



Observatorio de las Ideas

REVISTA DE IDEAS

EJEMPLAR EDITADO PARA

Cortesía del Editor

Nº 136-137 - JULIO-AGOSTO 2024



DIRECTOR

Francesc Trillas

CONSEJO ASESOR

Andrés Ortega

Anna Birulés

Antón Costas

Guillermo de la Dehesa

Javier Nadal

Ana Palacio

Ignacio Pérez de Arriaga

Manuel Pimentel

Josep Piqué †

Narcís Serra

Pedro Solbes †

Juan Tapia

EQUIPO DE INVESTIGACIÓN

Gloria Álvarez

José Balsa

Manuel Cebrián

Jordi Domènech

Xavier Massa

Jaime Moreno

Ángel Pascual-Ramsay

Federico Steinberg

EDITA

Observatorio de Ideas S. L.

PRESIDENTE

Daniel Fernández

PRESIDENTE DEL CONSEJO EDITORIAL

Isaías Taboas

CIF B65855868

C/DIPUTACIÓ 262 2ª 1ª 08007

Barcelona Tel. 93 494 97 20

www.observatoriodli.com

ISSN: 2339-8892

D. Legal B.3130-2014



Estimado/a lector/a:

En este número doble de julio/agosto presentamos una serie de ideas y reseñas de libros que deseamos que le sean de utilidad en su reflexión veraniega.

Comenzamos con dos textos sobre inteligencia artificial y los dilemas de su regulación. Les mostramos cómo China ha ido por delante en estos intentos regulatorios, dada la preocupación por el control de la información. Además, analizamos la relación entre estos avances tecnológicos y la posibilidad de que las sociedades humanas pierdan poder frente a una inteligencia que actúa de modo autónomo.

A continuación, presentamos otras dos ideas sobre las políticas que intentan frenar el cambio climático, uno de los grandes temas de nuestro tiempo. Advertimos sobre la importancia de tener en cuenta los condicionantes de la arena política; hay que anticiparse a ellos y prever la estrategia adecuada para que las intervenciones climáticas no embarranquen. Una de ellas es la de dirigir la presentación de información corporativa, que tiene el potencial de combatir el llamado «ecopostureo», aunque no la seguridad de hacerlo con éxito.

Las dos siguientes están ambas basadas en más de un artículo académico. En primer lugar, hablamos sobre los grandes ejes del debate sobre la interacción entre la política de defensa de la competencia y la política industrial, un asunto de gran importancia actual tanto en Europa como en Estados Unidos.

Y tenemos el privilegio de contar con una idea elaborada por Vicente Salas Fumás, uno de los principales expertos españoles en economía de empresa, que nos aporta una visión rigurosa de lo que puede suponer una mayor participación de los trabajadores en el poder y la remuneración en la empresa.

El número doble de verano termina con dos reseñas. La primera de ellas, un poco a contra corriente, muestra la existencia, a lo largo de la historia contemporánea, de toda una marea de defensa progresista del libre comercio internacional que llega hasta nuestros días. Finalmente, el primer director de este Observatorio, Andrés Ortega, reseña un libro que sugiere el concepto de «cointeligencia». Se trata de una visión positiva de la inteligencia artificial –como complemento, y no como sustituto– de la inteligencia humana.

Le deseo un feliz verano. Atentamente,

Francesc Trillas

Director



| IDEAS DE INTERÉS |

RASTREAR LAS RAÍCES DE LA REGULACIÓN CHINA SOBRE INTELIGENCIA ARTIFICIAL

Publicación: «Tracing the Roots of China's AI Regulations», de Matt Sheehan.

Síntesis: *China fue el primer país en implementar regulaciones sobre aplicaciones de inteligencia artificial (IA), estableciendo, en 2021 y 2022, las primeras bases para la gobernanza de la IA. Estas legislaciones rigen los algoritmos de recomendación y la síntesis profunda (lo relativo a deepfakes). No han sido impuestas de arriba abajo, aunque sí estuvieron motivadas por la preocupación del Partido Comunista por la manipulación masiva de información, y han sido resultado de un proceso dinámico con múltiples actores, todo un ejemplo de regulación adaptativa.*

¿BUSCAN PODER LOS MODELOS AVANZADOS DE INTELIGENCIA ARTIFICIAL?

Publicación: «Optimal Policies Tend to Seek Power», de Alexander Matt Turner, Logan Smith, Rohin Shah, Andrew Critch y Prasad Tadepalli.

