



COMBATBOTS · INVESTIGACIÓN · ENSAYO N.º 0

Capacidad frente a agresividad como doctrina: Ciber-guerreros de combate

Capacidad sí; agresividad como doctrina, no. Entretenimiento robótico y defensa preventiva.

COMBATBOTS

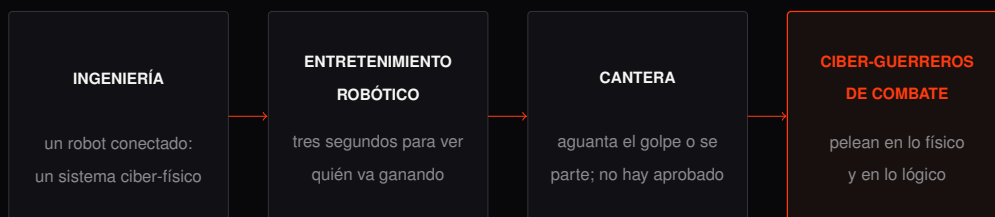
combatbots.ai

Resumen. COMBATBOTS monta arenas de combate robótico donde ya hay público y publica el diseño de sus robots con licencia abierta. Es entretenimiento —53,4 millones de visualizaciones— y a la vez una escuela de ingeniería adversaria: un robot de 500 gramos es un sistema ciber-físico completo, con un adversario al otro lado y una pila de software que también se ataca. A esos robots los llama “ciber-guerreros de combate” —máquinas que pelean en lo físico y en lo lógico—, y eso exige justificarse. Este ensayo lo hace en dos mitades. **Por qué los hacemos:** porque es divertido — el combate robótico combina tecnología, competición, estrategia y espectáculo en enfrentamientos impredecibles, y es lo impredecible, no la agresividad, lo que atrae. **Y cómo se justifica:** como deporte tecnológico, que fomenta la innovación, el aprendizaje y la competición responsable, y a la vez desarrolla capacidades nacionales en robótica e ingeniería en un entorno estratégico, competitivo y controlado. Quien quiera poner límites tiene que saber construir aquello que limita.

Palabras clave: combate robótico, disuasión, armas autónomas, verificación, ciberseguridad ciber-física, formación técnica, política tecnológica, sistemas ciber-físicos

1 Un banco de pruebas con público

LO QUE HACE COMBATBOTS



LO QUE SE LE CRITICA

LO QUE DE ELLO SE SIGUE



NINGUNA DE LAS CUATRO CONCLUYE «NO CONSTRUIR»

quien quiera poner límites tiene que saber construir aquello que limita

Figura 1. El argumento entero, antes de leerlo: arriba lo que hacemos y adónde lleva; abajo las cuatro objeciones que eso levanta y lo que de cada una se sigue.

COMBATBOTS monta arenas de combate robótico donde ya hay gente, publica el diseño de sus robots con licencia abierta y organiza competencias con reglas escritas antes de competir [9][10].

El porqué, dicho sin adornos: **porque es divertido**. Combina tecnología, competición, estrategia y espectáculo en enfrentamientos emocionantes e impredecibles, y es lo impredecible lo que engancha: el único estudio experimental que localizamos encontró que **lo que aumenta el disfrute es la incertidumbre sobre el resultado, no la agresividad** [1]. No sostenemos que esto “canalice” la agresividad, porque la hipótesis de la catarsis está refutada [6]. Y funciona: tres segundos bastan para entender quién va ganando, y los canales propios suman **66.400 seguidores y 53,4 millones de visualizaciones** entre YouTube y TikTok [11]. Hay otras modalidades de competición robótica; ninguna reúne las cuatro cosas a la vez ni sostiene sola una grada.

Y lo que se vende como espectáculo es, por dentro, **un sistema ciber-físico completo a escala de una persona**: mecánica, potencia, control, radio, autonomía, inteligencia

artificial, ciberseguridad y una superficie de ataque, en un objeto que cabe en las manos y en la cabeza de alguien.

Qué es un ciber-guerrero de combate

No es una persona: es el robot. Una máquina con sensores, actuadores y cognición, capaz de pelear en los dos planos — el físico, donde da y encaja golpes, y el lógico, donde ha de defenderse de que le tomen el mando o le engañen la percepción. El término es incómodo a propósito: nombra una capacidad de **dobles usos** en vez de disimularla. Lo que decide para qué sirve no es la máquina, sino la doctrina que la gobierna.

