EDICIONES
COMPLUTENSE

Teknokultura. Revista de Cultura Digital y Movimientos Sociales
ISSNe: 1549-2230

KARPETA

La regimentación tecnológica: Inteligencia artificial, fascismo, agresión y sociedad democrática[1]

Pompeu Casanovas Romeu
Instituto de Investigación en Inteligencia Artificial-CSIC, UAB (España)
y La Trobe University (Australia)

https://dx.doi.org/10.5209/TEKN.98266 Recibido: 02 de octubre de 2024 • Aceptado: 15 de febrero de 2025 • **REVISIÓN EN ABIERTO**

ESP **Resumen.** Los sistemas artificiales inteligentes son instrumentos sofisticados que nos ayudan a tomar decisiones y actuar en un mundo híbrido, simbiótico, entre humanos y máquinas. El presente artículo reflexiona sobre la relación entre inteligencia artificial (IA) y las formas políticas, especialmente en su versión autoritaria. Una vuelta al fascismo a través de medios digitales es una de las posibilidades que han sido señaladas por los que han advertido sobre los riesgos del desarrollo de una IA generativa que impulsa la autonomía y adaptabilidad de los sistemas. El artículo examina esta tesis e identifica algunos de los elementos a considerar —el uso bélico de la IA, el comportamiento estratégico de las grandes corporaciones tecnológicas, su mentalidad, y la filosofía práctica subyacente a los modelos de gobernanza. El artículo presta particular atención a la tesis de doctorado en filosofía de Alexander C. Karp, CEO de Palantir Technologies Inc., sobre la agresión y replantea la pregunta sobre el fascismo a partir de la idea de 'sobrecumplimiento' (*over-compliance*) y 'regimentación' (*regimentation*). El fascismo no es el problema en este nuevo entorno; la innovación, el poder y la gobernanza sí lo son.
Palabras clave: élite; ética de la tecnología; Palantir; regimentación; totalitarismo; valores.

ENG Technological regimentation: Artificial intelligence, fascism, aggression, and democratic society

ENG **Abstract.** Artificially intelligent systems are sophisticated tools that help us to decide and to behave independently in a hybrid, symbiotic world in between humans and machines. This article reflects on the relationship between artificial intelligence (AI) and political forms, especially in its authoritarian version. A return to fascism through digital means is one of the possibilities that have been pointed out by those who have warned about the risks of developing a generative artificial intelligence fostering the autonomy and adaptability of systems. The article examines this thesis and identifies some of the elements to consider—warfare, the strategic behaviour of large technological corporations, their mindset, and the practical philosophy underlying governance models. The article pays particular attention to Alexander Karp's PhD dissertation in Philosophy on aggression and reframes the question of fascism from the idea of 'over-compliance' and 'regimentation'. Fascism is not the issue in this new environment; innovation, power, and governance, are.
Keywords: elite; ethics of technology; Palantir; regimentation; totalitarianism; values.

Sumario. 1. Introducción. 2. Una pregunta incómoda. 3. La agresividad como modo de integración social. 4. El teatro de improvisación. 5. Implementación de la ética en los modelos regulatorios. 6. Reflexión final: ¿Un fascismo latente? 7. Disponibilidad de datos. 8. Declaración de uso de LLM. 9. Agradecimientos. 10. Referencias.

Cómo citar: Casanovas Romeu, Pompeu (2025). La regimentación tecnológica: Inteligencia artificial, fascismo, agresión y sociedad democrática. *Teknokultura. Revista de Cultura Digital y Movimientos Sociales 22*(2), 137-149. https://dx.doi.org/10.5209/tekn.98266

[1] Este artículo se enmarca parcialmente en el proyecto financiado SGR-532 (AGAUR) de la Generalitat de Catalunya, y en la Unidad Asociada IIIA-CSIC/IDT-UAB, constituida mediante la Resolución de la Presidencia del CSIC de 21/06/2023.

Wo der heilige Geist ausging, redet man mit mechanischen Zungen.
Donde se ha ido el sagrado espíritu, el hombre habla con lenguas mecánicas.
Theodor W. Adorno, *Jargon der Eigentlichkeit Zur deutschen Ideologie* ([1964,] 1969, p. 12).

1. Introducción

El presente articulo versa sobre fascismo, totalitarismo, autoritarismo e inteligencia artificial (IA). El fascismo es un tema habitual en la literatura sobre filosofía, tecnología y las formas políticas (Aquarone, 1965; Brose, 1987; Paxton 1990). Desde la historia intelectual, la tecnología y la cultura de Weimar ha recibido ya mucha atención (Gay [1968] 2000); Herf, 1986; Dierkes, Knie y Wagner, 1990; Gispen, 1992). También desde la filosofía. Baste recordar las páginas sobre la pérdida de sentido de la ciencia respecto a la técnica en la *Krisis* husserliana ([1936] 2008), especialmente las páginas sobre Galileo Galilei (p. 92), y las advertencias recurrentes de Martin Heidegger contra la comprensión meramente técnica del ser en *Die Frage nach der Technik* ([1954] 1997).

Es en la *Krisis* donde Edmund Husserl introduce el concepto de *Lebenswelt* (mundo de la vida) como el conjunto implícito de presupuestos subjetivos del conocimiento, y es en *Die Frage* donde Martin Heidegger elabora el concepto de *Gestell* (marco, disposición, andamio). Como se verá, deberemos referirnos a estos conceptos más adelante. Ambos filósofos vivieron, de modo ciertamente distinto, el ascenso del nacionalsocialismo. Pero, por lo menos, el segundo pudo ver el desenlace bélico de la energía nuclear, lo que marcó su pensamiento sobre la técnica. En esta época, 'tecnología' se asociaba con la energía atómica, más que con los procesos de información. Igualmente, es frecuente vincular el desarrollo de la tecnología de la información inmediatamente posterior a la Segunda Guerra Mundial con la situación geopolítica de la Guerra Fría, el advenimiento de la sociedad-red y los procesos de globalización. Siempre es recomendable volver al clásico de Herbert A. Simon ([1996] 2006) y a su autobiografía (1991). La RAND Corporation, fundada inmediatamente después de la Guerra con fines de investigación armamentística, tuvo un papel primordial en el nacimiento y desarrollo de la IA.

Proponemos tratar aquí este tema, puesto que ha habido ya reflexiones de distinto signo desde dentro y fuera de ella que relacionan el fascismo como forma política con el reciente desarrollo de Internet, la Web 3.0 y 4.0, y el uso de la IA generativa. Deberíamos delimitar de entrada algunos términos. 'Fascismo' se refiere al movimiento político centroeuropeo que dio lugar a los regímenes nacionalsocialista (Alemania), fascista (Italia) y franquista (España), incluyendo el Movimiento Nacional propulsado por el falangismo. 'Totalitarismo' es un término adoptado por la historia y la ciencia política posterior: apunta a la estructura y organización de los estados que se sitúan fuera de la democracia liberal, modificando el régimen de elecciones y de partidos, y sometiendo la voluntad de la mayoría al poder de una minoría de gobernantes. Los denominados 'estados totalitarios' se extienden así a los regímenes comunistas de la URSS, China y Corea del Norte, pero también las democracias pueden ser totalitarias si comparten los rasgos de falta de control efectivo de los gobernantes por parte de los ciudadanos, como sucede en la Rusia post-soviética. 'Autoritarismo' es un término, elaborado por cierto desde Yale por el politólogo y sociólogo español Juan José Linz, que permite el análisis y comparación de las distintas fases y grados de desarrollo de la autoridad política (en la diferenciación de pre- y post- regímenes totalitarios, por ejemplo). Linz denominaba 'autoritarios' a los regímenes no democráticos y no totalitarios con un pluralismo limitado, y elaboró una tipología en función de estos límites borrosos que permite el análisis y la comparación de la autoridad política en sus diversas fases y grados de desarrollo (Linz, 2000, pp. 158 y ss.). Algunos rasgos de formas políticas dictatoriales —como la falta de garantías jurídicas y procesales para los ciudadanos, el desapego respecto a los derechos humanos, y el personalismo o carisma de un líder— son compartidos por los tres conceptos. Otros, como el racismo, la construcción de una fuerte identidad nacional, y la vocación expansiva imperialista, lo son con matices.

El rasgo que aquí nos interesa es la relación entre el conocimiento (especialmente el científico y tecnológico) y el belicismo, incluyendo el impulso de una actitud agresiva de defensa en la cultura y la ciudadanía. Ésta puede dirigirse también hacia el interior, bien sea en la forma de rechazo hacia los disidentes políticos, los extranjeros o también los inmigrantes con marcadores primarios distintos (inmediatamente perceptibles, como el color de la piel) o que no muestren haber asimilado en su totalidad los rasgos secundarios de la identidad nacional, como la lengua, las formas religiosas o la cultura dominante. Más que al fascismo en su forma histórica, pues, nos referiremos a los rasgos autoritarios que consideramos que se han desarrollado últimamente como cultura digital. Utilizando un término de ciencias de la computación, sugeriremos, al final, comprender estos rasgos como pertenecientes a una cultura política tecnocrática regimentada. Una lectura semejante asume la posición del origen social de las formas políticas, ya observadas por la Escuela de Frankfurt y, a principios de los cincuenta, con diferencias respecto a la posición de Max Horkheimer y Theodor Adorno, por Hannah Arendt. Después del nazismo, el tema del desarrollo de la 'personalidad autoritaria' en el seno de las sociedades democráticas y la cultura de masas fue objeto principal de discusión (Jay, 1973), como lo fue para Arendt la relevancia de la soledad en común, la pobreza recurrente, y la presencia sistemática de la exclusión social como forma de ordenación interna de las sociedades occidentales. No puede explicarse el origen de los estados totalitarios sin aludir a la deriva totalitaria de las sociedades que los originaron (Arendt, 1951).

