es un problema perverso** ("*wicked problem*") para RAND (Black et al., 2024) que no se puede "resolver" totalmente, sino que se trata de un continuo "*proceso de creación de sentido en relación con el cambio sociotécnico*", **con soluciones parciales a algunos de los problemas planteados**.

La gobernanza de la IA en la guerra cognitiva se considera un problema perverso de este estilo cuyo uso en el ámbito **militar** genera nuevos problemas potencialmente imprevistos que deberán gestionarse.

> *Un "problema perverso" es aquel que atraviesa diferentes áreas de política, competencias institucionales y fronteras internacionales; uno que sea complejo y multicausal, con consecuencias inciertas y efectos en cascada de segundo y tercer orden que son difíciles de predecir; uno que sea interdependiente con otras cuestiones (por ejemplo, un deterioro de las relaciones entre Estados Unidos y China, o una ruptura del compromiso occidental con Rusia tras la invasión a gran escala de Ucrania por parte de esta última en 2022), lo que dificulta su abordaje de forma aislada; uno que es entendido e interpretado de manera diferente por los diferentes actores, sin consenso sobre cuáles son los problemas más grandes o urgentes, y mucho menos cómo resolverlos; uno que no tenga una solución única, sino que requiera una amplia gama de intervenciones de diferentes actores para cambiar comportamientos, sin que ningún actor posea todas las palancas necesarias; y que, en última instancia, es un desafío político, social y cultural tanto como militar-técnico, que requiere voluntad de compromiso.* Black et al. (2024)

La **creciente relevancia de las noticias falsas** en sociedades digitalizadas, y la sensación de que no es fácil conocer si algo es verdadero o no, se ha convertido en un problema de índole social para la población, que ha empezado a **exigir actuaciones por parte de los gobiernos para frenar su extensión.** Estos, a su vez, ha comenzado a responder con las herramientas habituales a su alcance: incremento de las prohibiciones en los procesos de difusión con mecanismos de censura previa, el endurecimiento de las penas en el ámbito judicial y algunas actuaciones de intervención de medios de comunicación.

Lahman (2024) avisaba de los riesgos de crear el "*Ministerio de la Verdad algorítmico*" y que *"permitir que los gobiernos occidentales desplieguen sistemas respaldados por IA para contrarrestar la amenaza percibida de una "guerra cognitiva" por parte de Rusia y otros actores autoritarios probablemente hará más daño que bien a los derechos fundamentales de los ciudadanos y al sistema democrático liberal que tales medidas supuestamente buscan preservar"*[122].

Estos riesgos reales no pueden impedir a los gobiernos e instituciones públicas tomar acciones encontrando el necesario equilibrio. El informe de la IE elaborado por Pujol y Xuan (2024) propone tres conjuntos de recomendaciones *"con el fin de proteger el orden liberal internacional de los riesgos de la guerra cognitiva y sus tecnologías emergentes":*

- *Intensificar los esfuerzos para comprender y crear conciencia sobre los mecanismos y las implicaciones de la guerra cognitiva en los sectores académico, industrial y de defensa para preparar eficazmente a los ciudadanos y a los responsables políticos.*

[122] https://verfassungsblog.de/european-security-and-the-threat-of-cognitive-warfare/

- *Desarrollar un marco de gobernanza integral para la guerra cognitiva y sus tecnologías emergentes, determinando cómo o en qué casos se aplica el derecho internacional existente, y estableciendo mecanismos claros para la atribución y la rendición de cuentas.*
- *Amplificar la colaboración entre las diversas partes interesadas para capitalizar el potencial de las tecnologías emergentes para abordar las vulnerabilidades de las democracias liberales y protegerlas de los riesgos de la guerra cognitiva".*

La **pregunta abierta a la sociedad** (sobre todo, a la occidental a la que pertenecemos) es conocer hasta qué punto está dispuesta a **renunciar a derechos democráticos en la defensa contra una guerra cognitiva potenciada por la IA** que es cada vez más visible, y cómo encontrar el equilibrio en una situación marcada por el desarrollo de la IA a un ritmo muy rápido que hace obsoletos los marcos regulatorios existentes. **Encontrar ese equilibrio no es sencillo**.

Se han expresado reservas sobre el aumento del uso militar de la IA derivadas del peligro de las emociones en el campo de batalla (Wilson, 2025). De hecho, **las emociones humanas pueden agravar situaciones, asesinatos por venganza o el incumplimiento de órdenes legítimas, mientras que sistemas de IA pueden ser una mejor alternativa porque las emociones humanas "negativas" son reemplazadas por un "campo de batalla más humano"**. Sin embargo, debido a la "razón de la caja negra", sus resultados pueden no ser más predecibles o confiables que el comportamiento humano.

En un contexto más amplio que el de la guerra cognitiva, aunque forma parte de ello, eventos como las cumbres de la **IA Responsable en el Ámbito Militar** (REAIM), por ejemplo, son ilustrativos del **interés en establecer límites éticos al uso de la IA**. Como resultado de dichas reuniones han surgido varios documentos, incluida la *Declaración Política de REAIM de 2023* (Steene and Jenks, 2023) sobre el *Uso Militar Responsable de la Inteligencia Artificial y la Autonomía*, y el documento final de la *Cumbre REAIM* celebrada en Corea del Sur en septiembre de 2024[123] que respalda la puesta en marcha de un **plan de acción**. En todo caso, son aún pasos iniciales. La siguiente *Cumbre REIM* tendrá lugar en España en febrero de 2026.

La ***Declaración de REAIM 2023*** comprende un **conjunto de directrices jurídicamente no vinculantes** que describen las mejores prácticas para el uso responsable de la IA en un contexto de defensa. Entre ellas, garantizar que los sistemas de IA militar sean auditables, tengan usos explícitos y bien definidos, estén sujetos a pruebas y evaluaciones rigurosas a lo largo de su ciclo de vida, y que las aplicaciones de altas consecuencias se sometan a una revisión de alto nivel y puedan ser desactivadas si demuestran un comportamiento no deseado.

Hay autores que, desde la perspectiva legal (Wanyana, 2205), preconizan la urgente necesidad de claridad sobre qué tipo de acciones en el ámbito cognitivo constituyen un **uso prohibido de la fuerza** en virtud del Artículo 2(4) de la Carta de las Naciones Unidas[124], tradicionalmente, interpretado como ataques militares que involucran fuerza física, pero

[123] https://overseas.mofa.go.kr/eng/brd/m_5676/view.do?seq=322676

[124] *"todos los Miembros se abstendrán, en sus relaciones internacionales, de recurrir a la amenaza o al uso de la fuerza contra la integridad territorial o la independencia política de cualquier Estado, o en cualquier otra forma incompatible con los propósitos de las Naciones Unidas".*

claramente ambiguo en este contexto de guerra cognitiva; y, en el caso de que se proceda a ello, **determinar el umbral** que deben cumplir dichas acciones para que se puedan considerar como fuerza prohibida. Debe tenerse en cuenta que el uso de determinadas armas puede crear **secuelas no letales** y no a corto plazo, pero que invalidan al soldado par su función durante periodos prolongados[125].

Otro artículo clave de la ***Carta de las Naciones Unidas*** en este contexto es el Artículo 51, calificando a un "*ataque armado como uso de la fuerza que implique destrucción sustancial de la propiedad, pérdida de vidas o lesiones"*. Desde esta perspectiva, las operaciones cognitivas que causan daño físico a las personas y a la propiedad se considerarían como fuerza prohibida.

Desde nuestro punto de vista, si la interpretación de guerra cognitiva, y su consiguiente clasificación, fuese la de simples operaciones encubiertas o conflictos por debajo del umbral de un conflicto armado debido a su naturaleza no cinética con efectos psicológicos y no físicos, se está olvidando el cambio profundo que se está produciendo acelerado por el uso de la tecnología. Su **capacidad de movilización de la población y acelerador de "conflictos cinéticos"** le confiere una relevancia mayor. Como indica Wanyana (2025) "*El umbral combinado de gravedad e intención también es crítico para determinar cuándo se puede emplear una respuesta militar en defensa propia contra un ataque cognitivo ofensivo."*

En nuestra opinión, el **desarrollo del marco ético para la guerra cognitiva está en sus inicios**. Su desarrollo requiere que todos los actores implicados, incluyendo los ciudadanos, participen en un **diálogo continuo** para abordar los desafíos y oportunidades emergentes que presenta la guerra cognitiva potenciada por la IA, y acordar un marco de actuación flexible y adaptado a la evolución de la tecnología.

5. Evolución tecnológica futura de la guerra cognitiva

5.1. En búsqueda de un nuevo dominio militar basado en el cerebro

En un informe de la OTAN de 2021 (du Cluzel, 2021) se anticipaba que "*el cerebro será el campo de batalla del siglo XXI, y los humanos son el dominio en disputa"*. Esta visión está en el origen del concepto de guerra cognitiva desarrollado en el presente documento. No es un objetivo novedoso. Desde hace miles de años se buscan formas de atacar la moral y la voluntad de lucha del adversario; sin embargo, sus efectos eran limitados porque la tecnología también lo era.

Es ahora, con el desarrollo y convergencia de múltiples tecnologías emergentes cuando surge un **sexto dominio de operaciones** denominado ***"cognitivo"***, que complementa los dominios de operaciones de aire, tierra, mar, espacio y ciber, considerados ya como clásicos en el ámbito de la defensa.

El sexto dominio cognitivo aprovecha que la tecnología existente, incluyendo la IA y las neurociencias, ha alcanzado la **capacidad de penetrar en la mente humana**, de forma invasiva o no, para conseguir sus objetivos de una manera indirecta. Es verdad, que el

[125] En particular, sucede con el uso de láseres con el objetivo de cegar a los soldados, o diversas neuroarmas que pueden generar depresiones o diversos tipos de síndromes incapacitantes.

reconocimiento de la vulnerabilidad de la mente humana ha permitido pensar en lograr a corto o medio plazo la **"*superioridad cognitiva*",** convirtiendo a los investigadores en neurociencias e IA en nuevos actores en el ámbito de la defensa y la seguridad. (McCreight, 2023) (León, 2024).

El dominio cognitivo actúa sobre la forma en la que el cerebro humano actúa ante estímulos recibidos a través de los sentidos (en el ser humano, la vista y el oído fundamentalmente). Pero, hasta hace muy poco tiempo, **no se disponía de tecnologías para actuar directamente sobre el cerebro humano y conseguir "leer" y "escribir en él.** En la última década, sin embargo, han empezado a surgir tecnologías que, si bien, están orientadas hacia el campo médico para ayudar a pacientes con discapacitades, han encontrado también interés en el dominio de la defensa.

Todas las grandes potencias han comenzado a explorar seriamente a dónde puede conducir este sexto dominio, a monitorizar su desarrollo en los próximos años, y a evaluar su uso militar. China y Estados Unidos son, posiblemente, las más decididas en iniciar este proceso con el desarrollo de programas y recursos específicos.

La filosofía del conflicto militar de China está evolucionando para considerar **el cerebro humano como un nuevo dominio operativo**. Imaginan para el futuro un sistema integrado de sistemas en el que los seres humanos se integrarán mejorados cognitivamente por las TIC como algún tipo de **desarrollo evolutivo transhumano**. La idea de *guerra inteligente* fue conceptualizada por China en el informe de la *"Defensa Nacional de China en la Nueva Era"* desde julio de 2019.

También lo está haciendo Estados Unidos. La figura 60 representa también la visión de la *Escuela de Guerra de Estados Unidos* (US *Army War College*) en su informe de 2023 sobre la **evolución del conflicto militar hacia 2035** basado en la convergencia de diversas tecnologías emergentes hacia lo que denomina *"guerra tecno-sentiente"* (Carlson et al., 2023). Los autores del informe consideran que la convergencia entre tecnologías de interfaz cerebro-máquina (BCI), las tecnologías de IA, las tecnologías cuánticas y las técnicas de modificación del genoma humano madurarán suficientemente en diez años como para provocar una nueva revolución en los conflictos militares.

Obsérvese que, de acuerdo con las estimaciones de los autores del informe citado, en solo diez años es casi seguro que tengamos enjambres de drones autónomos con comportamiento inteligente[126]. El segundo elemento, considerado altamente probable, es **disponer de soldados con prestaciones físicas y mentales incrementadas** (exoesqueletos, fármacos o interfaces cerebro máquina) y también se considera muy probable la guerra cognitiva ligada al campo de batalla de los medios sociales.

[126] Transcurridos dos años desde la publicación de ese informe, los acontecimientos de la guerra de Ucrania parecen indicar que, muy probablemente, será mucho antes de esa fecha.

Figura 60. Elementos de la futura guerra "tecno-sentiente". Fuente: Adaptada de: https://media.defense.gov/2024/Aug/30/2003535969/-1/-1/0/FINAL%20TWG%20BOOK%202%20MAY%2023%20V3%202.PDF

En la estimación que hicieron Carlson y sus colegas en 2023, se considera posible disponer en 2035 de "*humanos mejorados*", participando en operaciones militares junto a robots y vehículos no tripulados potenciados por IA con hardware neuromórfico especial. Puede que no se alcance totalmente el objetivo final en 2035 y se retrase a 2040 o más tarde, pero la tendencia es clara.