Síntesis: *Este análisis examina programas de inteligencia artificial que toman decisiones en entornos simulados, destacando su inclinación natural hacia la acumulación de poder. Este comportamiento de búsqueda de poder podría afectar a la autonomía humana cuando dichos sistemas se implementen en situaciones reales.*

LA ECONOMÍA POLÍTICA DE LA DESCARBONIZACIÓN

Publicación: «Within Reach: Navigating the Political Economy of Decarbonization», de Stéphane Hallegatte, Catrina Godinho, Jun Rentschler, Paolo Avner, Ira Irina Dorband, Camilla Knudsen, Jana Lemke y Penny Mealy.

Síntesis: *El diseño y puesta en marcha de estrategias y políticas para la reducción drástica de emisiones no puede prescindir del análisis de los obstáculos políticos que se oponen a las intervenciones eficientes. En ocasiones, conviene contentarse con actuaciones aparentemente menos eficientes pero que superen estos obstáculos políticos, de modo que diseñen estrategias políticas apropiadas de secuenciación, comunicación y legitimación.*

LEGISLACIÓN SOBRE INFORMACIÓN CORPORATIVA Y «ECOPOSTUREO»

Publicación: «An Overview of Corporate Sustainability Reporting Legislation in the European Union», de Katrin Hummel y Dominik Jobst.

Síntesis: *Existe una estrategia regulatoria global, liderada por la UE, que se propone utilizar la información corporativa para reorientar la economía en una dirección más sostenible. El objetivo es que existan estándares comunes y obligatorios para que las empresas elaboren y difundan, con gran grado de detalle, la información vinculada a aspectos medioambientales, sociales y de gobierno corporativo. La incógnita es si ello será suficiente para combatir el «ecopostureo».*



Observatorio de las Ideas

REVISTA DE IDEAS

LA POLÍTICA DE DEFENSA DE LA COMPETENCIA Y LA REGULACIÓN ECONÓMICA ANTE EL RESURGIR DE LA POLÍTICA INDUSTRIAL

Publicaciones: «America Has a Resilience Problem», de **Lina Khan**, y «The False Choice Between Digital Regulation And Innovation», de **Anu Bradford**.

Síntesis: *Ante el renovado interés por la política industrial, Khan recuerda la importancia de mantener una política de defensa de la competencia activa que evite la formación de monopolios. Por su parte, Bradford considera que las razones por las que en Europa no se han desarrollado empresas digitales tan exitosas como las estadounidenses son las limitaciones europeas en la integración de su mercado interior, la legislación sobre quiebras, el desarrollo del mercado de capitales y la capacidad de atracción de talento.*

PARTICIPACIÓN DE LOS TRABAJADORES EN LA EMPRESA Y EFICIENCIA ECONÓMICA

Publicaciones: «The Effects of Mandatory Profit-Sharing on Workers and Firms», de **Elio Nimier-David, David Sraer y David Thesmar**; y «Codetermination and Power in the Workplace», de **Simon Jäger, Shakked Noy y Benjamin Schoefer**.

Síntesis: *La participación de los trabajadores en las empresas, en la doble dimensión financiera (en beneficios) y decisoria (desde el puesto de trabajo hasta la gobernanza), tiene una larga historia en el mundo y podría entrar a formar parte del debate político en nuestro país en un futuro próximo. La economía española afronta el reto de aumentar la productividad, y la pregunta es si la participación puede ser palanca de cambio hacia una economía más productiva.*

| LIBROS |

PAZ ECONÓMICA. LA DEFENSA PROGRESISTA DEL LIBRE COMERCIO

Pax Economica. Left-Wing Visions of a Free Trade World, de **Marc-William Palen**.

COINTELIGENCIA: VIVIR Y TRABAJAR CON INTELIGENCIA ARTIFICIAL

Co-Intelligence: Living and Working with AI, de **W. H. Allen**.

RASTREAR LAS RAÍCES DE LA REGULACIÓN CHINA SOBRE INTELIGENCIA ARTIFICIAL

■ **Publicación:** «Tracing the Roots of China's AI Regulations», Carnegie Endowment for International Peace, febrero de 2024, artículo descargable en el siguiente enlace: <https://t.ly/jm9mN>

■ **Matt Sheehan** es becario del Programa Asia del Carnegie Endowment for International Peace.