El resto del artículo se distribuye de la manera siguiente. Las secciones 2 y 3 analizan en detalle la tesis de Alexander Karp (2002) de encontrar «un modelo teórico que haga posible una mejor comprensión de la naturaleza y función de la agresión en la vida cotidiana» [*ein theoretisches Modell skizzieren, das es ermöglicht, die Beschaffenheit und die Funktion von Aggression im Alltag besser zu*

verstehen] (p. 8). Karp es el CEO de Palantir Technologies Inc., quizás la mayor empresa estadounidense especializada en proporcionar tecnología de la información con objetivos de defensa e inteligencia militar. La sección 4 analiza la novedad del modelo de gobernanza propuesto por la corporación. La sección 5, sus propuestas éticas, comparadas con los principios de Asilomar. Finalmente, la sección de conclusiones responde a la pregunta sobre el fascismo, el totalitarismo y la IA, y ofrece algunas reflexiones sobre las limitaciones de abordar este problema y el impacto económico, social y político de la IA mediante los modelos de regulación tradicionales.

2. Una pregunta incómoda

Las preguntas aparentemente más simples son las más difíciles de responder. En el curso de IA de la Universitat Catalana d'Estiu (UCE), una señora, ya mayor, preguntó: «¿hay algún filósofo que esté a favor de la utilización de la IA con fines bélicos?». Aparentemente, todos estamos a favor de la paz y de detener o, al menos, mitigar los conflictos. Ésta es la mentalidad dominante, expresada en múltiples manifiestos y declaraciones públicas. Pero sí que los hay, y hay que tomarse en serio los argumentos para entender qué es lo que está sucediendo. Por ejemplo, en Estados Unidos, Alexander C. Karp, es filósofo por la Universidad Goethe de Frankfurt, con una tesis titulada *La agresión en el-mundo-de-la-vida: La extensión del concepto de agresión de Parsons mediante la descripción de la conexión entre jerga, agresión y cultura* [*Aggression in der Lebenswelt: Die Erweiterung des Parsonsschen Konzepts der Aggression durch die Beschreibung des Zusammenhangs von Jargon, Aggression und Kultur*] (2002). Se trata de un antiguo discípulo de Jürgen Habermas, cuya tesis defendía que las posiciones estratégicas agresivas producían la integración social. La tesis apuntaba a la agresividad y, también, a la sistemática ocultación de la misma en el discurso filosófico. Por eso vuelve a la reflexión de aquellos que la sufrieron en el período de entreguerras y de la segunda Guerra Mundial. Hay que prestarle atención ahora porque sitúa en el plano de la legitimación moral y política el uso de la IA para fines bélicos.

La tesis de Karp es particular porque no tiene estructura de tesis. Sin introducción, ni pregunta de investigación, ni conclusiones. Se abre directamente con lo que quiere defender el autor, proyecta el análisis funcional de Talcott Parsons al discurso de Adorno, y acaba con las reflexiones críticas de Martin Walser sobre la atención obsesiva al propio sentimiento de culpa del pueblo alemán. Es una tesis de lo concreto, que parece salida también de la experiencia del autor en Stanford en los años noventa, cuando discutía 'furiosamente' con su compañero Peter Thiel las mejores estrategias de mercado en la transición a la sociedad digital y las características económicas y políticas que ésta tenía y debía tener (Montfort, 2023). Es una tesis, asimismo, de carácter moral sobre la inadecuación de las políticas progresistas de izquierdas para adaptarse a la sociedad digital. Algo así como '¿qué es lo que está mal en lo que creo?.' En cierto modo, es la crónica de un desencanto. En una de sus últimas entrevistas en el *New York Times*, Maureen Dowd (2024) ilustra la 'sinceridad apasionada' de Karp con una de sus respuestas: «No puedes explicar por qué el país no funciona, pero puedes explicar por qué odias a tu vecino».

La tesis es una revisión del concepto de agresión en la obra de Adorno ([1964] 1969) sobre el *jargon* (jerga) de la autenticidad, comparándola y estructurándola mediante los análisis funcionalistas de Talcott Parsons. En *Certain primary sources and patterns of aggression in the social structure of the Western world* Parsons (1947) estuvo muy atento al ascenso del fascismo y lo que éste significaba para la cultura occidental. Ésta es una de las catorce obras que escribió entre 1935 y 1951 en defensa de la democracia. Uta Gerhardt, a quien por cierto Karp no cita, ha dedicado prácticamente toda su investigación a recomponer este aspecto de la obra de Parsons (Gerhardt, 1999, 2002, 2010, 2017). Theodor Adorno, en *Jargon der Eigentlichkeit* ([1964] 1969) efectuaba una crítica al lenguaje del existencialismo de Heidegger sobre el conocimiento, la subjetivación y las relaciones humanas. Lo calificaba de inservible para el objeto de la liberación humana de la ideología y las formas de alienación. «*Jargon*, la jerga, es la forma apropiada de la no-verdad en la Alemania de estos últimos años» escribe Adorno en la presentación de la segunda edición del libro (1967). *Eigentlichkeit*, autenticidad, es el término que Heidegger compartía con los representantes del régimen fascista. Este tipo de expresiones empezó con el fascismo. Adorno advierte que se dio antes de 1933, «pero sólo después de la guerra, con el rechazo del lenguaje nacionalsocialista, la jerga se volvió omnipresente», y el lenguaje escrito imitó el lenguaje oral de la radio (Adorno [1964] 1969, p. 19; 1973, p. 19). En 1967, en una conferencia sobre la nueva extrema derecha, insistió en este análisis para denunciar el ocultamiento en el lenguaje administrativo y en la aparente aceptación de las formas democráticas sus posiciones existenciales (Adorno, 1967; 2022, p. 25).

Karp acepta parcialmente la crítica de Theodor Adorno a la jerga del existencialismo, particularmente de Karl Theodor Jaspers y de Heidegger, asumida tanto por el lenguaje y el habla del nacionalsocialismo de los años treinta como, progresivamente, por el existencialismo de los años cincuenta. Escribía Adorno en 1964:

> Ante todo, su lenguaje modela el pensamiento de forma que se sienta cómodo con el objetivo de la sumisión, incluso donde cree que se puede resistir. La autoridad de lo absoluto es derrocada por la autoridad absolutizada. El fascismo no era simplemente una conspiración —aunque lo fue—, sino que fue algo que cobró vida en el curso de un poderoso desarrollo social. La lengua le proporciona asilo. Dentro de este cobijo un mal que cuece a fuego lento se expresa a sí mismo como si fuera la salvación (Adorno, [1964] 1969, p. 8; 1973, p. 5).

Según Adorno, el lenguaje proveniente de la filosofía y la teología que pretende expresar la experiencia de lo concreto invade todos los ámbitos de la vida pública, oficial y cotidiana y da lugar a un argot particular: existencial (*existentiell*), en la decisión (*in der Entscheidung*), encargo (*Auftrag*), llamada (*Anruf*),

encuentro (*Begegnung*), diálogo genuino (*echtes Gespräch*), declaración (*Aussage*), compromiso (*Anliegen*), vínculo (*Bindung aus*)... Podríamos añadir otros: 'apertura', 'yo / tú', 'el otro', 'diferencia'... Para Adorno, este lenguaje de lo 'auténtico' expresaría solamente los límites cerrados del sujeto individual (la 'subjetualidad)' sin asumir ni ningún efecto de entendimiento real, ni ninguna referencia a la realidad, ni ningún entendimiento de las causas económicas y políticas que generan la explotación y la alienación humanas. En la línea de la Escuela de Frankfurt, es un sistema, operado con 'lenguas mecánicas' (*mechanischen Zungen*) que fragmentan el sentido y lo adaptan a cada situación particular, una estructura alienada adicta a la autoridad que debe ser descubierta y combatida como ideología desde la dialéctica negativa del propio Adorno, i.e. manteniendo la tensión y contradicciones de lo real sin ninguna síntesis (ni reconciliación) final.

Tal y como había escrito en *Minima Moralia* justo después de la guerra, «de Nietzsche y Bergson, de las últimas filosofías recibidas, no queda más que el turbio antiintelectualismo en nombre de una naturaleza secuestrada por sus apologetas» (Adorno, 1951, párr. 120, 2006, p. 189). Esto no es más que otra manifestación de la tesis profunda de Adorno sobre la pérdida de control del individuo sobre sí mismo y de la colectividad sobre la sociedad, expresada tantas veces en su obra. Por ejemplo: «El fin objetivo de la humanidad es sólo otra expresión para referirse a lo mismo. Y significa que el individuo en cuanto individuo, en cuanto representante de la especie hombre, ha perdido la autonomía con la que poder hacer realidad la especie» Adorno ([1951] párr. 17; 2006, p. 35).