En nuestra opinión, esta evolución marcará un factor clave para la superioridad militar que anticipa una **batalla tecnológica** entre grandes potencias centrada en el **dominio de la convergencia entre las tecnologías indicadas**. La siguiente sección analiza expresamente el papel de la neurotecnología en la guerra cognitiva

5.2. El papel de la neurotecnología en la guerra cognitiva

Es creciente la importancia de las neurociencias y la neurotecnología en la sociedad; sin embargo, si bien se trata de una tecnología clave, al igual que otras ciencias sociales, para comprender y lograr los efectos pretendidos con la guerra cognitiva, **sus impactos todavía están lejos de poder aplicarse de forma rutinaria en el campo de batalla real.**

Combinada con otras tecnologías, la neurotecnología puede emplearse en el futuro para lograr efectos cinéticos; es decir, para herir, derrotar o destruir a los adversarios. También se pueden emplear para lograr efectos no cinéticos; por ejemplo, creando polarización o disrupción. Desde un punto de vista científico, es un **dominio altamente interdisciplinar** en el que convergen cuatro dominios disociados hasta el siglo XXI:

- **Nanotecnología** (tecnología nanorobótica, nanosensores, nanoestructuras, energía...),
- **Biotecnología** (tecnología biogenómica, Crisp-Cas9, neurofarmacología...),
- **Informática** (tecnologías de la información, informática, microelectrónica, IA...) y
- **Cognética** (tecnología cognitiva, ciencia cognitiva y neuropsicología).

Antes de hacer posible su uso militar, **será necesario comprender mucho mejor que hoy cómo el neuro procesamiento afecta el rendimiento de los combatientes** (y el miedo) y por qué las cosas son tan difíciles de adoptar en un entorno complejo como el militar. Solo entendiendo estas dificultades se podrá desarrollar tecnología eficaz.

Más allá de lo que actualmente es posible, si se consiguiera **conectar directamente en red los cerebros humanos**[127], se podría acelerar la transmisión en tiempo real de datos en el campo de batalla y mejorar las capacidades cognitivas de toma de decisiones de los mandos y facilitar la comunicación entre ellos y sus unidades (Latheef, 2023), (Vakilipour and Fekrvand, 2024). En el futuro, no solo se puede emplear para comunicarse entre dos cerebros humanos, sino también entre diferentes especies animales a través de la cual se podían enviar y recibir comandos motores o sensoriales.

Es probable que algunos países se esfuercen por aprovechar los *neurodatos* obtenidos del funcionamiento del cerebro para obtener ventajas en la capacidad de información, social, legal y militar, a medida que los países competidores invierten en investigación en neurociencias. Como ejemplo, China está buscando el desarrollo de tecnologías de redes cerebrales para mejorar las capacidades cognitivas de los comandantes y la comprensión de la situación en el campo de batalla. Para ello, no solo se necesita una comprensión de la neurociencia, sino también de la cognición, del comportamiento social y de su relación.

Las habilidades cognitivas de un combatiente son extremadamente importantes en el espacio de batalla moderno. Es necesario procesar grandes cantidades de datos/información de forma rápida y precisa y garantizar que la información obtenida de dicho procesamiento sea precisa y fiable. Los avances en la guerra cognitiva ofrecen un medio **para eludir el campo de batalla tradicional con una ventaja estratégica significativa**, que puede utilizarse para transformar radicalmente las sociedades occidentales (Grigsby et al., 2023).

Esta convergencia conduce hacia la potencial **militarización de las neurociencias en la guerra cognitiva.** La neurociencia y la tecnología se incorporan, y se incorporarán cada vez más ampliamente en las estrategias de seguridad nacional, en la recopilación y el análisis de inteligencia y en las operaciones militares proporcionando con ello un considerable incremento del poder; no es necesario actuar de forma invasiva en el cerebro, se puede hacer desde el exterior con técnicas no invasivas (cascos con sensores de actividad cerebral), como se ve en la figura 61, experimentándose en Estados Unidos (Past, 2024).

Este tipo de **neurocascos inteligentes** se han empezado a desarrollar para pilotos militares que necesitan participar en maniobras complejas que requieren preciosos segundos o minutos de su tiempo y atención. Las pantallas táctiles avanzadas o los comandos de voz les permiten realizar esas acciones más rápido, pero requieren respuestas de unos segundos que podría significar la diferencia entre la vida y la muerte.

Una solución futurista es disponer de un casco de alta tecnología que se conecta directamente al cerebro del piloto sin necesidad de implantes quirúrgicos, lo que permite la comunicación bidireccional de cerebro a computadora. **Este casco interpretaría las**

[127] Una interfaz de cerebro a cerebro (BBI), definida como una combinación de métodos de neuroimagen y neuroestimulación para extraer y entregar información entre cerebros directamente sin la necesidad del sistema nervioso periférico, es una técnica de comunicación prometedora.

ondas cerebrales y llevaría a cabo múltiples funciones en microsegundos. Teóricamente, permitiría a los pilotos llevar a cabo tareas sofisticadas, como dirigir drones al combate de forma remota, solo con el poder del pensamiento. O los pilotos podrían transmitir lo que ven a una base militar y recibir información a través del casco a sus cerebros.

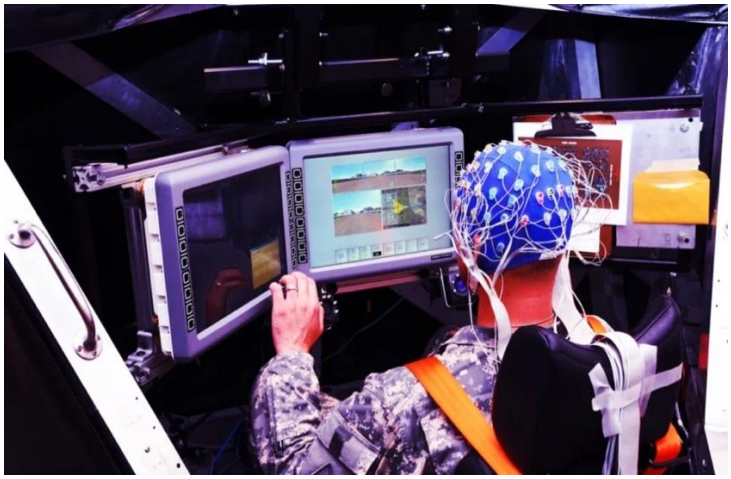


Figura 61. Neurocascos con sensores no invasivos. Fuente: Pabst, 2024

El tipo de neurotecnología no invasiva que se está desarrollando también podría mejorar la capacidad de un cerebro para recordar información clave, aprender nuevas habilidades más rápido e incluso realizar tareas difíciles con la ayuda de un especialista remoto.

Con un objetivo más cercano en el tiempo, el avión de combate ***Tempest*** de sexta generación que el Reino Unido, Japón e Italia están desarrollando debe incluir un **casco de vuelo que monitorice la actividad cerebral del piloto y otros datos biométricos, incluida la respuesta de la piel, la frecuencia cardíaca, la respiración, el seguimiento ocular y el electrocardiograma**. La IA a bordo podría, algún día, usar esta capacidad de monitorización para decidir si intervenir en caso de que el piloto necesite asistencia de emergencia[128].

Históricamente, ha existido un **rechazo intuitivo por el ciudadano al uso de este tipo de tecnología**. El término "*lavado de cerebro*", definido vagamente, surgió en 1950 en el contexto de la "Guerra Fría". En ese momento atrajo la atención y generó preocupaciones sobre los usos futuros de la psicología en la guerra y la vida doméstica, así como el potencial de las nuevas tecnologías para controlar y manipular las mentes humanas[129]. La tecnología, sin embargo, estaba muy lejos de poder actuar directamente en el cerebro, salvo algunos fármacos con efectos generales y no controlables del todo.

Killen (2023) describe algunos de los métodos experimentales ingeniosos y a veces transgresores para estudiar y proponer contramedidas contra los esfuerzos soviéticos de control mental. Detalla cómo estos procedimientos adquirieron una extraña vida propia,

[128] https://www.fanaticalfuturist.com/2025/01/us-militarys-next-gen-fighter-jet-helmet-lets-pilots-control-everything-with-thought/

[129] La frase "*batalla por las mentes de los hombres*" fue utilizada por primera vez por uno de los miembros fundadores de la *CIA (Central Intelligence Agency*) en Estados Unidos y popularizada por el presidente *Dwight Eisenhower*.

escapando de los confines del laboratorio de investigación para convertirse en parte de la contracultura de la década de 1960.

A principios de la década de 2000, estas técnicas **resurgieron en la guerra contra el terrorismo** (Eyre at al., 2023). Actualmente, muchos países están desarrollando y subsidiando activamente la investigación en neurociencia y tecnología bajo agendas de **doble uso o para su incorporación directa en programas militares**.

En la figura 62 puede verse un diagrama de los objetivos planteados por la iniciativa en Estados Unidos (*Brain Initiative*), en la que uno de los socios federales es DARPA (inicialmente con el objetivo de rehabilitación de personal militar y civil). También la figura 56 señala el **papel que juega la IA orientada al desarrollo de nuevos métodos para trabajar con los datos adquiridos** durante el estudio del cerebro.

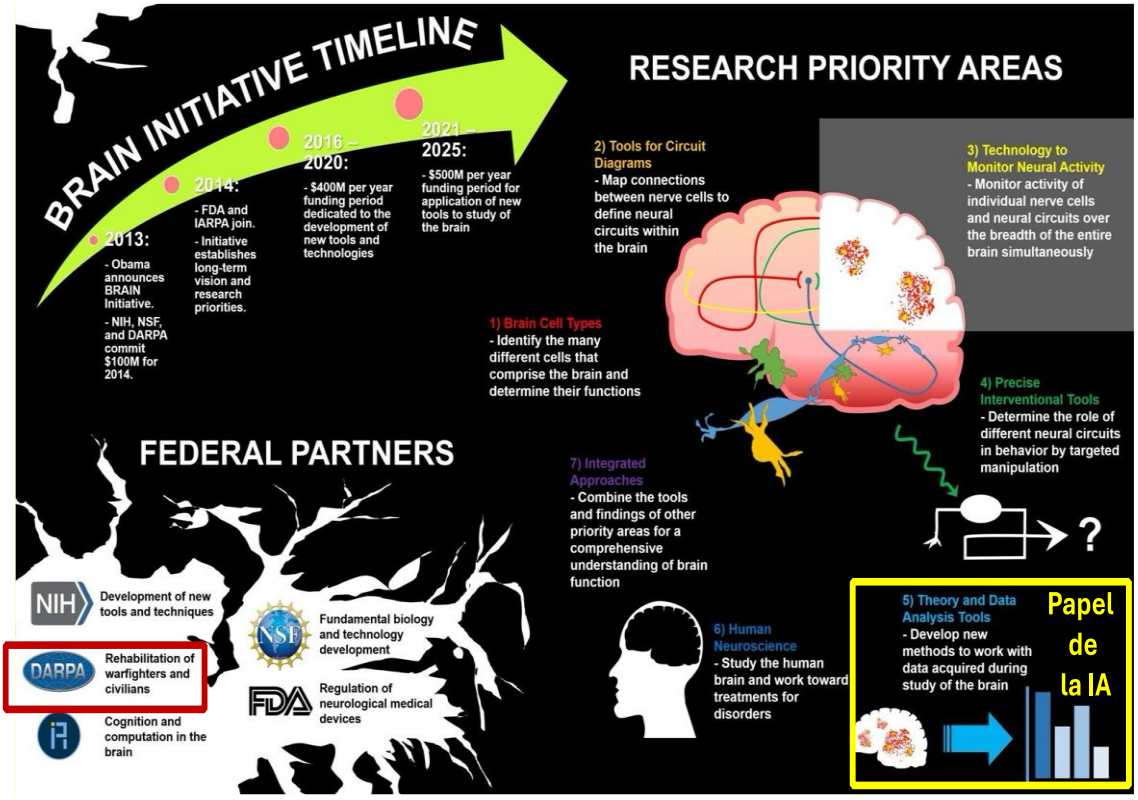


Figura 62. Desarrollo de la Iniciativa sobre el cerebro en Estados Unidos. Fuente: Adaptada de NIH The BRAIN Initiative. https://braininitiative.nih.gov/

Estos esfuerzos internacionales en la investigación sobre el cerebro podrían conducir, a medida que las tecnologías implicadas maduren, y las naciones reaccionen a los nuevos desarrollos tratando de contrarrestar y/o mejorar los descubrimientos de los demás, a una ***"competición para obtener capacidades militares ligada al cerebro".*** Sería el dominio de las neuroarmas.

Las **neuroarmas** pueden definirse como tecnologías y sistemas que podrían mejorar o dañar las capacidades cognitivas y/o físicas de un combatiente o de un objetivo, o atacar de otro modo a las personas o a la infraestructura crítica de la sociedad (Pabst, 2024). Las neuroarmas son herramientas avanzadas que aprovechan la neurociencia y la neurotecnología para influir o alterar el cerebro y el sistema nervioso humanos. Estas

armas pueden **mejorar o degradar las funciones cognitivas, emocionales y motoras de las personas con fines militares y de inteligencia**. Hacerlo a distancia puede ser disruptivo[130]. Las **neuroarmas** incluyen varios tipos: agentes bioquímicos, fármacos neurofarmacológicos, sistemas de energía dirigida e interfaces cerebro-máquina (*Brain Computer Interfaces, BCI*) ya sea de forma invasiva (p.ej. con implantes cerebrales) o no invasiva (p.ej. con sensores en cascos externos) (León, 2024). Veamos algunos ejemplos.