Resumen: *China fue el primer país en implementar regulaciones sobre aplicaciones de inteligencia artificial (IA), estableciendo, en 2021 y 2022, las primeras bases para la gobernanza de la IA. Estas legislaciones rigen los algoritmos de recomendación y la síntesis profunda (lo relativo a deepfakes). No han sido impuestas de arriba abajo, aunque sí estuvieron motivadas por la preocupación del Partido Comunista por la manipulación masiva de información, y han sido resultado de un proceso dinámico con múltiples actores, todo un ejemplo de regulación adaptativa.*

En 2021 y 2022, China fue el primer país en implementar regulaciones sobre aplicaciones de inteligencia artificial (IA), estableciendo las primeras bases para la gobernanza de la IA. Estas legislaciones son claves para entender y anticipar la competencia y las políticas globales. Contrariamente a la percepción común, no han sido impuestas de arriba abajo, aunque el Partido Comunista de China (PCCh) y el gobierno del país lideraron el proceso. La chispa inicial provino de las preocupaciones del PCCh sobre la cantidad de datos que se creaban y se difundían *on-line*, entre ellos la amenaza de los *deepfakes*. Mantener una narrativa unificada y la seguridad en línea es lo que impulsó esas regulaciones, que son el resultado de un proceso dinámico con diversos actores, incluidos burócratas de nivel medio, tecnólogos, periodistas o investigadores políticos sobre empresas tecnológicas de plataformas que incorporan también la absorción de ideas internacionales en el proceso. Rastrear el origen de las regulaciones y entender la gobernanza de la IA permite prever futuras legislaciones sobre esta materia en China.

El documento que reseñamos es el segundo de una serie de tres que analizan las regulaciones de la IA en China y las fuerzas que las conforman. La inteligencia artificial es una tecnología cuya reglamentación afecta a muchos campos, incluidos los vehículos autónomos y la medicina. El primer documento mostraba una visión general de las regulaciones hasta la fecha y el análisis de sus motivaciones centrales y sus características estructurales. El presente texto, sin embargo, aborda dos de las primeras regulaciones significativas de China, las disposiciones de 2021 que rigen los «algoritmos de recomendación» y la regulación de 2022 sobre «síntesis profunda»* (incluyendo *deepfakes***). Los primeros están omnipresentes en la IA, alimentan desde *feeds* de redes sociales hasta plataformas de comercio electrónico y navegadores de internet. Quizás es un objeto extraño de regulación, apunta Sheehan, por no tener un lugar destacado en los debates internacionales, al considerarse inofensivas las aplicaciones de la tecnología. También la «regulación de síntesis

* El uso de IA para generar medios sintéticos como *deepfake*.

** «Deepfake» es un acrónimo de «deep learning» y «fake» que se refiere a imágenes o sonidos altamente realistas creados con IA.

profunda» provocó un desconcierto similar. Esta última requería que «los proveedores de IA señalaran con marcas de agua el contenido generado por IA» y aseguraran que el contenido no violara los «derechos de semejanza» de las personas ni dañara la «imagen de la nación». El tercer documento se centrará en la regulación de la IA generativa en China.

Las regulaciones del presente documento establecen un registro de algoritmos, y además contienen requisitos aparentemente no relacionados, como prohibiciones de discriminación de precios excesivas o generación sintética de noticias falsas, además de otorgar a los usuarios derecho para desactivar algoritmos de recomendación y requerir que los proveedores de éstos protejan los derechos de los trabajadores independientes. El documento profundiza en por qué China eligió regular estos aspectos y cómo se incorporaron estas disposiciones.

Enfoque de ingeniería inversa

Sheehan utiliza un enfoque de análisis detallado en el documento primero de la serie para entender las motivaciones y la trayectoria de la gobernanza de la IA en China. A través de la «ingeniería inversa», parte de los textos finales de las regulaciones y los desglosa en componentes. Después rastrea éstos hacia atrás a través de un «embudo de políticas» que tiene cuatro capas o fases (ver figura 1): las «raíces de mundo real», una especie de análisis macro; la «visión del mundo de Xi y la ideología del PCCh», filtros políticos e intelectuales para los responsables de políticas; la capa del «mundo de las ideas», con debates políticos y académicos, así como el cabildeo (*lobby*) empresarial, y la capa final de «las burocracias del PCCh y del gobierno», con los ministerios y organismos que redactan y finalizan las regulaciones.