3. La agresividad como forma de integración social

En su tesis doctoral, Karp se planteó conceptualmente la 'integración funcional' de la agresividad, sin combatirla directamente, aceptándola como un hecho natural y otorgándole un papel incentivador, de guía (*driver*) y capacitador (*enabler*) de la conducta. Es decir, como componente clave de la integración social, incluyendo los elementos de irracionalidad y anomia que pueden afectar a la construcción compleja de la identidad. Y, lo que resulta más interesante, incorporando en ella el conflicto, la reflexividad, y las tensiones de culpa que provoca la memoria. Creemos que es mejor no prejuzgar desde posiciones políticas preconfiguradas. Es una obra que hay que leer empezando desde el final y no desde el principio, porque es la reflexión sobre Martin Walser y el holocausto la que proporciona el contexto alemán para entenderla, ya desde el primer párrafo:

> Este trabajo comenzó con la observación de que algunas afirmaciones poseen un efecto impulsor [*Äußerungen triebentlastende Wirkung haben*], no a pesar de su obvia irracionalidad sino debido a ella. Afirmaciones que son claramente auto-contradictorias ofrecen a los actores la oportunidad de reconocer formalmente el orden normativo de su entorno cultural y al mismo tiempo expresar deseos que se desaprueban y que violan las reglas de ese orden. Esto no genera ninguna sanción cultural ni social. Al contrario: estas afirmaciones solidifican los procesos de integración al hacer que la integración y sus costes sean psicológicamente soportables. Siguiendo a Adorno, denomino jerga [*Jargon*] a estas afirmaciones. La jerga no es solamente una mentira-de-vida [*Lebenslüge*], sino un tipo especial de mentira-de-vida. No sólo alivia, sino que integra al hablante en el círculo de quienes pertenecen a él. Mediante la jerga se edulcora el presente, se lo orienta hacia el futuro y, por lo tanto, se lo torna aceptable (Karp, 2002, p. 2).

Cuatro años antes, en 1998, con motivo de la concesión del Premio de la Paz del Libro Alemán, el escritor Martin Walser realizó un discurso de aceptación que convulsionó al público que lo escuchó y al que lo conoció a través de múltiples reproducciones en la prensa y los medios de comunicación. En su discurso *Erfahrungen beim Verfassen einer Sonntagsrede,* Walser compartía la experiencia del cansancio que le provocaban las múltiples y continuas alusiones al Holocausto en Alemania y la desazón que sentía ante las manifestaciones oficiales de culpabilidad. «[...] Ninguna persona seria niega Auschwitz; nadie que esté en su sano juicio duda sobre el horror de Auschwitz; pero cuando este pasado me asalta cada día en los medios de comunicación, noto que algo dentro de mí se rebela contra esta presentación incesante de nuestra desgracia» (Walser, 1998).

Walser no negaba la gravedad de lo ocurrido ni su impacto en la sociedad alemana después de la guerra. Otros escritores y filósofos de gran relevancia como Heinrich Böll, Günter Grass —con quien polemizó— y Habermas no tenían este sentimiento y habían centrado en gran parte su obra en el nacionalsocialismo y sus efectos. Walser solamente reclamaba el derecho a la intimidad para calibrar estos efectos en la vida cotidiana como actos de conciencia, y no creía que el dolor y la culpa pudieran ser objeto de continua manifestación pública. Si así sucedía, se desvirtuaba el acto por su misma presentificación en una memoria colectiva a quien se negaba la posibilidad de recuperar la relación normalizada consigo misma. En definitiva, se instrumentalizaba la conciencia personal con fines distintos al de una expiación soportable. Las críticas no se hicieron esperar y algunas fueron muy agrias. Especialmente la del antiguo Presidente del Consejo Judío, Ignatz Bubis, quien le acusó de antisemitismo y le tildó de 'intelectual incendiario'. Alexandre Mathäs (2020) subraya con justicia que los rituales de memoria también pueden ser usados como un medio para superar, olvidar o blanquear el pasado. Ante el cúmulo de actos presentes de memoria sobre el Holocausto, recuerda que el problema ya no es 'no olvidar nunca' sino más bien 'cómo recordar'.

Como vemos, Walser era tan 'sincero' como Karp. Es a la suplantación de la conciencia, i.e. a la jerga del perdón, la culpa y la expiación pública impuesta políticamente, ya sea por los medios, los intelectuales, los partidos o por el estado, a lo que apuntaba su tesis. Karp va más allá: Walser también forma parte del juego de la integración social, en el doble movimiento de obediencia y rebeldía que promueve, más que no impide, la normalización alemana como una

cultura integrada y responsable. La jerga crítica no por eso deja de serlo, puesto que no provoca recusación ni extrañamiento, sino que ella misma reproduce la integración: «La jerga y nuestra pregunta sobre el rol de la agresión en la cultura son el resultado del juego interactivo [*Zusammenspiels*] entre Martin Walser y su público cultural más amplio» (Karp, 2002, p. 10). Según Karp, este doble juego es lo que el análisis de la Escuela de Frankfurt en general, y el de Adorno en particular, no podía explicar. Estaba aún demasiado próximo al discurso de la ética de la virtud del bien y del mal en su versión marxista. En especial, la dimensión agresiva de la jerga quedaba fuera de su significado. Ahora no se trataba de los valores, ni de ninguna liberación por los valores, sino de su funcionamiento en una sociedad compleja y contradictoria. Había, pues, que variar el punto de mira para ver la totalidad del cuadro y, dentro de él, el punto de fuga, para volver al núcleo de conexión de los sentimientos personales con las variables sociales que los producen y reconstruir la distancia entre el presente y el pasado, y entre el individuo y la sociedad. Lo que hay que investigar es «cómo la jerga puede jugar un papel en la integración de los impulsos agresivos en una cultura coherente. La reformulación específicamente cultural de la agresión debe ser también un componente de este tratamiento» (Karp, 2002, p. 3).

Para llevar a cabo este análisis introduce el concepto de *Trieb* (*driver*, conductor, guía, impulso) modificando levemente la forma en que Parsons lo utilizaba para realizar el paso de las variables sociales (culturales) a la formación de la personalidad del individuo. Según Karp, el modelo funcionalista no puede captar los matices culturales del *Lebenswelt* —intenta captarlos 'científicamente'—, pero constituye un útil punto de partida para entender el rol de la agresión en la cultura y el desarrollo humanos. 'Personalidad, cultura y sistema' es la tríada adecuada para no simplificar de entrada el problema conceptual de construir la estructura comprensiva (*Gestell*) de la agresividad y la agresión humanas. *Trieb*, en su faceta no-racional (*Triebhaftigkeit*), refiere también al 'impulso, compulsión e instinto' freudianos que constituye una de las fuentes del análisis de Parsons.

Según Karp, la sociología 'positiva' de Parsons (y de Marx, Weber y Mead antes que él) traduce la agresión en formas de divergencia o desviación (*abweichende*) respecto a las normas que constituyen la cultura internalizada en el proceso de integración social. Las conductas transgresoras lo son respecto a estos valores o normas. Pueden tener un efecto positivo (como en la lucha por los derechos humanos) o negativo (como en la comisión de delitos violentos que llevan aparejada una sanción). La violación de las normas no tiene por qué tener siempre un valor negativo. La dimensión funcional debe complementarse con la histórica. Para ello recurre a la obra de Helmuth Plessner *Die Verspätete Nation* (*La nación tardía*; [1959] 1988). Plessner sostiene que la forma de estado-nación se impuso tardíamente sobre un pueblo que no había alcanzado el sustrato de nación (como comunidad) propio de la Ilustración política francesa e italiana. Al llegar tardíamente, el 'espíritu' (*esprit, Geist*) de la nación no pudo albergarse en una burguesía alemana opuesta a todo lo que viniera de Francia, empezando por el Código Civil de Napoleón. La 'seducción totalitaria' se produjo, pues, sobre bases culturales: no basta la explicación de la derrota en la 1ª Guerra Mundial. El Holocausto no deriva de la derrota y el antisemitismo es más profundo. Es esto lo que recobra Karp: la estructura y las funciones de la agresividad descritas por Parsons a nivel societal (1947) adoptan la forma de agresión a nivel personal a través del substrato cultural alemán que las sustentan. Como señala Karp (2002) «El trabajo de Plessner apunta a la necesidad de comprometerse seriamente con las cualidades significativas de la acción cultural» (p. 9).

La sociología positiva tiende a traducir la agresión en formas de conductas desviantes (respecto a normas y valores). Pero el comportamiento transgresor que tiene su origen en la incapacidad de cumplir con normas o valores puede identificarse en función del contexto y relaciones a que da lugar. Así, al integrar lo intolerable en un comportamiento aceptable, Walser muestra cómo puede integrarse la agresividad en su relación con el comportamiento cultural concreto de la comunidad y la audiencia con que interacciona. En nuestra opinión, resuenan también las tesis de Viktor Klemperer (*LTI — Lingua tertii imperii*, 1947) sobre el lenguaje del III Reich y de Daniel Goldhagen (*Hitler's willing executioners: Ordinary Germans and the holocaust*, 1996) en este posicionamiento de Karp.