Hasta aquí, un pasatiempo defendible. El problema empieza al nombrar lo que sale de la arena: **ciber-guerreros de combate**, máquinas hechas para pelear en lo físico y en lo lógico. De ahí la pregunta, en dos mitades:

¿Por qué hacemos robots de combate, y cómo se justifica éticamente?

2 Por qué “ciber-guerreros” robóticos y no “pilotos”

Porque **el que combate es el robot, no quien lo manda**. Lo que COMBATBOTS construye son robots: sistemas con sensores, actuadores y cognición, con la autonomía como objetivo declarado y no como añadido. Un piloto maneja una máquina que le dieron hecha; aquí se hace entera, y eso incluye la mitad que no se ve. Un combate entre sistemas ciber-físicos no se decide sólo en lo físico: **un exploit puede inhabilitar al contrincante en segundos, sin tocarlo**.

El radiocontrol de la familia que usan estos robots no cifra la clave de emparejamiento [2]; se puede inyectar un obstáculo inexistente en un LiDAR con una tasa de éxito del 75 % [7], y unas pegatinas sobre un objeto real hacen fallar a un clasificador visual en el 84,8 % de los fotogramas [15]. La pila de software tiene una superficie de ataque publicada [14], catalogada porque en robótica el fallo tiene “physical connection to the world” [24].

Y la capa que más deprisa se abre es la de IA generativa. Los modelos visión-lenguaje-acción traducen imagen e instrucción directamente en comandos motores [4][22], con una consecuencia que conviene decir sin rodeos: **atacar al modelo es atacar al movimiento**. Un parche visual degrada la política hasta anular la tarea [32], y sobre robots gobernados por modelos de lenguaje se ha demostrado el salto completo, con jailbreak de un cuadrúpedo comercial [27].

Un combate ciber-físico se pierde por la pila de software tan fácilmente como por el chasis. La diferencia es que la primera derrota no se ve.

3 Deporte tecnológico

El para qué cabe en una frase que el resto del texto defiende:

El combate robótico se justifica éticamente como **deporte tecnológico**: es divertido, fomenta la innovación, el aprendizaje y la competición responsable, y a la vez desarrolla capacidades nacionales en robótica e ingeniería, formando una cantera tecnológica en un entorno estratégico, competitivo y controlado.

Cada palabra de esa frase es discutible, y la crítica las discute todas.

4 La objeción

La crítica a los combates robóticos recorre desde la posición abolicionista de las campañas contra las armas autónomas [5] hasta la investigación sobre la entrada del capital riesgo en la industria militar [18], y en castellano tiene formulaciones públicas articuladas [13]. Reducida a lo esencial, sostiene cuatro cosas.

- ① **Que las bajas propias iguales a cero eliminan la restricción política que frenaba empezar**: una guerra limpia es una guerra más probable.
- ② **Que el código sustituye al derecho**, quitándole la ambigüedad y la lentitud que son su función.
- ③ **Que la máquina elimina la duda moral** del combatiente, que es el último filtro entre la orden y el disparo, y es también la razón por la que el CICR pide prohibir los sistemas impredecibles y los dirigidos contra personas [12].
- ④ **que el proveedor está fijando la doctrina**: una empresa cuya valoración depende de que haya conflicto no puede decidir cuándo se combate. González llega más lejos y sostiene que los sistemas nacidos de ese ecosistema serán “ineffective,

unpredictable, and unsafe” [18]; es decir, que construir capacidad por esa vía la degrada.

Todo eso es correcto, y sin embargo la crítica se interrumpe justo antes de la pregunta operativa: **¿qué hace España (y Europa) el lunes por la mañana frente a actores que ya están armados en robótica de combate y que ya han decidido no cooperar?** Las dos respuestas habituales —“construir otras narrativas” y “un cambio educativo, perceptivo, intelectual enorme”— no dicen quién hace qué, con qué medios ni en qué plazo. Una narrativa no verifica un tratado. Lo que sigue acepta el diagnóstico y discute la conclusión.

5 Tres doctrinas

El primer movimiento del crítico suele ser implícito: si esto es peligroso, no lo construyas. Pero la abstención unilateral no retira la tecnología del mundo. Retira tu capacidad de decidir cómo se usa. Es importante reconocer que bajo la etiqueta de “combate” o de “armarse” conviven tres posiciones incompatibles:

- **Disuasión.** Capacidad que hace que atacarte no salga a cuenta. Lógica defensiva, y su criterio de éxito es que *no pase nada*. Disuadir por negación —impedir que el atacante logre su objetivo— no exige amenazar con nada ofensivo [31], solo demostrar capacidad.
- **Agresividad táctica.** Pegar primero como ventaja negociadora; lo que Schelling separó de la disuasión bajo el nombre de *compellence* [30]. Su defecto es estructural: si todos la adoptan, el equilibrio es el primer golpe.
- **Depredación.** No negocies, arrasa. La abolición de la política, y lo que la crítica retrata bien.