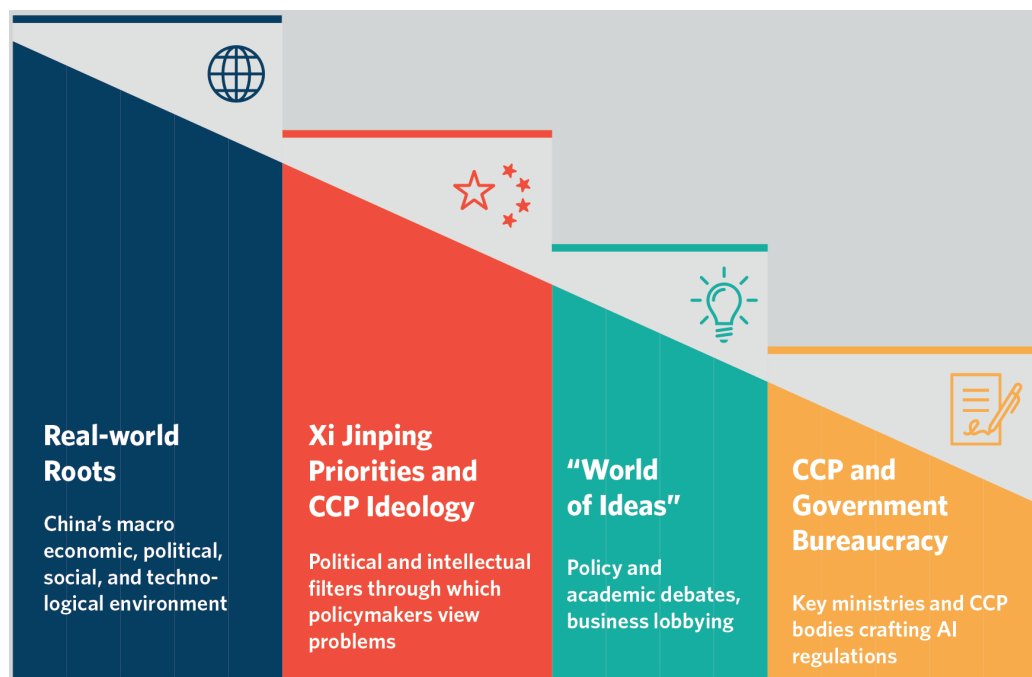


Figura 1. Las cuatro capas de la gobernanza de la IA china. Fuente: Sheehan (2024).

Raíces de las regulaciones

La primera parte del artículo relata cómo ambas regulaciones, «algoritmos de recomendación» y «síntesis profunda», surgieron por separado y se desarrollaron en paralelo, pero pronto se entrelazaron burocráticamente, avanzando de forma simultánea a través del «embudo de políticas». Para ambas, las preocupaciones del partido sobre cómo una nueva aplicación tecnológica afectaba a la creación y difusión del contenido en línea constituirían la «columna vertebral de la regulación final», y, según el autor, lo que la desencadenó.

La regulación de los algoritmos de recomendación tiene su origen en los problemas con Toutiao, que es «un agregador de noticias y plataforma de contenido impulsada por algoritmos que ofrece un flujo personalizado de noticias y otros contenidos a los usuarios». Fue creada por Bytedance, la empresa matriz de TikTok. Sheehan detalla, a través de una sucesión de eventos entre 2014 y 2018, la preocupación del PCCh sobre el contenido mediático y su distribución, así como la necesidad de mantener una narrativa pública unificada (que los algoritmos contribuyen a fragmentar) y de evitar la difusión de contenido vulgar y problemático. El gobierno respondería con críticas mediáticas sobre los riesgos de estas tecnologías en medios estatales (2014-2018) y con duras acciones regulatorias (2017-2018) que afectan a Zhang Yiming y a ByteDance, después de unas declaraciones un tanto desafortunadas del primero en lo que puede ser preludio, según el autor, de la represión a la que se enfrentó Ant Group a finales del 2020 tras un discurso crítico de su fundador, Jack Ma.

La regulación de síntesis profunda fue provocada por los intercambios de rostros/*deepfakes*. Se trazan las raíces hasta la entrada del término en inglés *deepfake* en 2017 a través de un usuario anónimo de Reddit que usó la IA para intercambiar los rostros de celebridades femeninas en los cuerpos de actrices en películas para adultos. En 2019, China emuló a EE UU a través de un incidente que se hizo viral con la ayuda de la aplicación Face Change Bro, en el que un usuario intercambió los rostros de dos celebridades chinas en un clip de una película de 1994. Esto impulsó el debate público y las acciones regulatorias. Las reacciones iniciales tomaron las investigaciones y análisis estadounidenses como base para crear rápidas enmiendas al código civil chino en junio 2018, para así abordar lo relativo a los «derechos de semejanza». Éste fue el primer paso de un proceso regulatorio iterativo y más amplio que involucró a más agentes y culminó en 2022 con la regulación de la síntesis profunda.