4. El teatro de improvisación

Formulémonos ahora la pregunta del *New York Times*: «¿Cómo puede ser que un estudiante de doctorado de filosofía alemana haya acabado liderando una empresa de análisis de datos que ha llegado a ser un importante contratista del Departamento de Defensa norteamericano y trabaja con los servicios de inteligencia mientras configura el futuro de la guerra autónoma?» (Dowd, 2024). Alexander Karp defendió su tesis doctoral en 2002. El siguiente año fundaba, junto con Peter Thiel, Stephen Cohen, Nathan Gettings y Joe Londale, Palantir. In-Q-Tel, una organización de capital riesgo sin ánimo de lucro de la CIA, fue su primer inversor. La compañía estaba valorada en 83.290 millones de dólares a 22 de septiembre de 2024. Cuatro meses más tarde, posiblemente debido a los últimos acontecimientos políticos y al desarrollo del panorama de relaciones internacionales, ha duplicado su valor. A 27 de enero de 2025 el valor de la empresa es de 179.840 millones. Tiene tres grandes plataformas: Palantir Gotham (inteligencia militar, lucha contra el terrorismo y crimen organizado), Foundry (integración de datos y análisis financieros) y Apollo (innovación y *Software-as-a-Service*, *SaaS*, para empresas). Es, sin duda, una de las grandes corporaciones globales de análisis de datos, con plataformas perfectamente diseñadas, orientadas al análisis. Por ejemplo, la arquitectura de Palantir Foundry tiene tres niveles (*layers*): semántico, kinético y dinámico, para realizar operaciones de circuito cerrado (*closed-loop operations*). Una ontología básica (*core-ontology*) integra los tres niveles. Lo que nos interesa aquí no es la estructura ni la evolución de la compañía, sino su estilo, la mentalidad de los emprendedores que la generaron y, en particular su dimensión intelectual, la cultura empresarial y política que promueven. Peter Thiel también tiene una licenciatura en filosofía, estudió con René

Girard y posee las primeras ediciones de los libros de Leo Strauss en inglés y alemán (Gellman, 2023). «Un reto fundamental, tanto en los negocios como en la vida, es integrar lo micro y lo macro de tal forma que todo tenga sentido», escribía Thiel (2012) en sus notas de clase de Stanford.

Ya hemos visto el bagaje de Alexander Karp. Hay diferencias entre ambos. Thiel forma parte desde hace muchos años de la derecha radical norteamericana de Silicon Valley, en contra de los derechos multiculturales, la discriminación inversa y la distribución de riqueza. A favor del *free speech* y del desarrollo sin límite de la tecnología, combina «una obsesión por el progreso tecnológico con políticas nacionalistas» (Chafkin, 2021, p. ivx). Profundamente norteamericano e implicado en su política, pese a sus orígenes alemanes. Recuerda el individualismo radical de los personajes de Ayn Rand en su descripción del sueño americano, un tecno-libertario capitalista a la contra que sólo pone el sistema en cuestión cuando cree que atenta contra la libertad. Karp, aparentemente, está más escorado hacia posiciones liberales, en el sentido tradicional de la izquierda en Estados Unidos, aunque coinciden en lo básico y, especialmente, en el patriotismo que les une a las creencias del complejo militar industrial que les proporciona contratos millonarios.

Más o menos en la misma época que Thiel y Karp fundaban Palantir, los estudiantes del Army War College de Carlisle, Pennsylvania, empezaron a hablar del mundo a que debían enfrentarse como VUCA (*Volatile, Uncertain, Complex, Ambiguous*), un mundo donde las estrategias tendrían que adaptarse a un cambio acelerado que impedía descripciones fijas y estables de la situación (Johansen, 2009). Así, como en un videojuego, la movilidad, la flexibilidad y proactividad tenían que ser articuladas y concretadas en posiciones agresivas, aunque fueran de vigilancia. Es la responsabilidad del liderazgo. Sin conexión con esta caracterización, Karp la había intuido en su tesis, y Palantir la adoptó desde el principio. Parafraseemos al NYT: «La posición de Karp es que corremos hacia este nuevo mundo nos guste o no. ¿Queremos ser dominantes en él, o queremos ser dominados por China?» (Dowd, 2024). Es un nuevo realismo político que no discute la violencia o la agresividad humanas, sino que sencillamente las asume sin crítica en nombre de un discurso práctico y pragmático de acción, natural, donde el reto es adquirir y situarse en posiciones de ventaja —económicas, tecnológicas, militares— sobre los adversarios. El planteamiento es de supervivencia. No adoptar esta posición implica aumentar el riesgo de desaparición, y esto se aplica tanto a las empresas en el mercado como a los estados nacionales en la geopolítica internacional. Recuerda las posiciones de Henry Kissinger en política y de Samuel Huntington en ciencia política. Hay que defender los valores de la cultura occidental.

El discurso es consciente y elaborado, y oponerse a él puede ser calificado, desde esta perspectiva, de hipocresía e incluso de inmoralidad, puesto que significa negar la realidad evidente del riesgo y dejar de actuar racionalmente en nombre de ideologías desconectadas del análisis de datos. Conviene, pues, no solamente reconocer, sino reforzar las conductas agresivas como las más efectivas para mitigar el riesgo. Pero cuidado: esto no debe hacerse de cualquier manera, sino dentro de una estricta moralidad que vale para marcar las interacciones a nivel individual y corporativo. Hay una red de *drivers* que plantean la cultura interna de la empresa, y ésta se presenta como un adalid de la disrupción, donde los títulos universitarios no cuentan, donde la valía se demuestra y reconoce solamente a partir de la competencia, y donde la innovación, y en concreto el dominio de la improvisación frente a las tareas rutinarias es lo que más se valora. En definitiva, la visión más allá de lo ahora existente. Palantir toma su nombre de 'palantiri', las esferas que otorgan una visión privilegiada al observador en la saga imaginada por John Ronald Reuel Tolkien, *The Lord of the Rings*. Este mundo simbólico y literario se refleja en el orden de la corporación, lleno de referencias a él. ¿Es necesario recordar que Tolkien reelaboró el discurso de la guerra del bien contra el mal en su mundo imaginario de hobbits, elfos y trolls?

David Johnson (2016) ha resumido vigorosamente los puntos esenciales de esta cultura corporativa, i.e. los mensajes que reciben los empleados, de la siguiente manera:

> (i) la educación que has recibido ha arruinado tu genio; (ii) en cada interacción, es crucial que conozcas cuál es tu estatus en relación a los demás [i.e. conocer tu lugar]; (iii) la clave para desencadenar la creatividad es decir que sí y 'sobreaceptar' las propuestas [i.e. explorarlas todas]; (iv) llegas a ser verdaderamente creativo cuando tu personalidad misma ha sufrido la disrupción, tienes un sentimiento de falta de decisión, y estás tan absorbido en tu trabajo que te sientes poseído por él. En definitiva, debes dejar atrás tu antiguo yo, asumir un nuevo papel en relación con tu estatus, y reconfigurar la personalidad según los dictados de tu trabajo (p. 25).

Johnson (2016) también señala la importancia que adquiere la expresividad en esta liberación personal autocontrolada: la intuición y flexibilidad de la interpretación requiere una atenta práctica de aprendizaje con los demás sometida a una disciplina férrea. Ésta es la que va a permitir ejercer el dominio de sí y, más allá, el control de las interacciones en las que se participa y que conducen a la innovación mediante la catálisis de esta guía corporal y mental. Este teatro de la improvisación teorizado por el antiguo miembro del *Royal Court Theatre,* Keith Johnstone, es de lectura casi obligada en la compañía. Hay que desaprender primero, puesto que la monotonía (*dullness*) «no es consecuencia de la edad, sino de la educación» (Johnstone [1976], 1981, p. 14). La espontaneidad nunca es espontánea; y no hay que creer en nada porque sea meramente conveniente. Así se entiende la formación de brigadas adaptadas al cliente que trabajan con él, en contra de la tendencia dominante de dejar de lado al usuario; y la elaboración de códigos éticos y jurídicos estrictos que se siguen más allá de lo normativamente necesario. En el sector financiero y bancario, hemos designado recientemente este fenómeno de sobreactuación como 'sobrecumplimiento' (*overcompliance*) (De Koker y Casanovas, 2024). Se trata de un instrumento de gobernanza jurídica interna que opera a

favor del banco frente a los servicios que, por ley, debe prestar a su cliente. No nos habíamos dado cuenta aún de lo importante que era en el interior de gigantes tecnológicos como Palantir, quien no solamente evita sanciones, hace frente con éxito a las acusaciones de malas prácticas y a las demandas, escoge a sus clientes y decide qué servicios les proporciona (SaaS), sino que logra integrar en una unidad las posibles disidencias internas, anticipa y neutraliza los movimientos de luchas (*turf battles*) y se presenta como un todo social coherente y compacto. Como una cultura, en definitiva, que incorpora modelos de regulación basados en ética y valores universales concretos a distintos niveles de abstracción que buscan ser coherentes consigo mismos y consistentes en su implementación. Es granular y busca la concreción y la simplicidad: de la posición geopolítica global a las prácticas que deben observar los empleados. Esta cultura no excluye de ningún modo la agresividad (como característica social) y la agresión (cuando ésta es necesaria). Se consideran componentes de la sociedad y del mercado, con el contrapeso de la cultura creada para liberar la creatividad individual y colectiva, defendiéndola tanto de injerencias externas como de actuaciones políticas sobrevenidas.