Neurofarmacología

Fármacos que influyen en la cognición y el comportamiento humano mediante la reingeniería de neuronas, moléculas y el sistema nervioso. Algunos ejemplos son:

- *Modafinilo*: fármaco de mejora cognitiva que mejora la conciencia situacional de los soldados, mejora la memoria, crea un estado de hiper alerta y permite a los soldados obtener "superioridad del sueño sobre el enemigo" (Krishnan, 2014). Permite a los soldados permanecer despiertos durante 90+ horas.
- *Captagona*: droga que altera el estado de ánimo que, en dosis suficientes, puede eliminar la empatía y el miedo y desencadenar la violencia y la paranoia. La droga preferida de Al-Qaeda e ISIS.
- *Propranolol:* fármaco que evita que los usuarios desarrollen recuerdos cargados emocionalmente y suprime la memoria y el miedo, disminuyendo así la incidencia del trastorno de estrés postraumático.

Armas de energía dirigida

Un tipo de neuroarmas al que se ha empezado a prestar atención son las denominadas **Neuroarmas de Energía Dirigida**. Utilizan la energía concentrada como micro, radio, ondas acústicas y electromagnéticas para crear armas que van desde letales a no letales. Su interés surgió hace unos años con el denominado "*Síndrome de La Habana"*[131].

Su uso puede ser una **opción en conflictos híbridos**, debido a sus características encubiertas y no atribuibles[132]. Algunos ejemplos son:

[130] Algunos autores como Putric (2022) empiezan a considerar la neuro-guerra y las neuro-armas como parte de un sexto dominio militar (junto a los dominios "clásicos" marítimo, terrestre, aéreo, espacial y ciber) con potencial relevancia en la lucha terrorista y contra terrorista, y en la defensa. En este documento se ha optado por hablar de un sexto dominio cognitivo en el que el uso de la neurotecnología forma parte de las tecnologías empleadas en la medida en la que afectan a la cognición.

[131] En 2016, diplomáticos estadounidenses y oficiales de la CIA en Cuba comenzaron a reportar síntomas de mareos, náuseas y dificultades cognitivas. Las personas que padecían esta dolencia, denominada "Incidentes de Salud Anómalos" por el gobierno de los Estados Unidos, pero más comúnmente conocida como "Síndrome de La Habana", creían que habían sido víctimas de un ataque coordinado. Los científicos recopilaron información y realizaron cálculos que sugerían que los síntomas podrían haber sido causados por un arma de microondas concentrada. Los estudios de seguimiento con resonancias magnéticas encontraron que los pacientes afectados tenían evidencia de lesiones cerebrales traumáticas, lo que sugiere que sus cerebros pueden haber sido dañados físicamente por el incidente. En todo caso, no se ha demostrado nada.

[132] https://cttp.sanford.duke.edu/wp-content/uploads/sites/16/2023/06/2021-Hayes_FSRP-Presentation_v3.pdf

- Sistema activo de denegación. Aprovecha las ondas de frecuencia extremadamente alta para simular la sensación de fuego y "crear una sensación de ardor en la piel" (Skelton, 2012).
- Dispositivo Acústico de Largo Alcance. Utiliza frecuencias de sonido audibles o inaudibles para paralizar, herir o matar enemigos. Capaz de emitir sonidos de hasta 160, 180 y 200 decibelios, suficientes para inducir náuseas, cavitación corporal, contracción muscular, pérdida temporal de la visión, daño pulmonar y de órganos, arritmia cardíaca, deformación celular y muerte

Interfaces Cerebro-Ordenador (Brain Computer Interfaces, BCI)
Tecnologías que conectan las ondas cerebrales con dispositivos externos como computadoras, armas y drones, lo que permite a los humanos controlar máquinas sin estar físicamente presentes.

Dron controlado por la mente. Un casco no invasivo monitoriza las señales cerebrales y las transfiere a un dron, lo que permite a las personas volar el dron con su mente. Hay muchas experiencias en este sentido no solo en Occidente, sino también en China o Rusia (la figura 63 presenta un ejemplo de control de un mini cuadróptero *Geoscan* por la empresa rusa *Neurorobotics,* que ha captado el interés del ministerio de Defensa de Rusia).

Figura 63. Control mental de un dron. Fuente: RIA Novosti/Telegram
https://www.eurasiantimes.com/russian-uses-brain-impulse-to-operate-uav-looks-to-completely/

Estos sistemas de control mental de drones se basan en la captura y análisis de las señales cerebrales empleando sistemas no invasivos como el sistema ***NIMBUS*** (*Neurological Intelligent Monitoring and Brain Utilisation System*) cuyo esquema de funcionamiento se indica en la figura 64.

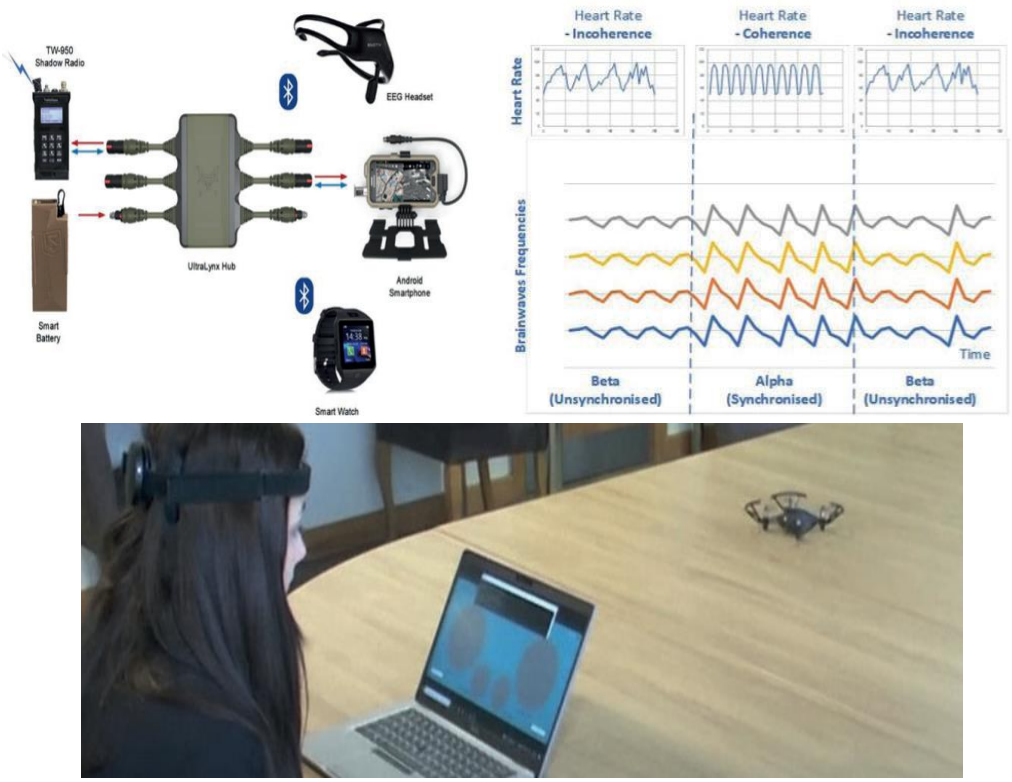


Figura 64. Sistema para el control mental de un dron UltraNimbus. Fuente: https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://neuro8.com/static/media/Nimbus.cd72131bee2e10595c77.pdf&ved=2ahUKEwjQ2fK467SLAxUy7AIHHSvcD1EQFnoECBgQAQ&usg=AOvVaw2Ralwc7EQz2Rlvs7jRMpKG

En septiembre de 2024, la empresa *Neuralink* (propiedad de Elon Musk) anunció el desarrollo de un implante denominado "***Blindsight***" para restaurar la visión a las personas ciegas, al que la ***FDA*** (*Food and Drugs Administration*) en Estados Unidos le ha otorgado la etiqueta de *Dispositivo Innovador*. A diferencia de otros implantes de retina innovadores que envían señales desde una cámara a electrodos que van directamente a los nervios de la retina, **Blindsight es un implante cerebral que estimula directamente el área de la corteza visual del cerebro**. El sistema no emplea el nervio óptico. En la figura 65 se ve el área cerebral estimulada y a la derecha el implante empleado por Neuralink.

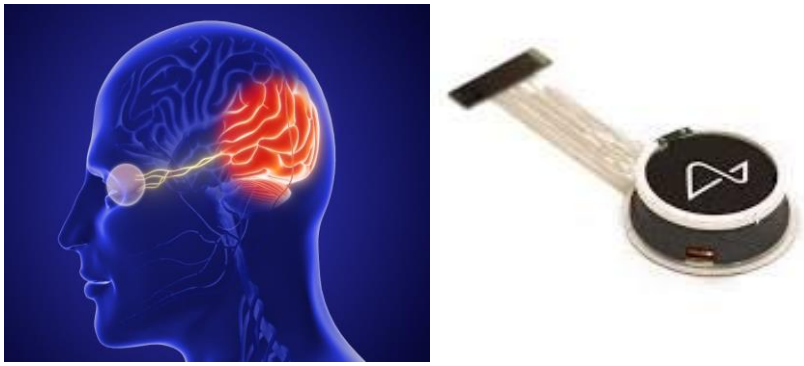


Figura 65. Neuromodulación con Neuralink para personas ciegos. Fuente: https://es.wired.com/articulos/blindsight-el-chip-de-elon-musk-para-devolver-la-vista-a-personas-ciegas

No es el único caso de avance disruptivo. "*Gennaris Bionic Vision System*", ojo biónico desarrollado por investigadores de la Universidad Monash de Australia, en colaboración con Alfred Health consiste en una potente minicámara montada en un casco personalizado que graba su entorno y lo transmite a un implante cerebral inalámbrico.

A diferencia de *Neuralink Blindsight*, que funciona en un implante de matriz de microelectrodos, el ***Gennaris Bionic Eye* utiliza múltiples (hasta 11) implantes diminutos (9 mm) que se pueden colocar en la superficie del cerebro para estimular partes particulares del cerebro**[133]. Cuando se implantan en la corteza visual, estos impulsos eléctricos pueden interpretarse como información visual, proporcionando al receptor formas y contornos básicos que pueden ayudar con la navegación, el reconocimiento de objetos y otras tareas cotidianas básicas[134].

Más allá de que la tecnología madure lo suficiente como para hacer realidad ese objetivo en el sector médico, o en personas sanas, aún muy lejos y nada claro que se consiga a corto plazo[135], lo que Musk también indicaba era su "visión" del futuro: *"Con el tiempo, tendrá el potencial de ser mejor que la visión natural y le permitirá ver en longitudes de onda infrarrojas, ultravioletas o incluso de radar"*. Si eso es así, y se consigue mejorar la capacidad sensorial y cognitiva de las personas, no solo de los pacientes, nos encontraremos en una nueva fase de la guerra cognitiva.

Antes estos avances, y las predicciones de su aceleración, no es extraño que las agencias de investigación militar de las grandes potencias hayan dedicado un interés creciente a financiar, analizar y experimentar con el uso militar de la neurotecnología. En junio de 2022, el Departamento de Defensa de Estados Unidos lanzó la **Iniciativa de Salud Cerebral de los Combatientes** (*Warfighter Brain Health Initiative*)[136] para reunir a las comunidades operativas y médicas en un enfoque más unificado para **optimizar la salud cerebral y contrarrestar las lesiones cerebrales traumáticas del personal militar** con los siguientes objetivos específicos:

- Optimizar el rendimiento cognitivo y físico.
- Identificar, monitorizar y mitigar las exposiciones cerebrales.
- Prevenir, reconocer y minimizar los efectos de las lesiones cerebrales traumáticas.
- Reducir o eliminar los efectos a largo plazo o tardíos de las lesiones cerebrales.
- Avance de la ciencia de la salud cerebral de los combatientes.

La **Unidad de Innovación de Defensa** (*Defense Innovation Unit, DIU*) ha colaborado con empresas como *Garmin* y *Oura Ring* para proporcionar dispositivos portátiles para **rastrear 165 biomarcadores** entre 8.500 soldados, generando análisis predictivos de salud. Además, los investigadores de la Fuerza Aérea están solicitando propuestas de la industria para computadoras portátiles y sensores que puedan detectar y contrarrestar la fatiga y el

[133] https://www.sify.com/science-tech/neuralink-blindsight-and-gennaris-bionic-eye-the-future-of-ophthalmology/

[134] https://www.monash.edu/mada/design-health-collab/projects/product-design-and-wearables/gennaris-bionic-eye

[135] https://spectrum.ieee.org/neuralink-blindsight

[136] https://www.health.mil/Military-Health-Topics/Warfighter-Brain-Health

estrés entre varios grupos, incluidos los combatientes, los profesionales médicos y los socorristas.

Este programa se une al denominado "***The Next-Generation Nonsurgical Neurotechnology (N3)***"[137] financiado por DARPA desde 2021, cuyo objetivo es el de desarrollar interfaces cerebro-máquina bidireccionales de alto rendimiento para militares sanos. Dichas interfaces permitirían disponer de la tecnología necesaria para diversas aplicaciones de seguridad nacional, como el control de vehículos aéreos no tripulados y sistemas de defensa cibernética activos o la combinación con sistemas informáticos para realizar múltiples tareas con éxito durante misiones militares complejas.