La postura de COMBATBOTS es la primera columna.

Figura 2. Las tres comparten plataforma, enlace de control y percepción. Lo que las separa está aguas abajo del material. COMBATBOTS aboga por la disuasión como respuesta a qué hacer frente a adversarios ya armados en el combate robótico.

Lo que distingue las tres no es el hardware: es la doctrina, la cadena de mando y las reglas de empeño.

De ahí que “¿construir o no?” esté mal planteado: la pregunta es qué doctrina gobierna lo que se construye, y ésa es una decisión política que no se delega en el fabricante.

Sobre la tesis de la agresividad

Si el adversario ya está armado, ¿no conviene además declararse agresivo, aunque sólo sea para que se lo piense? La respuesta breve es **no**, y no por razones morales. Declararse agresivo obliga a demostrarlo y convierte cada crisis en una prueba de credibilidad; Sagan formalizó el mecanismo como *trampa de compromiso* [28]. Y su premisa —que la firmeza pasada compra credibilidad futura— no resiste la evidencia: los adversarios juzgan por el poder y los intereses en juego *ahora*, no por el historial [26][25]. Es peor estrategia, además de peor ética.

Puestas las cuatro opciones reales una al lado de otra, lo que aparece es que **abstenerse no es la casilla neutral**:

Postura	Qué consigue	Qué pierde	Si el otro no coopera
Abstenerse	No mancharse	Capacidad de verificar, de comprar bien y de regular	El peor resultado: los sistemas los diseña quien no se abstuvo
Comprar fuera	Capacidad operativa rápida	Dependencia; ninguna posición para exigir condiciones	Malo: se tiene el arma y no el criterio
Agresividad declarada	Iniciativa a corto plazo	Obligación de demostrarla en cada crisis	Inestable: el equilibrio es el primer golpe
Disuasión con capacidad propia	Poder responder, verificar y poner límites	Coste, y el riesgo de alimentar la espiral	El menos malo, y el único que no exige acertar en el diagnóstico

Hacen falta capacidades nacionales.

6 Cuatro objeciones, cuatro requisitos

Elegida esa casilla, las cuatro piezas de la crítica dejan de pedir abandono y pasan a pedir requisitos, una por una y en el mismo orden:

- **① Bajas cero.** Ciertamente quitan el freno político — pero abstenerse no lo devuelve: lo devuelve tener con qué verificar y con qué exigir condiciones. Y esa salida puede no tener vuelta; Cohen y Levinthal lo llamaron *lockout* [8].
- **② Código en vez de derecho.** Riesgo real, y por eso una norma vale lo que valga su verificación. Y verificar sin conocimiento no existe: un inspector que no sabe leer el sistema firma un papel. Comprobar *a priori* que un software no esconde una opción autónoma es “virtually impossible” [19].
- **③ La duda moral.** Es la objeción más fuerte y es cierta. De ella no se sigue “no construir”, sino construir sistemas que **preserven el punto donde la duda opera**: un humano en la decisión, con información, con tiempo y con su decisión registrada — que es lo que pide la formulación más citada del control humano significativo [29], nacida en un documento de campaña y no en el derecho positivo [3].
- **④ El proveedor fija la doctrina.** Se le quita teniendo criterio propio, que no es un documento: es gente. Un funcionario que no entiende la pila no escribe el pliego, y un mando que nunca ha visto fallar un sistema autónomo no sabe cuándo desconfiar de él.

Con eso queda contestada la segunda mitad, en sus propios términos y no en otros

más cómodos. **Sí: hacer robots de combate construye capacidad de combate ciber-físico**, y no hay manera de hacerlos sin construirla. Pero esa capacidad es la misma para las tres doctrinas —el material no distingue—, así que lo que hay que declarar no es si se construye, sino qué la gobierna. La nuestra se escribe antes de competir y es comprobable desde fuera: disuasión, y las líneas que cierran este texto.

Quien quiera poner límites tiene que saber construir aquello que limita.