Así, por una parte, Karp aboga por la defensa extrema de los valores occidentales (incluyendo los europeos) liderados por Estados Unidos, el único país capaz de hacerlo porque ahí es donde se produce la mejor tecnología para ser usada en el campo de batalla. Por otra, también defiende acaloradamente el derecho a vender estos productos en el mercado, especialmente si los compradores son países comprometidos en la defensa de estos valores, como Israel (contra el terrorismo islámico) y Ucrania (contra la invasión rusa) (Karp, 2021, 2024a, 2024b). Los obstáculos que hay que vencer no son propiamente tecnológicos. Se encuentran en la desconfianza del propio Pentágono, la presencia de competidores con productos que no dan respuesta a las necesidades y en mentalidades ancladas aún en el pasado inmediato; los académicos que no entienden la urgencia de hacer frente a enemigos armados en el ámbito de la geopolítica internacional no son actores: realmente no cuentan. Como en el caso de TITAN (*Tactical Intelligence Targeting Access Node*), hay que anticiparse para dotar de instrumentos eficaces de integración de datos y construcción de sistemas autónomos mediante IA al ejército en los frentes de batalla. Según uno de sus últimos contratos (US Army, 2024, 6 de marzo): «TITAN es la estación terrestre de inteligencia, vigilancia y reconocimiento de próxima generación del Ejército, que aprovecha la inteligencia artificial y el aprendizaje automático para procesar rápidamente los datos de los sensores recibidos desde las capas espacial, de gran altitud, aérea y terrestre».

Al final, a fin de cuentas, el teatro de improvisación se convierte en teatro de operaciones. La última carta dirigida a los accionistas es, como siempre, muy clara y describe perfectamente los objetivos de la empresa en un mundo de confrontaciones bélicas, explicitando esta vez la estrategia de encuadrar las posibilidades de la inteligencia artificial generativa en el marco lógico y conceptual ofrecido por Palantir. Es como otro *Sturm un Drang* contemporáneo —Karp no puede dejar de citar a Johann Wolfgang-Goethe, incluso cuando informa sobre los beneficios—, la antesala del porvenir. Al mismo tiempo que la promesa de continuar construyendo software «*for the world as it is, not as we wish it might be*», insiste en la posibilidad de dominarlo y conducirlo:

> Los grandes modelos lingüísticos que han cambiado (*transfixed*) el mundo sólo podrán transformar el trabajo de una corporación multinacional o las operaciones de una agencia de defensa si su poder se desata en el contexto de un sistema de software empresarial con una precisa visión del mundo —sus objetos idiosincrásicos, su lógica y su física. Los modelos con billones de parámetros pueden imitar perfectamente a Goethe, pero sin nada más añaden poco valor a la empresa. Han nacido en este mundo sin ningún sentido de sus contornos o lógica, o incluso una concepción de la verdad o de los hechos básicos, y mucho menos del conocimiento colectivo y la percepción de las operaciones de una organización con medio millón de empleados. Los modelos de lenguaje que han capturado nuestra atención colectiva tienen la capacidad de aproximarse a la verdad. Son motores probabilísticos increíblemente capaces que, sin embargo, tienen poco sentido innato del norte o del sur y se esfuerzan por resolver problemas analíticos complejos por sí mismos. Estos modelos son necesarios, pero no suficientes. Se requiere algo más. Son animales salvajes, cuyo poder y capacidades deben ser domesticados y aprovechados. Y ahora estamos viendo lo que es posible cuando lo son (Karp, 2024c).

En los próximos epígrafes sugeriremos que este modelo de regulación no se asimila ni al fascismo ni al totalitarismo —donde el estado ocupa un papel central—, sino que se trata de un modelo social autoritario de regimentación. Hay que dejar libertad total, al mismo tiempo que se ejerce una dirección implícita y granular adecuada. Se trata de la manifestación política de la cultura de la regimentación cultural y lingüística que opera a través de los algoritmos de los grandes modelos de lenguaje (Schneider, 2022). Esto no es riguroso: es una metáfora, naturalmente, pero en nuestra opinión hay que otorgarle el valor descriptivo que posee. Forzando la metáfora, la imagen sería la de concebir la sociedad como una esfera de Bloch, un *qubit*, en la que la corporación debe interferir. En computación cuántica, según la Wikipedia, la interferencia «ayuda a controlar los estados cuánticos y amplificar los tipos de señales que están orientados hacia la respuesta correcta, y luego cancelar los tipos de señales que conducen a la respuesta incorrecta».

5. La implementación de principios éticos en los modelos de regulación

Vayamos ahora a la ética y los modelos de regulación. Para centrarnos en USA, puede resultar ilustradora la comparación entre los principios éticos de Palantir con los clásicos de Asilomar. Los primeros se presentan del siguiente modo:

> Defendemos una ética de la tecnología que se aplique a todos los contextos de uso. Cada uno de estos contextos implica su propio conjunto de demandas específicas de cada dominio, expectativas funcionales y obligaciones éticas. Este marco nos obliga a poner a la IA en el lugar que le corresponde: como una herramienta entre otras herramientas de diversa sofisticación e inexorablemente inserta en un mundo de acciones y consecuencias tangibles (Palantir, 2024a).

Los principios básicos aplicables son los siguientes:

- Centrarse en el sistema totalmente integrado, no solo en las herramientas que lo componen.
- Reconocer los límites de la tecnología.
- No resolver problemas que no se deben resolver.
- Cumplir con las mejores prácticas metodológicas para una sólida ciencia de datos.
- Mantener la IA responsable, capaz de rendir cuentas y orientada a los humanos.
- Promover el compromiso de las múltiples partes interesadas.
- Garantizar la conciencia técnica, de gobernanza y cultural en las aplicaciones de datos y tecnología.

Las veintitrés recomendaciones de Asilomar nacen en 2017 de una Conferencia con más de 1200 científicos implicados organizada por el Future of Life Institute (fundado en el MIT en 2015). Establecen principios para la investigación (*research issues*) y la ética (*ethics and values*); para el desarrollo futuro de la IA (*longer-term issues*). Los principios éticos consensuados son trece, y pueden sintetizarse en los siguientes conceptos: seguridad, transparencia (en fallos y uso por parte de la judicatura), responsabilidad, valores humanos y alineación con las máquinas, privacidad, libertad, beneficio y prosperidad compartida, control, exclusión de la subversión social. Y, especialmente, exclusión de una carrera de armamentos en IA con armas autónomas letales.

Aparte de las diferencias inmediatamente visibles, como en el caso de ésta última prohibición, los principios éticos de Palantir presentan una concreción y focalización en la construcción de sistemas de IA mucho mayor. Son: endógenos (aplicables a la corporación industrial en cuanto tal); técnicos (solamente se ocupan de acciones de diseño en IA cuyos algoritmos puedan reproducirse); explícitos (toda acción debe poder ser descrita en todos sus pasos con especificación de algoritmos); dinámicos (descritos específicamente para cada estadio de la cadena de producción); articulados (tienen presente específicamente su integración en todo el sistema); intencionales (todas las acciones se dirigen al objetivo final de la producción de un producto de calidad); contextuales (hay que tener en cuenta el ámbito tecnológico y sobre todo el contexto al que se aplican los sistemas). Hay tres rasgos que son destacables en ellos: se incorporan y deben mejorar la cadena de valor de la IA en el proceso de producción, distribución e implementación del producto, se distinguen (y apartan conscientemente) de tota norma moral y política; pretenden ser componentes de la emergencia de una cultura corporativa, i.e. una cultura colectiva, que no puede imponerse si no es a partir de la conciencia personal de cada uno de los empleados de la corporación. Por ello se insiste en la personalidad del empleado. Éste no tiene solamente un empleo: refleja (debe reflejar) una cultura, un *modus operandi* que alimenta y preserva la creatividad del empleado en beneficio de la propia empresa.

Los principios de Asilomar, en cambio, son exógenos, mucho más generales, y lo que reflejan es el *Lebenswelt,* (el mundo-de-la-vida)*,* el sentido común de los científicos que los han propuesto. Justo lo que quería evitar Karp a toda costa con los suyos propios, porque literalmente no sabemos, no podemos predecir el mundo que tendremos delante y al que nos enfrentaremos en los próximos años. Asumiendo el papel de abogado del diablo, desde el punto de vista de Karp, si lo hemos entendido bien, constituyen una jerga, una jerga ética tranquilizadora, sin que aquellos que los promueven tengan nada que perder en el juego y vean de qué forma puedan concretarse en un mundo conflictivo, ya en guerra entre los distintos estados-nación que han seguido el proceso de globalización del fin del milenio. En un mundo global, viene a decir, una ética global carece de sentido: la ética es siempre situada, concreta, y contribuye a producir un determinado estado de cosas que debemos anticipar y, en la medida de lo posible, dirigir. La intensidad es una de sus características. Antiguos empleados, como Nabel Qureshi, (2024), recuerdan su importancia en entrevistas de trabajo atípicas —«estuvimos una hora hablando sobre Wittgenstein». Sin embargo, estos argumentos pueden volverse contra quien los emplea, puesto que el *Lebenswelt* de los que integran la jerarquía, i.e. los que toman las decisiones sobre la política de la corporación, también cuenta. Y aquí no se ve cómo puede funcionar la ética concreta sin las decisiones corporativas basadas en un patriotismo que no sobreviene, sino que está en el fondo de las relaciones militares y geopolíticas de la cúpula. A contrario, Palantir entra en la ética existencial ajena puesto que define los valores de sus empleados y penetra en su integridad, no cuestiona internamente la actividad de los jueces, no cuestiona la legitimidad de los actores en las políticas internacionales de agresión de sus clientes, no comparte las políticas de distribución de riqueza y se convierte —como organismo colectivo— en un actor esencial en los conflictos bélicos. A nuestro entender, también promociona (y explícitamente) normas morales y políticas y se posiciona en la geopolítica global.