Mientras que las interfaces neuronales más efectivas y de última generación requieren cirugía para implantar electrodos en el cerebro, **la tecnología N3 no requeriría cirugía y sería portátil**, lo que haría que la tecnología fuera accesible a una población mucho más amplia de usuarios potenciales. Ya existen neurotecnologías no invasivas, como el electroencefalograma y la estimulación transcraneal de corriente continua, pero no ofrecen la precisión, la resolución de la señal y la portabilidad necesarias para las aplicaciones avanzadas de las personas que trabajan en entornos del mundo real. Desde el punto de vista técnico, **la tecnología N3 supera las limitaciones de la tecnología existente al ofrecer un dispositivo integrado que no requiere implantación quirúrgica**[138]. Algunos métodos experimentados en el programa N3 son solo mínimamente invasivos[139]. Por ejemplo, un equipo de Battelle está estudiando la forma de insertar en el cerebro minúsculos dispositivos denominados "transductores". Estos dispositivos modificarían las señales eléctricas del cerebro para que puedan leerse desde el exterior[140].

Un ejemplo de actuación gubernamental de carácter dual fuera de Estados Unidos y la UE es el **Programa militar de China sobre el cerebro** (*Military Brain Sciences, MBS*)[141]. En la Tabla 2 se pueden ver los objetivos planteados, de los que algunos de ellos (no solo el último de la tabla referido a las neuroarmas) tienen un valor dual.

[137] https://www.darpa.mil/research/programs/next-generation-nonsurgical-neurotechnology

[138] Se pretende que tenga la precisión necesaria para leer y escribir en 16 canales independientes dentro de un volumen de 16 mm3 de tejido neural en 50 ms. Cada canal es capaz de interactuar específicamente con regiones submilimétricas del cerebro con una especificidad espacial y temporal que rivaliza con los enfoques invasivos existentes. Los dispositivos individuales se deben poder combinar para interactuar con múltiples puntos del cerebro a la vez.

[139] https://english.mathrubhumi.com/features/technology/darpa-brain-helmet-military-operations-explainer-1.10067851

[140] En 2023, un equipo de la Universidad Tecnológica de Sídney, en Australia, presentó un dispositivo no invasivo que utiliza un sensor de grafeno para monitorizar la actividad cerebral. Este dispositivo permite a una persona controlar a un perro robótico usando sus pensamientos, seleccionando de un menú de nueve comandos diferentes.

[141] https://www.sciencedirect.com/science/article/pii/S1008127517303188?via%3Dihub

Comprender el cerebro	Comprender los factores de riesgo de las lesiones cerebrales causadas por las actividades militares
Proteger el cerebro	Prevención específica del daño cerebral causado por las actividades militares
Monitorizar el cerebro	Monitorización de la función cerebral a través de nuevas tecnologías y equipos
Dañar al cerebro	Promover la investigación y el desarrollo de armas sonoras, ligeras, explosivas, magnéticas y otros nuevos tipos de armas
Interferir el funcionamiento del cerebro	Causar disfunción cerebral y pérdida de control con métodos "sin humo"
Reparar el cerebro	Lograr la reconstrucción de la función cerebral con tecnología médica novedosa avanzada
Aumentar las capacidades cerebrales	Mejorar el nivel de la función cerebral del personal que realiza tareas especiales
Simular el cerebro	Inteligencia robótica inspirada en el cerebro y predicción de decisiones humanas
Armar el cerebro	Estudiar el armado del cerebro, con las interfaces del cerebro y la máquina como foco

Tabla 2. Programa militar de China sobre el cerebro. Fuente: https://www.sciencedirect.com/science/article/pii/S1008127517303188?via%3Dihub

El **Ejército de China** (*Popular Liberation Army, PLA*) está explorando activamente las operaciones cognitivas como un nuevo dominio de la guerra, posicionándolo junto a la tierra, el mar, el aire, el ciberespacio y el espacio. El PLA está desarrollando actualmente **tecnología portátil y un sistema de apoyo psicológico dedicado para manipular el estado mental de las tropas enemigas** mientras fortalece sus propias fuerzas contra tales esfuerzos.

Una iniciativa notable del PLA es el **Sistema Inteligente de Monitorización Psicológica**. Este sistema subraya la importancia que se concede al bienestar mental de los soldados. Cada vez son más las unidades de PLA que están equipadas con **brazaletes sensores inteligentes que proporcionan datos fisiológicos en tiempo real**. Cuando sea necesario, el sistema puede enviar de inmediato a un "consejero psicológico" para recibir asesoramiento. También permite el **registro continuo de la información facial**, lo que facilita la evaluación en tiempo real de los estados psicológicos de los soldados a través de la retroalimentación y el archivo de datos.

Además, el PLA utiliza **simulaciones de realidad virtual de alto estrés para mejorar la preparación para el combate de los soldados**. Los datos recopilados durante estas sesiones de entrenamiento se pueden utilizar para preparar mejor a los futuros soldados para el combate. En un nivel de tecnología más bajo, el PLA ha establecido salas de entrenamiento antiestrés, campos de entrenamiento psicológico conductual y salas grupales de entrenamiento contra el estrés, que actúan como estaciones espirituales para oficiales y soldados. Estas instalaciones ofrecen consulta psicológica, liberación emocional y ajuste físico y mental.

***Tianjin* se ha convertido en un centro para la innovación en interfaces cerebro ordenador (BCI) de China**. El Laboratorio *Haihe de Interacción Cerebro-Computadora e Integración Humano-Computadora* (脑机交互与人机共融 **海河**实验**室**) en la Universidad de Tianjin es un importante instituto de investigación en BCI con numerosos socios comerciales y militares. Dado el concepto de fusión civil-militar de China, la investigación

hace poca distinción entre las instituciones de investigación, el gobierno, el sector privado y el ejército de China.

Según el *Rastreador de la Universidad de Defensa de China de ASPI* (Impiombato et al., 2204), **la universidad de Tianjin está calificada como de alto riesgo** por su participación en la investigación y el espionaje de defensa, ya que lleva a cabo investigaciones para el Ministerio de Seguridad del Estado (agencia de inteligencia civil) de China, y se agregó a la Lista de Entidades de Estados Unidos en diciembre de 2020.

Para Putric (2022), los **altos costos de las neuroarmas** las posicionan como activos antiterroristas, dado que los terroristas, insurgentes y guerrilleros podrán ser objetivos de ataques neurológicos, pero será difícil que ellos los lleven a cabo. Desde un punto de vista positivo, las neuroarmas reducirán los daños colaterales, incluidas las víctimas civiles.

Las **tecnologías persuasivas emergentes interactúan de manera más intrusiva con los estados neurológicos y fisiológicos de los usuarios**, ofreciendo técnicas cada vez más sofisticadas para influir en el comportamiento humano. Los avances en IA generativa serán el principal impulsor de esta nueva generación de cambio tecnológico, particularmente a través de conceptos como:

- ***Hiperpersuasión***: se refiere a la aplicación de modelos de aprendizaje automático que explotan los sesgos cognitivos para influir en las creencias y los comportamientos a niveles sin precedentes. Los sistemas de IA pueden identificar, dar forma y explotar las demandas de los usuarios y, a continuación, suministrar contenidos personalizados. La *hiperpersuasión* es una forma más sutil y estratégica de explotar las vulnerabilidades de los usuarios individuales para guiar a las personas a tomar decisiones que no son las mejores para ellos.
- ***Hiperacicates*** *(hyper nudging):* se trata de una extensión impulsada por IA de los "acicates" (*nudging)* tradicionales[142] (intentar influir en las personas de una manera predecible, suave y reconocible), que utiliza grandes conjuntos de datos en tiempo real para implementar arrastres dinámicos y personalizados, al mismo tiempo que genera y recopila nuevos datos. Esto hace que los hiperacicates sean significativamente más agresivos, menos reconocibles y potencialmente más manipuladores.

Las implicaciones de la próxima generación de tecnologías persuasivas son profundas. Tienen el potencial de difuminar fundamentalmente la línea entre la libre elección y la coerción subconsciente, y entre el mundo real y el metaverso.

Las tecnologías emergentes, como las interfaces cerebro-ordenador y la realidad extendida, funcionan a la perfección en segundo plano mediante la monitorización y la respuesta directa a los estados y reacciones fisiológicos y neurológicos de los usuarios en tiempo real, y evitan las interfaces de usuario tradicionales. Dichos sistemas pueden ajustar continuamente las estrategias de influencia para mantener los comportamientos deseados y, en última instancia, eliminar la necesidad de un consentimiento activo.

[142] El acicate (*nudging)* digital convencional se basa en un diseño del entorno de elección en el que se presenta la información, que puede ejercer una influencia subconsciente en el resultado. Utiliza elementos de diseño de la interfaz de usuario para guiar el comportamiento de las personas en entornos de elección digital en formularios o preguntas. https://rednmr.wordpress.com/2021/06/10/inteligencia-artificial-y-acicates-hypernudge-y-nudging-de-precision/

En definitiva, la **carrera armamentista prevista por la superioridad cognitiva** da lugar a un **dilema de neuro seguridad** mucho más amplio que el concepto convencional de ciberseguridad: a medida que las neurociencias se imbrican con la seguridad, las prácticas neurocientíficas surgen simultáneamente como una vulnerabilidad central, condicionando su trabajo al cumplimiento de las prácticas de seguridad (Ördén, 2024).

Aún **es pronto para poder evaluar si las tecnologías emergentes mencionadas madurarán lo suficiente para que puedan desarrollarse productos de uso real en operaciones militares**. Con la información disponible en las fuentes abiertas empleadas para la elaboración de este documento, no parece sencillo que su adopción masiva por las fuerzas armadas se produzca a corto plazo.

5.3. Estimación de la aparición de eventos disruptivos relacionados con la guerra cognitiva hasta 2030

La evolución tecnológica relacionada con la guerra cognitiva puede plantearse en **cuatro fases tecnológicas** que se han desarrollado en el tiempo. En nuestra opinión, nos encontramos en la fase 3, con algunas experiencias en proyectos de investigación de la fase 4 que podrán ser efectivos a partir de 2030. Considerando las fechas simplemente orientativas, estas fases son:

- La **Fase 1**, transcurrida desde finales del siglo XVIII hasta el año 1980 (más de 200 años), se refiere al estado de **guerra cognitiva analógica**. En ese periodo no se había producido aún el proceso de digitalización masiva de la sociedad y la guerra cognitiva empleaba **técnicas analógicas de difusión de mensajes** empleando inicialmente la prensa, seguida posteriormente con la adopción social de la radio y la televisión. En esta fase era muy difícil disponer de mensajes "personalizados" para un usuario concreto que modulasen su comportamiento, y los mensajes iban acompañados de discursos políticos en eventos multitudinarios.

  Algunos eventos característicos de esta fase de **guerra cognitiva analógica** son:
  - Democratización del acceso a la prensa libre (en papel)
  - Mensajes comerciales o políticos difundidos en el espacio público (vallas, pasquines, uso de fotocopiadoras, eslóganes, mítines, etc.).
  - Consideración de la "propaganda" como un medio empleado por los gobiernos.
  - Emisoras de radio con programas de opinión y retransmisión en directo de discursos políticos.
  - Emisoras de TV con programas de opinión, entrevistas, retransmisión en directo y documentales
  - Campañas de propaganda oficial gubernamental para modular la opinión pública sobre temas concretos.

- La **Fase 2,** desde 1980 a 2015 (25 años), se refiere al estado de **guerra cognitiva digital** caracterizada por la explosión progresiva de la digitalización de la información, del despliegue de las redes de comunicación fijas y móviles, y el intercambio de mensajes, primeros textuales y luego multimedia, entre millones de personas a través de las redes sociales y las plataformas digitales. Debe tenerse en cuenta que los medios empleados en la fase 1 siguen existiendo.

Algunos eventos característicos de esta fase de **guerra cognitiva digital** son:

- Continuación de las técnicas empleadas en la Fase 1.
- Difusión del correo electrónico como medio de intercambio de información personal o grupal.
- Intercambio de información personal a través de redes físicas y, desde el siglo XXI, en redes móviles con la expansión de la conexión a Internet y los teléfonos móviles inteligentes.
- Expansión de los sitios Webs con interacción asíncrona, blogs, foros, etc. y almacenamiento en la nube.
- Emergencia de las redes sociales de intercambios de texto (p.ej. Facebook, Twitter).
- Análisis de datos (*big data*) por parte de empresas de marketing y gobiernos.
- Emergencia del problema de ciberseguridad sobre la información personal.

- La **Fase 3**, desde 2015 a 2030 (15 años), se refiere al estado de **guerra cognitiva inteligente** en el que **la IA se convierte en la tecnología habilitadora** para un nuevo desarrollo de herramientas de desinformación personalizadas y de modulación del comportamiento de los individuos. Nos encontramos, en nuestra opinión, en medio de esta fase, y el presente informe refleja esta situación. Debe tenerse en cuenta que los medios empleados en las fases 1 y 2 siguen existiendo.

  Algunos eventos característicos de esta fase de **guerra cognitiva inteligente** son:

  - Uso de la **IA generativa** para la construcción y difusión de "narrativas" orientadas a modular la opinión pública o de segmentos específicos.
    - Monomodal (texto) y multimodal (texto, voz, imagen, video).
    - Generación automática de noticias falsas (*fake news*).
    - Herramientas para la detección de información falsa
  - Expansión de las **redes sociales multimedia** con miles de millones de usuarios, combinado con el acceso en todo momento desde dispositivos inteligentes.
    - Papel de los gestores de comunidad ("*community managers*") en el filtrado y difusión de la información,
    - Papel de los *"Influencers"* en la sociedad.
    - Emergencia de bots (redes sociales híbridas)
  - Plataformas digitales con **agentes inteligentes**.
    - Algoritmos de recomendación.
    - Análisis de comportamiento de usuarios.
    - Personalización sencilla de agentes inteligentes.
  - Información contenida en la Web **oscura** (*dark web*) y en la Web **profunda** (*deep web*)
  - Preocupación creciente sobre la **ciberseguridad de la información** (p.ej. alteración de información, robo y rescate de información personal o institucional).
  - **Sistemas de múltiples agentes inteligentes** para la gestión automática de campañas de (des)información.