7 Conclusión

Las dos mitades, juntas. **Hacemos robots de combate porque es divertido**, y porque un deporte que engancha por lo impredecible arrastra detrás tecnología, competición y estrategia. **Y se justifica** porque el mismo material sirve a tres doctrinas y lo que decide no es el material: **capacidad sí, agresividad como doctrina no**. Lo que sostiene esa elección no es una preferencia moral sino un argumento de ingeniería: la restricción ética, traducida a requisitos, describe un sistema más difícil de construir que el que se la salta.

Nada de eso lo convierte en otra cosa. No es un arma, no enseña la guerra y no sustituye a la formación jurídica y moral que hace falta para decidir sobre vidas. Es un banco de pruebas con público.

Queda una objeción que no cerramos porque no sabemos hacerlo. Toda capacidad construida empuja al vecino a construir más [20], y la evidencia cuantitativa es incómoda: armarse se asocia con más conflicto, y el estudio que corrige el sesgo de selección de los trabajos previos concluye además que no contribuye a la disuasión [16]. La respuesta clásica —que el dilema se atenúa cuando lo defensivo es distinguible de lo ofensivo [21]— está seriamente discutida [23]. Lo más honesto que sabemos decir es lo de Glaser: dejar de preguntar si armarse causa guerra y preguntar si *este* esfuerzo está bien ajustado a su entorno [17]. Eso concede la evidencia y conserva la posición, a cambio de renunciar a que la capacidad defensiva sea siempre estabilizadora. No lo es.

LA LÍNEA · CRUZARLA NO MATIZA ESTE TEXTO: LO INVALIDA

LO QUE ESTO ES	LO QUE NO ES
+ investigar sistemas ciber-físicos	× fabricar o vender sistemas de armas
+ formar personas	× la agresividad como doctrina
+ publicar diseños, CAD y simulación	× sustituir la formación jurídica y moral
+ entretenimiento civil, con reglas	× afirmar que un torneo enseña la guerra

Figura 3. El alcance, declarado, que es lo único que hace falsable todo lo anterior. La columna de la derecha no es una promesa de intenciones: es la lista de cosas cuya aparición dejaría este texto sin valor y obligaría a reescribirlo, no a matizarlo.

La disuasión defensiva es la única doctrina que no obliga a resolver esa duda a tu favor.

No exige declarar que el adversario ya desertó. Sólo exige que no le salga a cuenta comprobarlo.

Referencias

- [1] Vernon L. Allen and David B. Greenberger. Enjoyment of destruction: The role of uncertainty. *Journal of Nonverbal Behavior*, 4(2):87–96, 1979. doi: 10.1007/BF01006353.
- [2] Jonathan Andersson. Icarus: hijacking DSMx radio-control devices. PacSec 2016, Tokio, 2016. Charla de conferencia; no se localizaron actas formales.
- [3] Article 36. Killer robots: UK government policy on fully autonomous weapons. Policy Paper, Londres, 2013.
- [4] Anthony Brohan, Noah Brown, Justice Carbajal, et al. RT-2: Vision-language-action models transfer web knowledge to robotic control. In *Conference on Robot Learning (CoRL)*, pages 2165–2183, 2023.
- [5] Pere Brunet, Tica Font, and Joaquín Rodríguez. Robots asesinos: 18 preguntas y respuestas. Centre Delàs d'Estudis per la Pau, Barcelona, 2021. El PDF no lleva año explícito; sus referencias internas llegan a diciembre de 2020.