Los principios de Asilomar son ciertamente más generales y reactivos. Pero muestran con bastante claridad el rechazo de la comunidad científica a las prácticas de manipulación y sesgo de los algoritmos denunciadas, entre otras, por Cathy O'Neil (2016) y al uso incierto en el ámbito judicial de programas como COMPAS (Rubel et al., 2020). El escándalo de Cambridge Analytica vendría a corroborar todos los temores (Cadwalladr, 2019). Palantir mira hacia otro lado; y más bien se extraña ante la reacción. ¿Qué pensar de todo ello? ¿Constituye la IA una invitación al fascismo o al totalitarismo? ¿Es esto una posibilidad real?

6. Reflexiones finales: ¿Hay un fascismo latente?

En puridad, no podemos descartar que la adopción de las tecnologías de la IA no conlleve un riesgo de destrucción, aunque sea mínimo. Éste es el sentido de la advertencia y de la carta de 23 de mayo de 2023 de Geoffrey Hinton, Yoshua Bengio, Demis Hassabis, Sam Altman y otros expertos y CEOs. Pero si la pregunta es si las formas políticas totalitarias pueden ser fomentadas por el tipo de tecnología escogida, especialmente por la IA de última generación, la denominada IA generativa (LLM y MFM) diríamos, intuitivamente, que la pregunta no puede ser planteada de esta manera. Las tecnologías no son neutras, pero tampoco culpables de entrada. Depende de los objetivos, intenciones y planes que tengan aquellas personas que las diseñan, y especialmente las que las explotan. De la misma manera que Internet no crea nuevas emociones, pero puede intensificar las que ya existen de forma natural (alegría, tristeza, miedo, ira o amor) mediante la forma de articular las transacciones, la IA puede intensificar los propósitos y materializaciones de cualquier forma política, democrática, oligárquica o dictatorial (Russell, 2013; López de Mántaras, 2024).

Resulta posible efectuar un análisis de riesgos de la IA utilizando los medios econométricos clásicos (Noriega y Casanovas, 2024). El análisis de situaciones de crisis, catástrofes y riesgos latentes permite parametrizarlo, considerando también las variables temporales. Sin embargo, abordar el problema para mitigar los riesgos no es suficiente para comprender de qué forma se generan y desarrollan. Para ello es necesario entender cómo se producen las inversiones, la gobernanza corporativa y las decisiones éticas y políticas de los grandes actores, y de qué forma la mentalidad subyacente a sus patrones de conducta afecta al mercado y a las políticas internacionales (incluyendo todas las fases de los conflictos bélicos). Si algo hemos aprendido del análisis de Palantir, es que las grandes dicotomías no pueden ayudarnos a explicar bien lo que sucede en el nuevo entorno digital. El riesgo de desaparición tiene precedentes en los fascismos y totalitarismos del siglo XX, y especialmente en el uso bélico de la energía nuclear que se produjo a consecuencia de los avances tecnológicos y la escalada del conflicto. Esta situación perduró en la guerra fría que siguió. Pero los factores de la transformación que se está produciendo ahora con relación al conocimiento no es equivalente a los que dieron lugar a los totalitarismos del siglo XX. Por una parte, el cambio es más profundo, porque no solamente afecta a las condiciones productivas, sino a las propiedades de cognición, intelección y procesamiento de información que hace sólo cuatro décadas se producían únicamente mediante el uso de la inteligencia natural humana. Por otra, la aparición de Internet, la web de datos enlazados y la IA no ha ocurrido en el vacío, sino en las condiciones generadas por la globalización.

Si las afirmaciones precedentes resultan aceptables, la pregunta inicial sobre la relación entre fascismo e IA debe modificarse y ampliarse. No existe una relación de causalidad directa que pueda probarse inductivamente en una prospectiva que resulte fiable. El fascismo probablemente no pueda volver de la forma en que lo conocimos en las sociedades occidentales (en la organización del estado, la relación con el mercado, y la configuración de una sociedad civil sometida a los valores y principios de una pequeña élite política ultranacionalista). Pero esto no obsta para que algunos de sus componentes, empezando por su propensión a la agresión, la violencia y la imposición totalitaria de formas político-jurídicas sobre la población, no puedan sobrevivir mediante otros medios y adoptando formas diversas. Las grandes corporaciones tecnológicas no son fascistas, y tampoco lo son los gobiernos que las sostienen. Sin embargo, la imposición de patrones de conducta basados en la inteligencia de una minoría que configura también una cultura dominante oligárquica y exclusiva resulta preocupante. Por una parte, esta cultura tiene la capacidad de impregnar toda la red, las instituciones y los comportamientos de los usuarios (como consumidores en el mercado y como actores sociales y políticos). Es la cultura carismática del líder, al que se sigue en las redes sociales, periódicos y medios de información, y, en otro nivel, con el que negocian los gobiernos. Por otra parte, sus procesos de decisión se producen siguiendo las reglas del incremento de valor y la competición en todos los terrenos del capitalismo. El neoliberalismo se diferencia del liberalismo clásico, entre otras cosas, por su asunción del comportamiento agresivo como una de las propiedades naturales del ser humano que debe aceptarse y alimentarse. Esto no es en sí una legitimación de la violencia. Pero al final, la conducta que promueve es compatible con ella, pretende que ésta puede ser controlable, y por tanto la legitima.

Hay que tomar en cuenta este hecho, porque la capacidad de investigación de las corporaciones anticipa el futuro inmediato de una forma más precisa que hace diez años. Son las corporaciones las que tienen intuiciones certeras y factibles sobre la innovación en tecnología y la evolución del mercado. Son también ellas las que lideran y, en determinados casos, personalizan la defensa de las minorías vulnerables —Peter Thiel ha hecho pública con naturalidad que ha formado una familia homosexual, y Alexander Karp su tradición judía. Los estados vienen detrás. Por eso funciona la estrategia de sobrecumplimiento ético y jurídico que hemos descrito en este artículo. Los modelos regulatorios —especialmente los modelos que utilizan lenguajes formales de cumplimiento— permiten ahorrar costes, tiempo y esfuerzo. Y también pueden reforzar las políticas corporativas que evitan conflictos y sanciones de las agencias gubernamentales (privacidad y protección de datos, sobre todo) y demandas que lleguen a los tribunales por parte de los usuarios.

Nuestra sugerencia es que en lugar de pensar directamente en términos muy amplios aplicables al tipo de sociedad y estado —como fascismo y totalitarismo— nos concentremos en los procedimientos y procesos de gobernanza para detectar los elementos autoritarios que puedan ponerse en práctica. Este es un trabajo aún por hacer, pero vamos a recordar aquí una diferencia que resulta esencial para entender la construcción normativa de los denominados sistemas sociotécnicos y los sistemas multiagente. Casi desde el principio de la formalización de normas en lógica deóntica (Sergot y Jones, 1992), los científicos computacionales cayeron en la

cuenta de la diferencia entre 'regular' y 'regimentar'. Son verbos y acciones de una gran tradición en política y derecho. Regular implica considerar a los agentes como autónomos: pueden decidir si cumplir o violar las normas, bajo su responsabilidad. Regimentar, en cambio, implica reducir su autonomía e imbuir el cumplimiento dentro de su ámbito de acción, reduciendo el de decisión. Literalmente, no pueden violar las normas. Amir Chopra y Munindar P. Singh (2018) lo describen claramente en su discusión ética de la responsabilidad de agentes y sistemas inteligentes:

> La regulación es la idea de que las normas relevantes para el entorno se especifican y su satisfacción queda en manos de los agentes. La regulación, por lo tanto, promueve la autonomía pero al mismo tiempo produce estándares explícitos de corrección. La regimentación es la idea de que las interacciones entre los agentes pueden implementarse en máquinas; sus estándares de corrección están implícitos (p. 49).

Un sistema que opere en contextos reales debe buscar el equilibrio o proporción entre una arquitectura social (regulatoria) y una arquitectura técnica (regimentativa). En la tradición medieval, 'regimiento' no solamente significaba autoridad, sino que aludía también al cúmulo de conocimientos necesarios para la decisión. Así, el libro de Pedro de Medina (1563) *Regimiento de navegación* «contiene las cosas que los pilotos han de saber para bien navegar: y los remedios y avisos que han de tener para los peligros que navegando les pueden suceder» (p. 1).