- La **Fase 4**, desde 2030 en adelante, se refiere al estado que denominaría de **guerra neurocognitiva**, en el que se producirá una **convergencia real entre la IA y la neurotecnología**, auxiliada por nuevos conceptos de procesamiento neuromórfico y

sensores inteligentes implantados (o subcutáneos) que puedan actuar con el cerebro humano para incrementar capacidades.

Algunos eventos característicos de esta fase en relación con la guerra cognitiva podrían ser los siguientes:

- Maduración de **neurotecnologías con la capacidad de interpretar (leer y escribir) la actividad cerebral** empleando técnicas invasivas, mínimamente invasivas o no invasivas en personas sanas (fuera del contexto de la medicina).
- **Comunicación cerebro-cerebro bidireccional** (probablemente limitado en la próxima década a dos personas y no grupal).
- Chips implantados con capacidad de procesamiento neuronal y **neuromodulación funcional del comportamiento humano**.
- **Interfaces más transparentes y ubicuas.** La incorporación a nuestra cotidianeidad de los móviles los hace mucho más potentes para interiorizar las narrativas que dispositivos menos corporeizados.
- **Sistemas holográficos** de comunicación que pueden hacer más convincentes los mensajes y las narrativas.
- **Comunicación extendida** de tipo "conversacional" con no-humanos.
    - humano-animal y humano-robot.

Aunque referidos a una ventana temporal restringida a la presente década (2020-2030), es posible estimar la **aparición de algunos eventos disruptivos de carácter tecnológico** (entendidos como la primera introducción en el mercado de un producto o servicio disruptivo) **relacionados con la guerra cognitiva**.

Su análisis y estimación del impacto social que pueden generar puede permitir la **adopción temprana de medidas tanto tecnológicas como operativas** con el fin de minimizar o maximizar su impacto en función del objetivo que se desee desde una perspectiva dual.

Debe tenerse en cuenta que si se lograse el **abaratamiento sustancial del uso de estas técnicas (**p.ej. cascos, implantes) también se va a **incrementar su ritmo de adopción en personas sanas** para mejorar sus prestaciones en actividades civiles o militares, y, por tanto, su impacto y relevancia en la sociedad (León, 2024).

Los eventos identificados más relevantes son los que se indican en la figura 66. En la parte inferior de la figura se han indicado los relacionados con la **inteligencia artificial** y en la parte superior los relacionados con las **neurociencias**, al considerarse las dos tecnologías habilitadoras más relevantes, dependientes, a su vez, de la microelectrónica[143]. Se han separado en la figura 66 los ya sucedidos hasta 2024 y los previsibles hasta 2030.

[143] Las probabilidades de ocurrencia de los eventos identificados corresponden con opiniones subjetivas del autor.

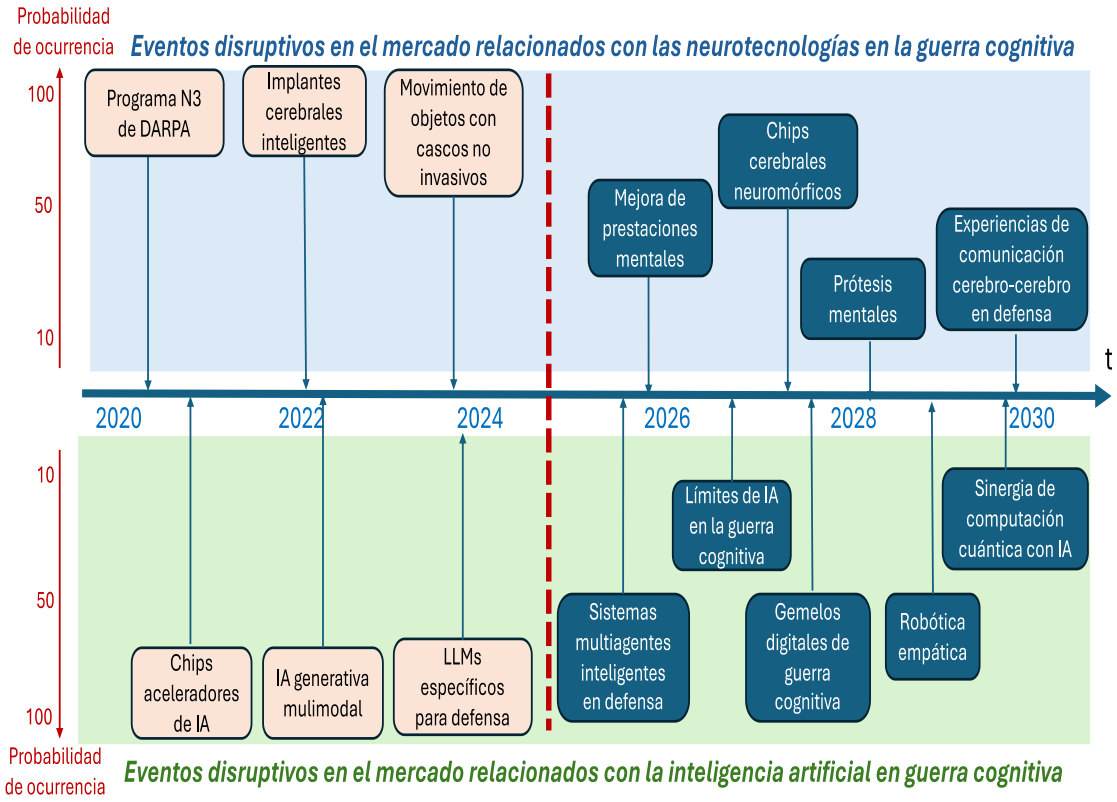


Figura 66. Eventos disruptivos relacionados con la guerra cognitiva. Fuente: elaboración propia

La figura destaca dos **eventos relevantes del uso de la IA en la guerra cognitiva**, con gran impacto, caso de suceder, aunque con probabilidades de materialización no muy elevadas (menores del 50%).

- Hacia 2027 se deberían producir los primeros pasos de un acuerdo en relación con el uso de técnicas de guerra cognitiva estableciéndose ciertos límites en la aplicación de la IA. No concedemos a la aprobación de un acuerdo formal una probabilidad elevada, pero sí es probable el establecimiento por parte de organismos internacionales de determinados "*códigos de conducta*" o directrices de aplicación voluntaria entre países.
- Al final de la presente década se ha incluido como evento la emergencia de la sinergia entre la computación cuántica y la IA. En relación con la guerra cognitiva, esta sinergia permitiría disponer de la capacidad de procesamiento suficientes para disponer de gemelos digitales ciberfísicos del cerebro y poder conocer en tiempo real el efecto de determinados neuro impulsos para la neuromodulación del comportamiento.

También se destacan dos **eventos que parecen relevantes en el uso de la neurotecnología en la guerra cognitiva**, también con probabilidades no elevadas (menores del 50%).

- Emergencia de **prótesis mentales** en personas sanas de manera rutinaria. Se trata, por tanto, de un uso más allá del ámbito médico, sino con el objetivo de incrementar las prestaciones del ser humano. Su extensión a un número considerable de personas sanas dependerá de muchos factores, entre ellos los

regulatorios que limiten su uso, y el coste e invasividad para que sean aceptados por la sociedad.

- **Comunicaciones cerebro a cerebro** aplicadas a operaciones de defensa. Aunque nuestra estimación de que suceda de forma abierta en el mercado tiene probabilidades bajas, estimamos que la tecnología madurará lo suficiente para poder llevar a cabo experiencias en entornos controlados.

En el caso de que todos estos eventos se produzcan, **la tecnología asociada a la guerra cognitiva habrá dado un paso de gigante** y deberemos estar preparados para ello.

6. Conclusiones y recomendaciones de actuación

6.1. Conclusiones

Desde una perspectiva militar puede considerase que, en último término, **todas las guerras tienen una dimensión cognitiva**. Pijpers (2024) decía que *"pensadores como Tucídides o Von Clausewitz argumentan que la esencia de la guerra es someter a un enemigo, es decir, asegurarse de que el actor opuesto (voluntaria o involuntariamente) se convenza de que debe cambiar su comportamiento y actuar bajo nuestra voluntad"*. Esto es lo que las tecnologías aplicadas a la guerra cognitiva pueden lograr.

Para Hagen (2024), *la guerra cognitiva es una nueva forma de pensar sobre cómo los actores usan el poder, influyen en los paisajes geopolíticos y derriban a los oponentes*. Para Claverie (2025) la guerra cognitiva puede, sobre todo, **aprovechar las tecnologías digitales para alterar funciones cognitivas específicas (memoria, atención, comunicación, emociones, etc.) en individuos específicos**. Por ejemplo, el envío de mensajes de texto personalizados a los miembros del parlamento atrapados en una sesión de votación sobre sus familiares, o el envío de fotos de niños muertos a los responsables de la toma de decisiones militares que participan en una operación. **El objetivo es alterar el pensamiento a corto plazo influyendo en la atención, la toma de decisiones y la reacción.**

De igual forma, **también la guerra cognitiva transforma la forma de abordar la ciberseguridad,** ampliando su ámbito para identificar y mitigar las campañas de desinformación, los ataques de ingeniería social y otras formas de manipulación cognitiva; todos ello, activos clave para mantener la democracia y la seguridad nacional. Para Hagen (2024) los esfuerzos en ciberseguridad deben incluir el desarrollo de herramientas y metodologías analíticas capaces de detectar y contrarrestar campañas sofisticadas de desinformación, creada por IA generativa y difundidas a través de redes sociales, y eso concede a la IA un papel clave.

Las tecnologías de IA ofrecen **soluciones prometedoras para la detección y eliminación de información errónea o falsa** mediante procesos de verificación automatizados. La IA permite identificar debilidades en la opinión pública y la información, evitando futuros ataques. Dado el carácter dual de la IA, **las mismas herramientas que pueden prevenir la información falsa también se pueden usar para "atacar" la mente de las personas**.

Dada la relevancia que ha adquirido, y la necesidad de la sociedad de responder a sus crecientes efectos en un contexto de amenazas, conflictos y guerra híbrida como los expuesto, parece que todo sería posible. Como ocurre con el uso de toda tecnología

empleada en defensa, no es así. Los valores y principios que deben informar la **definición de un marco ético consensuado de la guerra cognitiva** es un objetivo urgente.

Del presente documento pueden extraerse las siguientes **conclusiones**:

- La **relevancia de los conflictos híbridos** ha crecido en las últimas dos décadas con implicación de actores estatales y no estatales dotados de sofisticados conocimientos.
- El **carácter multidimensional de los conflictos híbridos** obliga a actuar en varias dimensiones simultáneamente empleando técnicas procedentes de las TIC, especialmente de la IA, de la psicología y de las neurociencias.
- El **uso de herramientas de IA junto a big data** ha supuesto un cambio disruptivo en la forma en la que se detecta la escalada de los conflictos híbridos, se generan campañas de desinformación, y se crean capacidades de mitigación de los riesgos asociados.
- La **IA generativa** ha empezado en los últimos cinco años a adquirir un mayor protagonismo en la generación de narrativas e información multimedia combinada con ciberataques.
- La **convergencia entre la IA y la neurotecnología** abre una nueva fase en la guerra cognitiva que se desarrollará plenamente en la próxima década.
- Todos los estados se han provisto de unidades **especializadas dotadas de tecnología avanzada de IA para luchar de forma integrada en conflictos híbridos** en estrecha interacción con entidades civiles a nivel nacional y enfatizando la cooperación internacional con países aliados.
- La **OTAN** y la **UE** prestan una atención creciente a este fenómeno focalizada en las acciones llevadas a cabo por Rusia y, en menor medida, China e Irán desde 2010 sobre países occidentales.
- Se detecta una **creciente convergencia entre la neurotecnología y la IA aplicada a la guerra cognitiva** que alimenta el desarrollo de neuroarmas de carácter disruptivo.

En definitiva, siguiendo a Pauwels (2024) *"Lo que vemos materializarse ante nuestros ojos es un tipo de guerra polimorfa que fusiona los ciberataques y las operaciones de información y que es librada por los Estados o sus representantes, a veces en situaciones hostiles que no cumplen claramente con el umbral legal de un conflicto armado, en la "zona gris" entre la guerra y la paz"*. Y **es en ese contexto de guerra cognitiva en el que la inteligencia artificial juega, y jugará más en el futuro, un papel preeminente que será necesario dominar.**

Barry et al., (2025) señala que *"los enfoques occidentales de la guerra carecen actualmente de un concepto suficiente para la integración de la IA en el nivel estratégico. Sin un marco rector para la integración estratégica, la adopción de la IA corre el riesgo de dar pasos en falso operativos y tácticos"*. Se requiere un largo proceso para conseguirlo, porque implica **adaptar los procesos organizativos militares**, llevar a cabo una **planificación modulada por la evolución previsible de la tecnología** y acometer la **capacitación continua de los recursos humanos** especializados para aprovechar el poder cognitivo de la IA.