- [6] Brad J. Bushman. Does venting anger feed or extinguish the flame? catharsis, rumination, distraction, anger, and aggressive responding. *Personality and Social Psychology Bulletin*, 28(6):724–731, 2002. doi: 10.1177/0146167202289002.
- [7] Yulong Cao, Chaowei Xiao, Benjamin Cyr, Yimeng Zhou, Won Park, Sara Rampazzi, Qi Alfred Chen, Kevin Fu, and Z. Morley Mao. Adversarial sensor attack on LiDAR-based perception in autonomous driving. In *ACM CCS 2019*, pages 2267–2281, 2019. doi: 10.1145/3319535.3339815.
- [8] Wesley M. Cohen and Daniel A. Levinthal. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1): 128–152, 1990. doi: 10.2307/2393553. Resultado sobre empresas; el traslado al Estado es analogía nuestra.
- [9] COMBATBOTS. Bases y normativa del torneo COMBATBOTS. Categoría 1 lb Plastic. Documento de competición, 3D Printer Party Valladolid, 2026.
- [10] COMBATBOTS. Kit combatbots: Cad, lista de materiales y simulación en ROS 2. <https://combatbots.ai/kits.html>, 2026.
- [11] COMBATBOTS. Tracción y eventos: inventario verificado de canales propios y de torneos celebrados. Documentación interna, [wiki/traccion-y-eventos.md](#), 2026. Comprobado el 2026-07-29; excluye Instagram, no medible sin sesión.
- [12] Comité Internacional de la Cruz Roja. ICRC position on autonomous weapon systems. Ginebra, 2021.
- [13] Eva Cruz, Jaime García Cantero, and José María Lassalle. Palmer Luckey: de manga a armas que aterrorizan. *Cibermillonarios*, episodio 9. Cadena SER, 2026.
- [14] Bernhard Dieber, Benjamin Breiling, Sebastian Taurer, Severin Kacianka, Stefan Rass, and Peter Schartner. Security for the robot operating system. *Robotics and Autonomous Systems*, 98:192–203, 2017. doi: 10.1016/j.robot.2017.09.017.
- [15] Kevin Eykholt, Ivan Evtimov, Earleence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. Robust physical-world attacks on deep learning visual classification. In *IEEE/CVF CVPR 2018*, pages 1625–1634, 2018. doi: 10.1109/CVPR.2018.00175.
- [16] Douglas M. Gibler, Toby J. Rider, and Marc L. Hutchison. Taking arms against a sea of troubles: Conventional arms races during periods of rivalry. *Journal of Peace Research*, 42(2):131–147, 2005. doi: 10.1177/0022343305050687.
- [17] Charles L. Glaser. When are arms races dangerous? rational versus suboptimal arming. *International Security*, 28(4):44–84, 2004. doi: 10.1162/0162288041588313.
- [18] Roberto J. González. How big tech and silicon valley are transforming the military-industrial complex. Technical report, Costs of War Project, Watson Institute, Brown University, 2024.

- [19] Mark Gubrud and Jürgen Altmann. Compliance measures for an autonomous weapons convention. Working Paper 2, International Committee for Robot Arms Control (ICRAC), 2013.
- [20] John H. Herz. Idealist internationalism and the security dilemma. *World Politics*, 2(2):157–180, 1950. doi: 10.2307/2009187.
- [21] Robert Jervis. Cooperation under the security dilemma. *World Politics*, 30(2): 167–214, 1978. doi: 10.2307/2009958.
- [22] Moo Jin Kim, Karl Pertsch, Siddharth Karamcheti, et al. OpenVLA: An open-source vision-language-action model. In *Conference on Robot Learning (CoRL)*, 2024.
- [23] Keir A. Lieber. Grasping the technological peace: The offense-defense balance and international security. *International Security*, 25(1):71–104, 2000.
- [24] Víctor Mayoral-Vilches, Lander Usategui San Juan, Bernhard Dieber, Unai Ayucar Carbajo, and Endika Gil-Uriarte. Introducing the Robot Vulnerability Database (RVD). arXiv:1912.11299, 2019.
- [25] Jonathan Mercer. *Reputation and International Politics*. Cornell University Press, Ithaca, 1996. Verificado por catálogo; no se cita página concreta.
- [26] Daryl G. Press. *Calculating Credibility: How Leaders Assess Military Threats*. Cornell University Press, Ithaca, 2005. Verificado por catálogo; no se cita página concreta.
- [27] Alexander Robey, Zachary Ravichandran, Vijay Kumar, Hamed Hassani, and George J. Pappas. Jailbreaking LLM-controlled robots. arXiv:2410.13691, 2024.
- [28] Scott D. Sagan. The commitment trap: Why the united states should not use nuclear threats to deter biological and chemical weapons attacks. *International Security*, 24(4):85–115, 2000. doi: 10.1162/016228800560318.
- [29] Filippo Santoni de Sio and Jeroen van den Hoven. Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 5:15, 2018. doi: 10.3389/frobt.2018.00015.
- [30] Thomas C. Schelling. *Arms and Influence*. Yale University Press, New Haven, 1966. Verificado por ficha editorial; no se cita página concreta.
- [31] Glenn H. Snyder. *Deterrence and Defense: Toward a Theory of National Security*. Princeton University Press, 1961. Verificado por catálogo; no se cita página concreta.
- [32] Taowen Wang, Cheng Han, James Chenhao Liang, et al. Exploring the adversarial vulnerabilities of vision-language-action models in robotics. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025. arXiv:2411.13587.

Generado desde una única fuente (essay.md) por `research/tools/build_research.py`. La versión web, con el mismo contenido, está en <https://combatbots.ai/research/capacidad-vs-agresividad/>. Licencia CC BY-SA 4.0.