En la ingeniería social y política contemporánea de las grandes corporaciones, esta prudencia y equilibrio se somete a la idea de que las interacciones entre agentes pueden implementarse también en los empleados, consumidores y ciudadanos. Se trata pues de una regimentación implícita, no de una regulación, en la que el conocimiento no fundamenta o basa la decisión, sino que la determina. Este exceso es lo que llama la atención en la ética de las grandes corporaciones como Palantir. Exceso de confianza, exceso de cumplimiento, exceso de creatividad, exceso de moralidad, exceso de agresividad. Y la creencia de que el atrevimiento nunca es excesivo y el diálogo social no es necesario. Todo ello lleva a la reducción del espacio y del derecho público. Paradójicamente, es el reverso de la democracia deliberativa de Habermas y de la contención de la violencia de Girard, dos de los mentores principales de Karp y de Thiel. Se trata quizás también de la arrogancia de los *nerds* respecto a las mismas convenciones sociales que los han encumbrado pero que ellos no respetan, las nuevas 'creencias de lujo' de las clases privilegiadas (Henderson, 2024).

Nuestra idea es que la respuesta a estas prácticas puede ser asimismo tecnológica. Es mejor observarlas atentamente, aprender en la medida de lo posible, y construir modelos regulatorios que puedan ejecutarse para suscitar la emergencia de ecosistemas sociales que incorporen valores éticos, el contenido de los derechos humanos y, de paso, constituyan un límite para este nuevo poder corporativo. Creemos que es posible construirlos de manera que proporcionen seguridad y efectividad jurídicas al mismo tiempo que sean aceptables y asumibles por parte de la ciudadanía, los consumidores y las empresas. Pero de qué forma puede operar la democracia de datos o 'enlazada' (Poblet et al., 2019), cómo pueden 'imbuirse' principios éticos en los sistemas inteligentes mediante la ingeniería de valores (Noriega et al., 2023), y cómo crear las condiciones para constituir estos ecosistemas jurídicos mediante la formulación de modelos de regulación ejecutables (Casanovas et al., 2025), son problemas que quedan fuera del objeto de este artículo. En él, solamente hemos pretendido plantear la dificultad de relacionar la inteligencia artificial con las formas políticas, y especialmente con la emergencia de modelos autoritarios. Recordando el dicho de Henry-Louis Mencken (1917/2021, p. 158): «Explanations exist; they have existed for all time; there is always a well-known solution to every human problem — neat, plausible, and wrong» [Las respuestas están ahí; siempre lo han estado. A cada problema humano le corresponde una solución aparentemente impecable: ordenada, convincente... y sin embargo, equivocada].

7. Disponibilidad de datos

Deseo declarar mi adhesión al movimiento de ciencia abierta. Este artículo es de naturaleza analítica, por lo que no se han utilizado datos empíricos que pudieran compartirse en repositorios comunes. La gran mayoría de textos clásicos citados se encuentran ya disponibles en abierto.

8. Declaración de uso de LLM

Este artículo no ha utilizado ningún texto generado por un LLM (ChatGPT u otro) para su redacción.

9. Agradecimientos

Este artículo tiene su origen en la conferencia plenaria sobre fascismo e IA pronunciada en catalán por Pompeu Casanovas en el Lycée Polyvalent Charles Renouvier, Universitat Catalana d'Estiu (UCE), Prades (Francia) el 19 de agosto de 2024. Participaron también Marta Poblet y Jordi Casassas, quien operó como moderador. El presente artículo corrige algunas imprecisiones y fallos de memoria. El acto fue registrado y se halla accesible en https://www.youtube.com/watch?v=CAxR6mhCwdc. El tema fue parcialmente desarrollado también en el curso impartido por Joan González Guardiola (UIB), Marta Poblet (Data Tank; RMIT) y Pompeu Casanovas titulado *Pensament filosòfic i intel·ligència artificial* entre el 17 y el 20 de agosto en la misma UCE. Agradezco a Enric Plaza y Marco Schorlemmer sus comentarios a una primera versión de este artículo. Me ha resultado sumamente útil la lectura y la crítica de Pablo Noriega en el contexto de nuestro trabajo en común. La responsabilidad por el resultado final es solamente mía.

10. Referencias

Adorno, Theodor W. (1964). *Jargon der eigentlichkeit, zur deutschen ideologie*. Suhrkamp. Reed. 1969. Versión en inglés de Knut Tarnowski y Frederic Will, *The jargon of authenticity*. Northern Western University Press, 1973.

Adorno, Theodor W. (1951). *Minima moralia. Reflexionen aus dem beschädigten leben.* Suhrkamp

Verlag. Versión castellana: *Minima Moralia. Reflexiones desde la vida dañada*. Trad. Joaquín Chamorro Mielke. Akal, 2006.

Adorno, Theodor W. (1967). *Aspekte des neuen rechtsradikalismus. Ein vortrag mit einem nachwort von volker weiß*, suhrkamp, 2019. Traducción catalana, Gustau Muñoz, *Aspectes de la nova extrema* dreta. Catarroja: Ed. Afers, 2022.

Akata, Zeynep, Dan Balliet, Maarten De Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Antske Fokkens et al. (2020). A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer 53*(8), 18-28. https://ieeexplore.ieee.org/document/915387

Aquarone, Alberto (1965). *Organizzazione dello stato totalitario.* Einaudi.

Arendt, Hanna (1951). *The origins of totalitarianism.* Traducción de Guillermo Solana, *Los orígenes del totalitarismo*, Taurus, 1974.

Brose, Eric Dorn (1987). Generic fascism revisited: Attitudes toward technology in Germany and Italy, 1919-1945. *German Studies Review*, *10*(2), 273-297. https://www.jstor.org/stable/1431103

Cadwalladr, Carole (2019, 16 de abril). *Facebook's role in Brexit — and the threat to democracy.* TED Talk. You Tube. https://www.ted.com/talks/carole_cadwalladr_facebook_s_role_in_brexit_and_the_threat_to_democracy

Casanovas, Pompeu, Hashmi, Mustafa, De Koker, Louis y Lam, Ho-Pun (2025, en prensa). Compliance, regtech, and smart legal ecosystems: A methodology for legal governance validation. En Woodrow Barfield and Ugo Pagallo (eds.), *Research Handbook on the Law of Artificial Intelligence*, vol. II, Edward Elgar Publ.

Chafkin, Max (2021). *The contrarian: Peter Thiel and Silicon Valley's pursuit of power*. Bloomsbury Publishing.

Chopra, Amit K. y Munindar P. Singh (2018). "Sociotechnical systems and ethics in the large." En *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 48-53.

De Koker, Louis y Casanovas, Pompeu (2024). 'De-Risking', de-banking and denials of bank services: An over-compliance dilemma? En D. Goldbarsht y L. de Koker (Eds.) *Financial crime, law and governance. Ius gentium: Comparative perspectives on law and justice* (pp. 45-70). Springer. https://doi.org/10.1007/978-3-031-59547-9_3

Dierkes, Meinolf, Knie, Andreas y Wagner, Peter (1990). Engineers, intellectuals and the state: Assessing technology in the Weimar Republic. *Industrial Crisis Quarterly 4*(3),155-174. https://doi.org/10.1177/108602669000400302

Dowd, Maureen (2024, 17 de agosto). Alex Karp has money and power. So, what does he want? *New York Times,* https://www.nytimes.com/2024/08/17/style/alex-karp-palantir.html

Gay, Peter (2001). *Weimar culture: The outsider as insider* [1968]. W.W. Norton.

Gellman, Barton (2023, 9 de noviembre). Peter Thiel is taking a break from democracy, *The Athlantic*, https://www.theatlantic.com/politics/archive/2023/11/peter-thiel-2024-election-politics-investing-life-views/675946/

Gerhardt, Uta (1999). A world from brave to new: Talcott Parsons and the war effort at Harvard University. *Journal of the History of the Behavioral Sciences*, *35*(3), 257-289. https://doi.org/10.1002/(SICI)1520-6696(199922)35:3<257::AID-JHBS5>3.0.CO;2-A

Gerhardt, Uta (2002). *Talcott Parsons: An intellectual biography*. Cambridge University Press.

Gerhardt, Uta (2010). Talcott Parsons's Weberian analysis of national socialism. En Christopher Hart (Ed.) *Essays in honour of Talcott Parsons* (pp. 233-266). Midrash Editions.

Gerhardt, Uta (Ed.) (2017). *Talcott Parsons on national socialism* [1993]. Routledge/Taylor Francis.

Gibson, James J. (1966). *The senses considered as perceptual systems*. Allen and Unwin.

Gispen, Kees (1992). National socialism and the technological culture of the Weimar Republic. *Central European History 25*(4), 387-406. https://www.jstor.org/stable/4546293

Heidegger, Martin (1997). *Die frage nach der technik* [1954]. Versión castellana: *La pregunta por la técnica*. Traducción de Jorge Aceved. En Martin Heidegger, *Filosofía, ciencia y técnica* (pp. 111-149). Ediciones Universitarias.

Henderson, Rob (2024). *Troubled: A memoir of foster care, family, and social class*. Gallery Books.

Herf, Jeffrey (1986). *Reactionary modernism: Technology, culture, and politics in Weimar and the Third Reich*. Cambridge University Press.