Desde nuestro punto de vista, **España no ha prestado demasiada atención a los medios tecnológicos y operativos necesarios basados en IA para la guerra cognitiva**, en un sentido amplio, centrándose más en los aspectos de ciberdefensa. Las recomendaciones

que se presentan en la siguiente sección tienen como objetivo mejorar esa situación de partida y acelerar la adopción de la IA en este contexto.

6.2. Recomendaciones de actuación

Tras el análisis realizado en las secciones anteriores de este documento, se presenta seguidamente un conjunto limitado de **recomendaciones** para mejorar la posición española frente a la evolución previsible de la guerra cognitiva y desarrollar las defensas necesarias frente a la interferencia extranjera.

R1. Incrementar las capacidades del Mando Conjunto del Ciberespacio (MCEE) para la guerra cognitiva en estrecha sintonía con los desarrollos que se lleven a cabo en la OTAN y la UE, así como con la Estrategia de Seguridad Nacional.

Se trata con ello de que el MCEE disponga de herramientas y métodos avanzados basados en IA que permitan una detección rápida de campañas de desinformación, noticias falsas, y evitar su difusión e impacto.

Asimismo, será necesario asegurar que su evolución aproveche las capacidades de otros actores españoles relevantes en la Estrategia de Seguridad Nacional y, expresamente con el Departamento de Seguridad Nacional, el resto de CERT de referencia nacionales (Instituto Nacional de Ciberseguridad y el Centro Criptológico Nacional) y organismos responsables de la Ciberseguridad y Ciberinteligencia, en el ámbito nacional e internacional.

R2. Incrementar las capacidades en guerra cognitiva del Centro de Inteligencia de las Fuerzas Armadas (CIFAS) desde un enfoque multidisciplinar.

El CIFAS como órgano responsable de facilitar la inteligencia militar precisa asegurar y extender sus capacidades para alertar sobre situaciones internacionales susceptibles de generar crisis que afecten a la Defensa Nacional

R3. Incrementar el número y cualificación tecnológica de las personas implicadas en la guerra cognitiva mediante formación técnica especializada desde un enfoque multidisciplinar.

España deberá disponer de un grupo de personas formadas en el uso de las técnicas y herramientas emergentes necesarias para la guerra cognitiva sometidas a procesos continuados de actualización de conocimientos con la colaboración del sector empresarial y académico.

R4. Extender progresivamente el concepto de operación multidominio a la guerra cognitiva, complementando los dominios de aire, mar, tierra, espacio y ciber en las actuaciones futuras de las FAS.

Integración progresiva del dominio cognitivo en las operaciones multidominio con énfasis en su uso en situaciones de guerra híbrida como parte de la futura Estrategia Tecnológica de la Defensa.

R5. Experimentar en entornos controlados el uso de herramientas de IA relacionadas con las noticias falsas en entornos de defensa y seguridad en cooperación con la industria y los centros académicos.

La rápida evolución de las tecnologías implicadas en la guerra cognitiva aconseja disponer de programas de experimentación y evaluación de nuevas metodologías y herramientas con la participación del sector público y el privado, con especial atención a startups con productos y servicios disruptivos basados en IA, ingeniería social y neurotecnologías.

R6. Monitorizar el desarrollo y uso de herramientas de IA en la guerra cognitiva por parte de las grandes potencias y actualizar con ello periódicamente las estrategias nacionales sobre las mismas.

Extender los observatorios tecnológicos relacionados con la defensa hacia aspectos relacionados con la guerra cognitiva, con el fin de disponer de informes actualizados de vigilancia tecnológica y prospectiva de interés para la defensa y la seguridad nacional y proponer, en función de ello la adquisición de las herramientas más adecuadas.

R7. Incluir el desarrollo y experimentación de tecnologías duales relacionadas con la guerra cognitiva en los programas estatales de investigación e innovación.

Expresamente, se deberá procurar dedicar atención en el contexto del programa COINCIDENTE e incrementar la participación española en los grupos de trabajo de la OTAN relacionados.

R8. Colaborar con los medios de comunicación en su defensa de la veracidad de la información que se difunda a la sociedad, obligando al uso de herramientas de verificación y la concienciación de la sociedad en estos aspectos.

Promover acuerdos con los medios de comunicación para dotarles de medios para luchar contra la desinformación, así como la puesta en marcha de campañas de sensibilización del ciudadano.

R9. Acometer las modificaciones legales y reglamentarias necesarias para dotar a las FCSE (Fuerzas y Cuerpos de seguridad del Estado), a los medios de comunicación y a las entidades públicas relevantes de los recursos humanos y materiales necesarios para ejercer su función en el contexto de la desinformación como parte de la guerra cognitiva.

Inclusión de la evolución de la guerra cognitiva desde las perspectivas tecnológica, operativa y de impacto social en los informes anuales que se confeccionen en el marco de la Seguridad nacional.

7. Referencias

1. Aguado, Juan M., de Haro, Verónica, Gómez de Ágreda, Ángel y Pérez-Escolar, Marta (2024); El ecosistema de la desinformación: actores, estrategias y redes de valor “Desinformación y Defensa. Conflictos híbridos, entorno cognitivo y operaciones de influencia”, Editorial Dikinson, ISBN: 978-84-1070-699-6

2. Ålander, M. (2024). Death by a Thousand Paper Cuts: Lessons from the Nordic-Baltic Region on Countering Russian Gray Zone Aggression. Carnegie Endowment for International Peace. November 14, 2024. https://carnegieendowment.org/research/2024/11/russia-gray-zone-aggression-baltic-nordic?lang=en
3. Asowo, P., Lal, S., Ani, U.D. (2025). An Ensemble Modelling of Feature Engineering and Predictions for Enhanced Fake News Detection. In: Bramer, M., Stahl, F. (eds) Artificial Intelligence XLI. SGAI 2024. Lecture Notes in Computer Science(), vol 15447. Springer, Cham. https://doi.org/10.1007/978-3-031-77918-3_16
4. Aznar, F. (2024). La Guerra, Teoría para comprender los conflictos del Siglo XXI. Ed. El Viejo Topo. ISBN 978-84-19778-95-6.
5. Baldwin, H. (2024). Critical Dual-Use Technologies: Commercial, Regulatory, Societal and National Security Challenges. Economics and Security Committee (ESC). 051 ESC 24 E rev.2 fin. NATO Parliamentary Assembly. 26 August 2024. https://www.nato-pa.int/document/2024-dual-use-technologies-report-baldwin-051-esc
6. Barry, W., Metcalf, C., and Wilcox, B. (2025). Strategic Centaurs: Harnessing Hybrid Intelligence for the Speed of AI-Enabled War. Modern War Institute. January 17, 2025. https://mwi.westpoint.edu/strategic-centaurs-harnessing-hybrid-intelligence-for-the-speed-of-ai-enabled-war/
7. Bebber, R. and Masshal, A. (2024). Cognitive Competition, Conflict, and War: An Ontological Approach. Hudson Institute. May 2024. https://www.hudson.org/defense-strategy/cognitive-competition-conflict-war-ontological-approach-robert-jake-bebber
8. Bernal, A., Cartel, C., Singh, I., Cao, C. y Madreperla, O., (2020). Cognitive Warfare. Johns Hopkins y NATO. Fall 2020. https://www.ejiltalk.org/cognitive-warfare-does-it-constitute-prohibited-force/#:~:text=Bernal%20et%20al,destabilising%20public%20institutions.
9. Beznosiuk, M. (2025). Russian hybrid warfare: Ukraine's success offers lessons for Europe. UkraineAlert. The Atlantic Council. June 5, 2025. https://www.atlanticcouncil.org/blogs/ukrainealert/russian-hybrid-warfare-europe-should-study-ukraines-unique-experience/
10. Black, J. Eken, M., Parakilas, J., Dee, S., Ellis, C., Suman-Chauhan, K., Bain, R.J., Fine, H., Aquilino, M. Lebret, M., et al. (2024). Strategic competition in the age of AI. Emerging risks and opportunities from military use of artificial intelligence. RAND Europe. Sep 6, 2024. https://www.rand.org/pubs/research_reports/RRA3295-1.html
11. Boswinkel, L., Finlayson, N.B., Michaelis, J. and Rademaker, M. (2022) Weapons of mass influence Shaping attitudes, perceptions and behaviours in today's information warfare. The Hague Centre for Strategic Studies. April 2022. https://hcss.nl/wp-content/uploads/2022/04/Weapons-of-Mass-Influence-Information-Warfare-HCSS-2022-V2.pdf
12. Bruze, E., Paskauskas, R.A., Piesarskas, E., Krilavicius, T., Versinskiene, E., Stankeviciute, S., Versinskas, E. y Cardarilli, M. (2021). Quantum as a disruptive technology in Hybrid Threats. JRC Report. JRC126379 Ispra: European Commission, 2021. https://euhybnet.eu/wp-content/uploads/2021/12/Quantum-research-article_final_.pdf
13. Cao, Y., Li, S., Yan, Z., Dai, Y., Yu, P., Sun, L. (2025). A Survey of AI-Generated Content (AIGC). ACM Comput. Surv. 57, 5, Article 125. May 2025, 38 pages. https://doi.org/10.1145/3704262
14. Carlson, L., Arvelo, F., Cain, P., Meyer, S., Hatcher, D. (2023). Techno Warfare 2035. United States Army War College Class. 2023,
15. Casamadrid, F.R. (2025). La autocensura como forma de violencia. Chasqui. Revista Latinoamericana de Comunicación. N.º 158, abril - julio 2025 (Sección Monográfico, pp.

77-94) ISSN 1390-1079 / e-ISSN 1390-924X. Disponible en https://dialnet.unirioja.es/descarga/articulo/10190776.pdf
16. CCN (2024). CCN-CERT IA-04/24: Ciberamenazas y Tendencias. Centro Criptológico Nacional. Octubre de 2024. https://www.ccn-cert.cni.es/es/informes/informes-ccn-cert-publicos/7274-ccn-cert-ia-04-24-ciberamenazas-y-tendencias-edicion-2024/file.html
17. Cialdini, R. (2016). Pre-Suasion: A Revolutionary Way to Influence and Persuade. 2016. Random House International. ISBN 9781847941435
18. Claverie, B., Du Cluzel, F. (2022). Chapter 2 – "Cognitive Warfare": The Advent of The Concept of "Cognitics" in the Field of Warfare. Proceedings of the first NATO scientific meeting on Cognitive Warfare (France) 21 June 2021. March 2022. https://www.researchgate.net/publication/359991886_Cognitive_Warfare_The_Advent_of_the_Concept_of_Cognitics_in_the_Field_of_Warfare
19. Claverie, B. (2025). Cognitive warfare: the new battlefield exploiting our brains. Polytechnique Insights, February 5th, 2025. https://www.polytechnique-insights.com/en/columns/geopolitics/cognitive-warfare-the-new-battlefield-exploiting-our-brains/
20. Colussi I.A. (2025). International Trade Sanctions related to Dual-Use Goods and Technologies. Athens Journal of Law - Volume 2, Issue 4 – Pages 237-252. Octubre 2016. https://doi.org/10.30958/ajl.2-4-3
21. Cominotto, L. (2025). Hybrid Warfare: historical evolution till the Ukrainian conflict. Opinio Juris. March 2025. https://www.opiniojuris.it/articoli-in-lingua/hybrid-warfare-historical-evolution-till-the-ukrainian-conflict/
22. Courtois, B. (2024). Combatting disinformation The AI war. Sopra Steria Next. https://www.soprasteria.com/insights/details/combatting-disinformation-the-ai-war
23. Danley, L. and Colder, B. (2023). The Building Resilience to Cognitive Warfare Technical Exchange Meeting. MITRE Corporation. Public Release 23-00371-18. September 2023. https://www.mitre.org/news-insights/publication/building-resilience-cognitive-warfare-technical-exchange-meeting
24. Deppe. C. and Schaal, G. (2024). Cognitive warfare: a conceptual analysis of the NATO ACT cognitive warfare exploratory concept. Front Big Data. 2024 Nov 1;7:1452129. doi: 10.3389/fdata.2024.1452129. PMID: 39552633; PMCID: PMC11565700. 01 November 2024. https://pmc.ncbi.nlm.nih.gov/articles/PMC11565700/
25. Divon, T., Krutrök, M.E. (2025). The rise of war influencers: Creators, platforms, and the visibility of conflict zones. Platforms & Society Volume 2: 1–18. Last updated: 2025-04-28. https://doi.org/10.1177/29768624251325721
26. du Cluzel, F. (2021). CognitiveWarfare. NATO ACT Innovation Hub, 45. Available at: https://www.innovationhub-act.org/sites/default/files/2021-01/20210122_CW%20Final.pdf.
27. EDA (2023). Beyond 2040 - EDA analysis warns on future warfare trends and technology imperatives for European defence. European Defence Agency. 23 October 2023. https://eda.europa.eu/news-and-events/news/2023/10/23/beyond-2040---eda-analysis-warns-on-future-warfare-trends-and-technology-imperatives-for-european-defence#:~:text=The%20main%20identified%20trends%20from,on%20space%20based%20enabling%20and
28. Edwards, C., Seidenstein, N. (2025). The Scale of Russian Sabotage Operations Against Europe's Critical Infrastructure. The International Institute for Strategic Studies (IISS). August, 2025. https://www.iiss.org/globalassets/media-library---content--migration/files/research-papers/2025/08/pub25-095-the-scale-of-russian-sabotage-operations.pdf

29. EEAS (2023). 1st EEAS Report on Foreign Information Manipulation and Interference Threats. Towards a framework for networked defence. European Union External Action. February 2023. https://www.eeas.europa.eu/eeas/1st-eeas-report-foreign-information-manipulation-and-interference-threats_en
30. EEAS (2024). 2nd EEAS Report on Foreign Information Manipulation and Interference Threats. A Framework for Networked Defence. Report on FIMI threats. European Union External Action. January 2024. https://www.eeas.europa.eu/eeas/2nd-eeas-report-foreign-information-manipulation-and-interference-threats_en#:~:text=Click%20to%20see%20full%20PDF%20document
31. EEAS (2025). 3rd EEAS Report on Foreign Information Manipulation and Interference Threats. Exposing the architecture of FIMI operations. European Union External Action. March 2025. https://www.eeas.europa.eu/sites/default/files/documents/2025/EEAS-3nd-ThreatReport-March-2025-05-Digital-HD.pdf
32. Entrust (2025). 2025 Identity Fraud Report. Entrust Cybersecurity Institute. 2025. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.entrust.com/sites/default/files/documentation/reports/2025-identity-fraud-report.pdf&ved=2ahUKEwjTrprkyKyLAxUXbvEDHdZtAksQFnoECBMQAQ&usg=AOvVaw1nfyXvtWZtrBqLGhyQmFvE
33. Ergo Sum Team, Jackson, T., Richardson R., Ogletree, K., Moss, C. Liberato, R. (2024). Weaponizing Artificial Intelligence Generated Content. United States Army War College. 1st May 2024. https://media.defense.gov/2024/Aug/30/2003535907/-1/-1/0/ERGO%20SUM%20MACHINA%20FINAL%20PRODUCT%20PDF%201.PDF
34. Eyre, H.A., Hynes, W., Ling, G., Occhipinti, J.A., Ayadi, R., Matthews, M., Abbott, R., and Love, P, (2023). From Neuroweapons to 'Neuroshields': Safeguarding Brain Capital for National Security. Rice University's Baker Institute for Public Policy. August 10, 2023, https://doi.org/10.25613/MS7G-YX19
35. Fenstermacher, L., Uzhca, D., Larson, K., Vitiello, C., Shellman, S. (2023). New Perspectives on Cognitive Warfare. Dist A – Reference Number RH-23-124383 Case Reviewer: Keith Klug Case Number: AFRL-2023-2500). https://www.researchgate.net/profile/Laurie-Fenstermacher/publication/371570719_New_perspectives_on_cognitive_warfare/links/64f2510948c07f3da3cdce07/New-perspectives-on-cognitive-warfare.pdf?__cf_chl_tk=NmVomxf8u_kqVUzKHGAfs99yVCHkoqGQgJ6p4ERJqJo-1740918068-1.0.1.1-WLJfio3W8m9NdXOYT5YA4dVOoP2n4tkURO_ousIvEPw
36. Genini, D. (2025). Countering hybrid threats: How NATO must adapt (again) after the war in Ukraine. New Perspectives, 33(2), 122-149. https://doi.org/10.1177/2336825X251322719
37. Grigsby, C., Bridges N., McKinley, R.A., et al. (2023) Developing cognitive neuroscience technologies for defence against cognitive warfare. In: Masakowski YR, Blatny JM (eds) Mitigating and Responding to Cognitive Warfare. NATO Science and Technology Organization. https://apps.dtic.mil/sti/trecms/pdf/AD1200226.pdf
38. Gutiérrez-Caneda, B., & Vázquez-Herrero, J. (2024). Redrawing the Lines Against Disinformation: How AI Is Shaping the Present and Future of Fact-checking. Tripodos, (55), 55–74. https://doi.org/10.51698/tripodos.2024.55.04
39. Hagen (2024). Cognitive Warfare, Cybersecurity, and the AI Challenge. 16 de febrero de 2024. https://www.linkedin.com/pulse/cognitive-warfare-cybersecurity-ai-challenge-raymond-andr%C3%A8-hagen-fvsef/
40. Hasselbach, C. (2025). Arresto por el Nord Stream plantea preguntas delicadas. DW. 22 de agosto, 2025. https://www.dw.com/es/arresto-por-el-nord-stream-plantea-preguntas-delicadas/a-73737714

41. Hoffman, F. (2007). Conflict in the 21st Century: The Rise of Hybrid Wars. Potomac Institute for Policy Studies. December 2007. https://www.academia.edu/22883467/The_Rise_of_Hybrid_Wars
42. Hung, T. C., and Hung, T. W. (2022). How China's cognitive warfare works: a frontline perspective of Taiwan's anti-disinformation wars. J. Glob. Sec. Stud. 7:ogac016. doi: https://doi.org/10.1093/jogss/ogac016
43. Impiombato, D., Attrill, N., Zhang, A., Ryan, F., Allen, B. (2024). Persuasive technologies in China: Implications for the future of national security. Policy Brief. ASPI (Australian Strategic Policy Institute). November 2024. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.aspi.org.au/report/persuasive-technologies-china-implications-future-national-security&ved=2ahUKEwi0zJnF2_2LAxVHT6QEHTrIEm4QFnoECBYQAQ&usg=AOvVaw1dDJDIsdl3AtDcDP85KrIf
44. Jones, S.G. (2025a). The Tech Revolution and Irregular Warfare: Leveraging Commercial Innovation for Great Power Competition. Center for Strategic & International Studies. January 30, 2025. https://www.csis.org/analysis/tech-revolution-and-irregular-warfare-leveraging-commercial-innovation-great-power
45. Jones, S.G. (2025b). Russia's Shadow War Against the West. Center for Strategic & International Studies. March 18, 2025. https://www.csis.org/analysis/russias-shadow-war-against-west
46. JRC (2021). The landscape of Hybrid Threats: A conceptual model. EUR 30585 EN 2021. Joint Research Centre. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC123305/conceptual_framework-reference-version-shortened-good_cover_-_publication_office_1.pdf
47. Jungwirth R., Smith H., Willkomm E., Savolainen J., Alonso Villota M., Lebrun M., Aho A., Giannopoulos G. (2023). Hybrid threats: a comprehensive resilience ecosystem, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/37899, JRC129019. https://www.hybridcoe.fi/wp-content/uploads/2023/04/CORE_comprehensive_resilience_ecosystem.pdf
48. Kalpokas, I. (2024). Post-Truth and Information Warfare in their Technological Context. Applied Cybersecurity & Internet Governance. ACIG, vol. 4, no. 2, 2024. https://doi.org/10.60097/ACIG/190407
49. Kania, E. B. (2022). Minds at War China's Pursuit of Military Advantage through Cognitive Science and Biotechnology. PRISM 8, NO. 3. https://ndupress.ndu.edu/Portals/68/Documents/prism/prism_8-3/prism_8-3_Kania_82-101.pdf
50. Kertcher, C., Zwilling, M. (2025). The meaning of sentiment analysis of UN speeches on the Russia-Ukraine war: a comparative study using VADER and BERT NLP techniques. Front. Polit. Sci., 04 April 2025. Sec. Political Science Methodologies Volume 7 - 2025 | https://doi.org/10.3389/fpos.2025.1546822
51. Khan AN, Ihalage AA, Ma Y, Liu B, Liu Y, Hao Y (2021). Deep learning framework for subject-independent emotion detection using wireless signals. PLoS ONE 16(2): e0242946. February 3, 2021. https://doi.org/10.1371/journal.pone.0242946
52. Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools. Psychological Science in the Public Interest, 21(3), 103-156. https://doi.org/10.1177/1529100620946707
53. Kormych, B., Malyarenko, T, Wittke, C. (2023). Rescaling the legal dimensions of grey zones: Evidence from Ukraine. Global Policy. Volume 14, Issue 3. June 2023 https://doi.org/10.1111/1758-5899.13233
54. Kosal M., y Putney J. (2023). Neurotechnology and international security: Predicting commercial and military adoption of brain-computer interfaces (BCIs) in the United

States and China. Politics Life Sci. 2023 Apr;42(1):81-103. doi: 10.1017/pls.2022.2. PMID: 37140225 https://pubmed.ncbi.nlm.nih.gov/37140225/
55. Lahmann (2024). European Security and the Threat of 'Cognitive Warfare'. Beware of the Algorithmic Ministry of Truth. Verfassungsblog. Europe's Geopolitical Coming of Age: Adapting Law and Governance to Harsh International Realities. 3 noviembre de 2024. https://verfassungsblog.de/european-security-and-the-threat-of-cognitive-warfare/
56. Latheef S. (2023). Brain to Brain Interfaces (BBIs) in future military operations; blurring the boundaries of individual responsibility. Monash Bioeth Rev. 2023 Jun;41(1):49-66. doi: https://doi.org/10.1007/s40592-022-00171-7 . PMID: 36550229.
57. Lavoix. H., Taboy, T., Valantin, J.M. (2025). La IA en las raíces de la guerra: gramática de una nueva geopolítica. Grand Continent. 8 de enero de 2025. https://legrandcontinent.eu/es/2025/01/08/la-ia-en-las-raices-de-la-guerra/
58. Lebret, M., Ogden, T., Black, J. (2024). Cross-cutting technologies in Chinese space activities: Raising the risk of hybrid threats. Hybrid CoE Paper 22. December 2024. https://www.hybridcoe.fi/publications/hybrid-coe-paper-22-cross-cutting-technologies-in-chinese-space-activities-raising-the-risk-of-hybrid-threats/
59. León, G. (2023). Relevancia geopolítica de las tecnologías duales. Consecuencias y oportunidades pare reforzar la soberanía tecnológica de la Unión Europea, UPM Press. ISBN: 978-84-18661-44-0
60. León, G. (2024). Tendencias en las tecnologías para el aumento de capacidades humanas. Academia de las Artes y las Ciencias Militares (ACAMI). 2024. Ed. Astorga: ELC-Ediciones La Crítica. https://www.acami.es/publicacion/el-aumento-de-las-capacidades-humanas-libro/
61. León, G. (2025). Informe Draghi y la defensa y seguridad europeas. Ensayos y Monografías nº 13. Academia de las Ciencias y las Artes Militares. https://www.acami.es/wp-content/uploads/2025/03/Informe-Draghi-perspectiva-europea-de-Defensa-web.pdf
62. López-Garay, M. (2025). Las redes sociales como armas de influencia masiva y la necesidad de una doctrina informativa. Documento de opinión nº 68/2025. Instituto Español de Estudios Estratégicos (IEEE). 15 de septiembre de 2025.
63. Lupiáñez, M. (2024). Cómo hacer frente a un ataque cognitivo: Prototipo de detección de la propaganda y manipulación en operaciones psicológicas dirigidas a civiles durante un conflicto. Revista del Instituto Español de Estudios Estratégicos n.º 22 - Año: 2023 - Págs.: 61 a 93. Mayo 2024. https://publicaciones.defensa.gob.es/media/downloadable/files/links/r/e/revista_ieee_22.pdf
64. Mantua, J. (2023). China's Focus on the Brain Gives it an Edge in Cognitive Warfare. July 6, 2023. https://irregularwarfare.org/articles/chinas-focus-on-the-brain-gives-it-an-edge-in-cognitive-warfare/
65. Marahrens, S., and Schröfl, J. (2024). The Russia-Ukraine Conflict from a Hybrid Warfare Cognitive Perspective. The Defence Horizon Journal. April 25, 2024. https://tdhj.org/blog/post/russia-ukraine-hybrid-cognitive-warfare/
66. Marma, K.J. (2025). Cognitive Warfare: The Invisible Frontline of Global Conflicts. Modern Diplomacy, February 12, 2025. https://moderndiplomacy.eu/2025/02/12/cognitive-warfare-the-invisible-frontline-of-global-conflicts/
67. McCreight, R. (2023). Neuro-Cognitive Warfare: Inflicting Strategic Impact via Non-Kinetic Threat. Small Wars Journal. September 2023. https://smallwarsjournal.com/2022/09/16/neuro-cognitive-warfare-inflicting-strategic-impact-non-kinetic-threat/