Hinton, Geoffrey, Bengio, Yoshua, Hassabis, Demis, Altman, Sam et al. (2023, 23 de mayo). Open Letter. Statement on AI risk, *Center for AI safety*. https://www.safe.ai/work/statement-on-ai-risk#open-letter

Husserl, Edmund (1936). *Die krisis der europäischen wissenschaften und die transzendentale phänomenologie: Eine einleitung in die phänomenologische philosophie* [1936]. Versión castellana: *La crisis de las ciencias europeas y la fenomenología trascendental.* Traducción de Julia V. Iribarne. Prometeo Libros. https://archive.org/details/husserl-e.-la-crisis-de-las-ciencias-europeas-y-la-fenomenologia-trascendental-epl-fs-1936-2022

Johansen, Bob (2009). *Leaders make the future. The new leadership skills for an uncertain world.* Berrett-Koehler Publishers.

Johnson, David V. (2016). How Palantir has made corporate orthodoxy out of experimental theatre, *The Baffler*, *32*, 20-25. https://www.jstor.org/stable/43959272

Jones, Andrew JI, y Sergot, Marek (1993). On the characterisation of law and computer systems: The normative systems perspective. En J.J. Mayer, R. Wieringa (eds.). *Deontic logic in computer science: Normative system specification* (pp. 275-307). John Wiley & Sons, Inc.

Johnstone, Keith (1981). *Impro: Improvisation and the theatre* [1976]. Methuen Drama Ld.

Karp, Alexander C. (2002). *Aggression in der lebenswelt: Die erweiterung des parsonsschen konzepts der aggression durch die beschreibung des zusammenhangs von jargon, aggression und kultur.* Inauguraldissertation Zur Erlangung des Grades eines Doktors der Philosophie im Fachbereich Gesellschaftswissenschaften der Johann Wolfgang-Goethe Universität zu

Frankfurt am Main. https://publikationen.ub.uni-frankfurt.de/opus4/frontdoor/deliver/index/docId/5389/file/00000316.pdf

Karp, Alexander C. (2021, 18 de marzo). *Palantir CEO Alex Karp in conversation with CNBC's Wilfred Frost.* You Tube. https://www.youtube.com/watch?v=kBlGMHiPf1U

Karp, Alexander C. (2024a, 7 de marzo). *Palantir CEO Karp on TITAN, AI warfare technology.* You Tube. https://www.youtube.com/watch?v=P2eRjbUW53s

Karp, Alexander C. (2024b, 6 de septiembre). *How Palantir is leading the AI revolution and creating a safer world.* You Tube. https://www.youtube.com/watch?v=Xg-KFZHd7uM

Karp, Alexander C. (2024c, 5 de agosto). *Letter to Shareholders*. https://www.palantir.com/q2-2024-letter/en/

Kurzweil, Ray (2024). *The singularity is nearer: When we merge with AI*. Random House.

Linz, Juan José (2000). *Totalitarian and authoritarian regimes*. Lynne Rienner Publishers.

López de Mántaras, Ramón (2024). *100 cosas que hay que saber sobre inteligencia artificial.* Lectio Ediciones.

Mathäs, Alexander (2020). The Presence of the past: Martin Walser on memoirs and memorials, *German Studies Review 25*(1), 1-22. https://doi.org/10.2307/1433242

Mechler, Reinhard, Laurens M. Bouwer, Thomas Schinko, Swenja Surminski, y Joanne Linnerooth-Bayer (Eds.) (2019). *Loss and damage from climate change: Concepts, methods and policy options*. Springer Nature. https://link.springer.com/book/10.1007/978-3-319-72026-5

Medina, Pedro de (1563). Regimiento de navegación. Simon Carpintero. Accesible en la Universidad de Texas: https://collections.lib.utexas.edu/catalog/utblac:f3f58d71-29b8-4187-a119-5a2dc70bd047

Mencken, Henry-Louis (1917, 16 de noviembre). The Divine Afflatus. New York Evening Mail [reeditado en *Prejudices: Second Series*. Jonathan Cape, 2021, p. 158, y reproducido por el Proyecto Gutenberg, https://www.gutenberg.org/cache/epub/53467/pg53467-images.html#Page_155]

Montfort, Kristin de (2023, 30 de noviembre). *Alex Karp's Aggression in the Lebenswelt*, https://kristindemontfort.substack.com/p/alex-karps-aggression-in-the-lebenswelt Edición y traducción al inglés de la tesis doctoral de Alexander Karp (2002).

Neufeld, Michael J. (1990). Weimar culture and futuristic technology: The rocketry and spaceflight fad in Germany, 1923-1933. *Technology and Culture 31*(4), 725-752. 10.1353/tech.1990.a901648

Noriega, Pablo, Harko Verhagen, Julian Padget, and Mark d'Inverno (2023). En N. Fornara, J. Cheriyan,A. Mertzani, (Eds.) Coordination, organizations, institutions, norms, and ethics for governance of multi-agent systems XVI. COINE 2023. *Lecture Notes in Computer Science*, vol 14002 (pp. 77-94). Springer https://doi.org/10.1007/978-3-031-49133-7_5

Noriega, Pablo and Pompeu Casanovas (2024). From the Pascal wager to value engineering: A glance at AI risks and how to address them. ECAI - 27TH European Conference on Artificial Intelligence, Workshop on AI Value Engineering and AI Compliance Santiago de Compostela, 19 de octubre. https://jurisinformaticscenter.github.io/VECOMP2024/

O'Neil, Cathy (2016). *Weapons of math destruction. How big data increases inequality and threatens democracy*. Crown Books.

Palantir (2024). *The ontology-powered operating system for the modern enterprise* https://www.palantir.com/platforms/foundry/

(2023, 19 de septiembre). *Palantir technologies' approach to AI ethics*. https://www.palantir.com/pcl/palantir-ai-ethics/

Parsons, Talcott (1947). Certain primary sources and patterns of aggression in the social structure of the Western world. *Psychiatry 10*(2), 167-181. Reeditado en G. Lanyi and W. McWilliams (Eds.), *Essays in sociological theory* (1949) y The structure of group hostility, en *Crisis and continuity in world politic*s (2nd ed.), Random House, 1973.

Paxton, Robert O. (1998). The five stages of fascism. *The Journal of Modern History*, *70*(1), 1-23. https://www.jstor.org/stable/10.1086/235001.

Plessner, Helmuth (1988). *Die verspätete nation. Über die verführbarkeit bürgerlichen geistes* [1959]. Berlin: Suhrkamp. Hay traducción castellana de Kilian Lavernia: J. Fischer y J.L. Villacañas (Eds.) (2017), *La nación tardía: Sobre la seducción política del espíritu burgués (1935-1959)*. Biblioteca Nueva.

Poblet, Marta, Pompeu Casanovas y Víctor Rodríguez-Doncel (2019). *Linked democracy: Foundations, tools, and applications*. Springer Brief 750, Springer Nature. https://link.springer.com/book/10.1007/978-3-030-13363-4

Qureshi, Nabeel (2024, 7 de noviembre). Reflections on Palantir: A retrospective of an eight-year stint. https://medium.com/@nabeelqu/reflections-on-palantir-52433cf95439

Rubel, Alan, Castro, Clinton y Adam Pham (2020). Algorithms, agency, and respect for persons, *Social Theory and Practice*, *46*(3),547-572. https://www.jstor.org/stable/45378044

Russell, Stuart J. (2023). *Human compatible: AI and the problem of control*. Penguin.

Russell, Stuart J. y Peter Norvig (2016). *Artificial Intelligence: A modern approach* (1995), 2nd Edition Prentice Hall.

Simon, Herbert A. (1991). *Models of my life*. Basic Books, Sloan Foundation Series.

Schneider, Britta. (2022). Multilingualism and AI: The regimentation of language in the age of digital capitalism. *Signs and Society*, *10*(3) (2022), 362-387. https://doi.org/10.1086/721757

Simon, Herbert A. (1996). *The Sciences of the Artificial*, 3rd edition MIT Press. Traducción castellana de Joan Josep Vallbé y Pablo Noriega, *Las Ciencias de lo Artificial*, Granada: Ed. Comares, 2006.

Thiel, Peter (2012). CS183: Startup - Peter Thiel Class Notes. https://blakemasters.tumblr.com/peter-thiels-cs183-startup

US Army (2024, 6 de marzo). Army tactical intelligence targeting access node (TITAN) Ground Station Prototype — Award. Press Release. *Program Executive Office — Intelligence, Electronic Warfare & Sensors.* https://peoiews.army.mil/2024/03/06/army-tactical-intelligence-targeting-access-node-titan-ground-station-prototype-award/

Walser, Martin (1998, 11 de octubre). *Erfahrungen beim Verfassen einer Sonntagsrede*. Dankesrede von Martin Walser zur Verleihung des Friedenspreises des deutschen Buchhandels in der Frankfurter Paulskirche am 11. Oktober 1998. *Ansprachen aus Anlass der Verleihung*. Hg. Börsenverein des deutschen Buchhandels.Versión en inglés: *German history in documents and images*. Vol. 10. *One Germany in Europe, 1989-2009*. Writer Martin Walser reflects on the difficulties of living with German Guilt.