68. Miller, S. (2023). Cognitive warfare: an ethical analysis. *Ethics Inf Technol* **25**, 46 (2023). https://doi.org/10.1007/s10676-023-09717-7
69. Murray, R. C. (2021). Hybrid Wars: Technological Advancements and the Generational Evolution of Warfare. Small Wars Journal. 9 September 2021. https://smallwarsjournal.com/2021/09/09/hybrid-wars-technological-advancements-and-generational-evolution-warfare/
70. Nayuni, I.E.S. (2024). Generative AI models for Fake News detection. Cyber Security and Networks Forensic. Insights2Techinfo. https://insights2techinfo.com/generative-ai-models-for-fake-news-detection/?utm_source=perplexity
71. Nikoula, D., McMahon, D. (2024). Cognitive Warfare: Securing Hearts and Minds. Information Integrity Lab. University of Otawa, July 2024. https://infolab.uottawa.ca/common/Uploaded%20files/PDI%20files/InfoLab%20-%20Cognitive%20Warfare,%20Securing%20Hearts%20and%20Minds.pdf
72. NIST (2025). Managing Misuse Risk for Dual-Use 3 Foundation Models. Second Public Draft. NIST AI 800-1 2pd 1 U.S. AI Safety Institute. January 2025 https://doi.org/10.6028/NIST.AI.800-1.2pd
73. Nørgaard, K. & Linden-Vørnle, M. (2021). Cyborgs, Neuroweapons, and Network Command. Scandinavian Journal of Military Studies. 4. 94-107. 10.31374/sjms.86.
74. Ördén, H. (2024). The neuropolitical imaginaries of cognitive warfare. PRIO Volume 55, Issue 6. October 2024 https://doi.org/10.1177/09670106241253527
75. Pandey, S.K. (2023). Navigating the Grey Zone: A Comprehensive Analysis of Hybrid Threats, Asymmetric Warfare, and Their Implications for Coastal and Maritime Security Strategies in the 21st Century. LEX SCRIPTA MAGAZINE OF LAW AND POLICY, VOL-2, ISSUE-4 ISSN-2583-8725. https://www.lexscriptamagazine.com/storage/app/public/uploads/Shivam%20Kumar%20Pandey%20pdf-2.pdf
76. Parnell, T. (2023). Brexit and Disinformation. In: Cap. 12. The Routledge Handbook of Discourse and Disinformation. Routledge November 2023. https://doi.org/10.4324/9781003224495-15
77. Pauwels, E. (2024). Preparing for Next-Generation Information Warfare with Generative AI. CIGI Paper No. 310. Center for International Governance Innovation. December 11, 2024. https://www.cigionline.org/publications/preparing-for-next-generation-information-warfare-with-generative-ai/
78. Pelevina, N., & Salojärvi, V. (2025). YouTube as a narrative battlefield: Brazilian social media influencers and the Russian war in Ukraine. The Communication Review, 1–24. 19 Aug 2025. https://doi.org/10.1080/10714421.2025.2545676
79. Pijlers, P. (2024). Legislation as an Instrument of Cognitive Warfare. En "Aspects of Cognitive Warfare". The Defense Horizon Journal. April 2024. ISBN:978-3-200-10166-1. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://tdhj.org/wp-content/uploads/2024/11/2024_TDHJ-Hybrid-CoE-Cgnitive-Warfare-2024_web-v2.pdf&ved=2ahUKEwjtrJzkjv6LAxW47AIHHegQGC0QFnoECBcQAQ&usg=AOvVaw10msZl5iY4oF6jWFAmal4j
80. Pilati., F., Venturini, T. (2025). The use of artificial intelligence in counter-disinformation: a world-wide (web) mapping. Front. Polit. Sci.. Volume 7 February, 2025. Disponible en: https://www.frontiersin.org/journals/political-science/articles/10.3389/fpos.2025.1517726/full
81. Praks, H. (2024). Hybrid CoE Working Paper 32: Russia's hybrid threat tactics against the Baltic Sea region: From disinformation to sabotage. May 30, 2024. https://www.hybridcoe.fi/wp-content/uploads/2024/05/20240530-Hybrid-CoE-Working-Paper-32-Russias-hybrid-threat-tactics-WEB.pdf

82. Pujol, I. (2024). La guerra cognitiva convierte la mente en un campo de batalla. Revista Española de Defensa. Octubre 2024.
83. Pujol, I., Xuan, Q. (2024). The Battle for the Mind. Understanding and addressing cognitive warfare and its enabling technologies. IE University's Center for the Governance of Change (CGC). 19 abril 2024. https://www.ie.edu/university/news-events/news/deceptive-use-ai-technology-is-turning-mind-main-battlefield-21st-century-according-ie-university/
84. Putric, T. (2022). Neuroweapons: The Future of Warfare. (LAPS-AP/DEMS4709) 2022. https://yourreview.journals.yorku.ca/index.php/yourreview/article/view/40734/36981
85. Putric, T. (2022). The Sixth Domain: Neurowarfare, Neuroweapons, and the Future of Counterterrorism. Revue YOUR Review (York Online Undergraduate Research). https://yourreview.journals.yorku.ca/index.php/yourreview/article/view/40734
86. Querishi, W.A. (2020). Fourth- and Fifth-Generation Warfare: Technology and Perceptions. (Digital Sandiego VOL. 21: 187, 2019). En Notre Dame Journal of International and Comparative Law": 174-178, 2020. https://digital.sandiego.edu/cgi/viewcontent.cgi?article=1293&context=ilj
87. Ramírez, D. (2023). Large Language Models: los nuevos actores de acceso al conocimiento. Instituto Español de Estudios Estratégicos. IEEE. https://www.ieee.es/Galerias/fichero/docs_analisis/2023/DIEEEA86_2023_DAVRAM_Conocimiento.pdf
88. Rauta, W. (2025). Countering state-sponsored proxies: Designing a robust policy. Hybrid CoE Paper No. 23. ISBN 978-952-7591-18-5 (web). February 2025. https://www.hybridcoe.fi/publications/hybrid-coe-paper-23-countering-state-sponsored-proxies-designing-a-robust-policy/
89. Rivera, J.P., Mukobi, B., Reuel , A., Lamparth, M., Chandler, S., Schneider, J. (2024). Escalation Risks from Language Models in Military and Diplomatic Decision-Making. FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency. 05 June 2024. Pages 836 – 898 https://doi.org/10.1145/3630106.3658942
90. Sharp, J. Melrose, J., Madahar, B., et al. (2022). Robustness of Artificial Intelligence for Hybrid Warfare. STO-MP-IST-190. NATO. 2022 https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-190/MP-IST-190-17.pdf
91. Schwarz, E. (2024). The (im)possibility of responsible military AI governance. Humanitarian Law & Policy. December 12, 2024. https://blogs.icrc.org/law-and-policy/2024/12/12/the-im-possibility-of-responsible-military-ai-governance/
92. Sheikh, M. (2025). Top 16 sentiment analysis tools to consider in 2025. Sprout blog. July 23, 2025. https://sproutsocial.com/insights/sentiment-analysis-tools/#:~:text=Sprout%20also%20automates%20analysis%2C%20listening,identify%20common%20trends%20and%20issues.
93. Simmons, D., Raisch, K., Gueller, K., Chun, N., McCray, M. (2024). Back to the Futures. A Look into the Future Through the Lens of the Past. United States Army War College. 9 May 2024. https://media.defense.gov/2024/Aug/30/2003535905/-1/-1/0/BOOK_BACK%20TO%20THE%20FUTURES%20TEAM%20ANDERSON%2020MAY24%201.PDF
94. Soldatov, A. y Borogan, I. (2025). Arsonist, Killer, Saboteur, Spy. While Trump Courts Him, Putin Is Escalating Russia's Hybrid War Against the West. Foreiugn Affairs. March 20, 2025. https://www.foreignaffairs.com/russia/arsonist-killer-saboteur-spy-vladimir-putin-donald-trump?s=EDZZZ005ZX&utm_medium=newsletters&utm_source=fatoday&utm_campai

gn=Arsonist%2C%20Killer%2C%20Saboteur%2C%20Spy&utm_content=20250320&utm_term=EDZZZ005ZX
95. Sprengel, F.C. (2021). Drones in hybrid warfare: Lessons from current battlefields. Hybrid CoE Working Paper 10.
96. Steene, S. b Jenks, C. (2023). The Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy. Articles of War. Nov 13, 2023. https://lieber.westpoint.edu/political-declaration-responsible-military-use-artificial-intelligence-autonomy/
97. Sumsub (2024). Identity Fraud Report 2024-2025. Sum and Substance (UK), 2024. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://sumsub.com/guides-reports/identity-fraud-report-2024-2025/&ved=2ahUKEwiut_nFzqyLAxV5UKQEHXSWOcoQFnoECBAQAQ&usg=AOvVaw1MhWiaR3E8W32xX_feg3r-
98. Takagi, K. (2024). Cognitive Centric Warfare: Modelling Indirect Approach in Future Warfare. Journal of Information Warfare. June 30, 2024. https://www.hudson.org/corruption/cognitive-centric-warfare-modelling-indirect-approach-future-warfare-koichiro-takagi
99. Troianovski, A. (2022). Vladimir Putin habla de 'desnazificar' Ucrania. ¿Por qué? New York Times. 17 de abril de 2022. https://www.nytimes.com/es/2022/04/17/espanol/nazi-ucrania-putin.html
100. Unver, H.A. (2024). Artificial intelligence (AI) and human rights: Using AI as a weapon of repression and its impact on human rights. IN-DEPTH ANALYSIS Requested by the DROI subcommittee. European Parliament. PE 754.450. May 2024. https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA(2024)754450_EN.pdf
101. Vakilipour P, Fekrvand S. (2024). Brain-to-brain interface technology: A brief history, current state, and future goals. Int J Dev Neurosci. 2024 Aug;84(5):351-367. doi: 10.1002/jdn.10334. Epub 2024 May 6. PMID: 38711277.
102. Wanyana, R. (2025). Cognitive Warfare: Does it Constitute Prohibited Force? EJIL: Talk! Blog of the European Journal of International Law. January 30, 2025. https://www.ejiltalk.org/cognitive-warfare-does-it-constitute-prohibited-force/
103. Waterman, S. (2024). As Space Gets More Crowded, Space Force Needs New AI Tools to Keep Up: Experts. Air&Space Forces Magazine. Dec. 6, 2024. https://www.airandspaceforces.com/experts-space-domain-awareness-ussf-ai-tools/
104. Wilbor, D. (2025). The Challenge of AI-Enhanced Cognitive Warfare: A Call to Arms for a Cognitive Defense. Small Wars Journal. 22 January 2025. https://smallwarsjournal.com/2025/01/22/the-challenge-of-ai-enhanced-cognitive-warfare-a-call-to-arms-for-a-cognitive-defense/
105. Wilson, J. (2025). AI, war and (in)humanity: the role of human emotions in military decision-making. Humanitarian Law & Policy. February 20, 2025. https://blogs.icrc.org/law-and-policy/2025/02/20/ai-war-and-in-humanity-the-role-of-human-emotions-in-military-decision-making/
106. Ziemer, C.T., Rothmund, T. (2024). Psychological Underpinnings of Misinformation Countermeasures: A Systematic Scoping Review. Journal of Media Psychology, vol. 36, no. 6, pp. 397–409, Nov. 2024, https://doi.org/10.1027/1864-1105/a000407

ANEXO. Listado de acrónimos

No se incluyen acrónimos relativos a nombres de empresas o universidades, pero sí de organizaciones, programas o proyectos que han sido mencionados en el texto.

1. AIGC: Artificial Intelligence Generated Content
2. API: Application Interface
3. APM: Advanced Persistent Manipulators
4. ASPI: Australian Strategic Policy Institute
5. BBI: Brain to Brain Interface
6. BCI: Brain Computer Interface
7. CCN: Centro Criptológico Nacional
8. CIA: Central Intelligence Agency
9. CIFAS: Centro de Inteligencia de las Fuerzas Armadas
10. CNN: Convolutional Neural Network
11. COINCIDENTE: Cooperación en Investigación Científica y Desarrollo en Tecnologías Estratégicas
12. DARPA: Defense Advanced Research Programs Agency
13. DDoS: Distributed Denegation of Service
14. DIU: Defense Innovation Unit
15. DMA: Digital Markets Act
16. DSA: Digital Services Act
17. EDT: Emerging Disrupting Technology
18. EEAS: European External Action Service
19. FANDANGO: FAke News discovery and propagation from big Data ANalysis and artificial intelliGence Operations
20. FAS: Fuerzas Armadas
21. FDA: Food and Drugs Administration
22. FEI: Foro de Empresas Innovadoras
23. FIMI: Foreign Information Manipulation and Interference,
24. GRU: Russian military intelligence
25. HaaS: Hacking as a Service
26. Hybrid CoE: Hybrid Centre of Excellence
27. HYFUTEC: Hybrid Warfare: Future and Technologies
28. H2020: Horizon 2020
29. IA: Inteligencia Artificial
30. ISAC: Information Sharing and Analysis Centre
31. ISIS: The Islamic State of Iraq and Syria
32. IVERES: Identificación, Verificación y Respuesta.
33. JRC: Joint research Centre (EU)
34. LLM: Large Language Model
35. LSTM: Memoria a Corto y Largo Plazo
36. MaaS: Malware as a Service
37. MBS: Military Brain Sciences
38. MCEE: Mando Conjunto del Ciberespacio
39. MIOP: Military Instrument of Power
40. ML: Machine learning
41. NATO: North Atlantic Treaty Organization
42. NIH: National Institute of Health

43. NIMBUS: Neurological Intelligent Monitoring and Brain Utilisation System
44. NIST: National Institute of Standards and Technology
45. NLP: Natural Language Processing
46. NSB: National Security Bureau
47. N3: Next-Generation Nonsurgical Neurotechnology
48. ONU: Organización de las Naciones Unidas
49. OTAN: Organización del Tratado del Atlántico Norte
50. PLA: Popular Liberation Army (China)
51. PSYOPS: Psychological Operations
52. RA: Realidad Aumentada
53. REAIM: Responsible AI in the Military Domain
54. RTVE: Radio Televisión Española
55. RuMoD: Russian Ministry of Defense,
56. RV: Realidad virtual
57. RuMoD: Russian Ministry of Defense
58. SAR: Sistema de Alerta Rápida (UE)
59. TIC: Tecnologías de la Información y las Comunicaciones
60. UE: Unión Europea
61. UK: United Kingdom
62. VPN: Virtual Private Network
63. VRE: Virtual Reality Environment
64. WEF: World Economic Forum
65. WoS: Web of Sience
66. XAI Inteligencia Artificial Explicable