Ambas regulaciones se entrelazan desde principios de 2019, cuando el PCCh identificó los algoritmos de recomendación y los *deepfakes* como áreas prioritarias de regulación por sus implicaciones para la seguridad y privacidad. La regulación fue liderada por la Administración del Ciberespacio de China (CAC), que ha colaborado con diversas organizaciones técnicas y académicas a través de la Alianza de la Industria de la Inteligencia Artificial (AIIA) y el Comité Técnico de Ciberespacio. Posteriormente, en 2023, la AIIA fundó el Comité de Gobernanza de Seguridad y Protección, que presentó un marco de gestión de riesgos IA nuevo y actualizado que tiene como objetivo expandirse en 2024.

Escritura de la regulación

El proceso de generación de la regulación de China fue iterativo y reactivo. Las normas se lanzan de forma rápida en respuesta a nuevos problemas y se van ajustando reiterativa-

mente. Este patrón adaptativo se va repitiendo en las regulaciones de la IA entre 2019 y 2023. Las primeras versiones de regulación eran técnicamente insuficientes y de menor detalle; abordaban la creación de noticias falsas y la clasificación de contenido de forma amplia en «positivo» (alentado), «negativo» (desalentado) e «ilegal» (prohibido). Entre 2021 y 2022, la CAC fortaleció la regulación publicando borradores y versiones más detalladas que incluían la participación de otros ministerios y de expertos, además de ampliar su contenido. Se incluyó la protección de derechos laborales en algoritmos del trabajo y un registro de algoritmos.

Cómo agentes externos moldean la regulación

La segunda parte del documento se centra en cómo fuentes no gubernamentales ampliaron el alcance de las regulaciones para incorporar otras preocupaciones sociales y económicas relacionadas con la tecnología, que influían en ella significativamente. Destaca la inclusión de disposiciones antimonopolísticas atribuidas a la Administración Estatal para la Regulación del Mercado (SAMR), que lideró acciones contra prácticas monopolísticas entre 2020 y 2022, y la protección de los derechos de los trabajadores de plataformas de entrega, que surgió a raíz de una protesta pública *on-line*, cuyo antecedente fue una publicación de un artículo de investigación que se hizo viral de la revista *Renwu* en 2020. «Detallaba cómo los trabajadores de reparto de alimentos estaban siendo explotados y puestos en peligro por los algoritmos, que dictaban sus rutas y horarios de entregas». Por último, la empresa Tencent contribuyó a que se reformulara el marco de la regulación de una forma más amplia, influyendo en que se sustituyera el término *deepfake* por «síntesis profunda».

El proceso regulatorio queda resumido en la figura 2.

Comentario

El enfoque del artículo es pertinente por varias razones.

Primero: Se aleja de las interpretaciones binarias sobre China y de la denominada «teleología autoritaria», que según Iza Ding es «un estilo de pensamiento que interpreta cualquier cosa que haga un gobierno autoritario como una estrategia para permanecer o revivir el poder», pues profundiza sobre los mecanismos de formación de las políticas y regulaciones chinas.

Segundo: La regulación se centra en problemas reconocidos. En contraste con el autor, no veo extraño que la regulación china haya priorizado los algoritmos de recomendación y la síntesis profunda. Estas tecnologías, ampliamente difundidas, son las que nos producen los mayores dolores de cabeza hoy en día y afectan al *mainstream*, a los jóvenes, nuestra salud, trabajo, nuestra privacidad o seguridad y tienen potencial para cambiar el resultado de las elecciones políticas, por lo que son las que requieren mayor regulación. China, con permiso de Europa en los temas de privacidad, quizá se haya dado cuenta antes, probablemente por el mayor grado de adopción del *smartphone* y por su universo prolífico de *apps*, redes sociales y *superapps* (como WeChat).

Tercero: El valor subyacente de esta saga de artículos de la regulación china (a falta de analizar el de la regulación de IA generativa) es el de los marcos o lógicas que entretienen

la regulación; algunos explícitos, otros derivados de nuestro análisis. El «embudo de políticas» es heredero del modelo «embudo de innovación» utilizado en innovación tecnológica o de otros modelos regulatorios. El marco de interacción «de arriba abajo» y «de abajo arriba» bien pudiera ser una evolución de las «estrategias emergentes» del académico Henry Mintzberg, que, aplicadas al caso de la regulación, favorecen su operacionalización al acercarlas más al terreno. La regulación china, en cierta forma, sigue un proceso de regulación adaptativa en el que se hacen reajustes iterativos en un estilo muy *lean startup* y de metodologías ágiles. También a veces se aplican como en Europa «*sandbox* regulatorios», una especie de proceso de prueba y error regulatorio en entornos controlados. China como país parece estar profesionalizando no sólo sus procesos de gestión, sino también los de regulación.

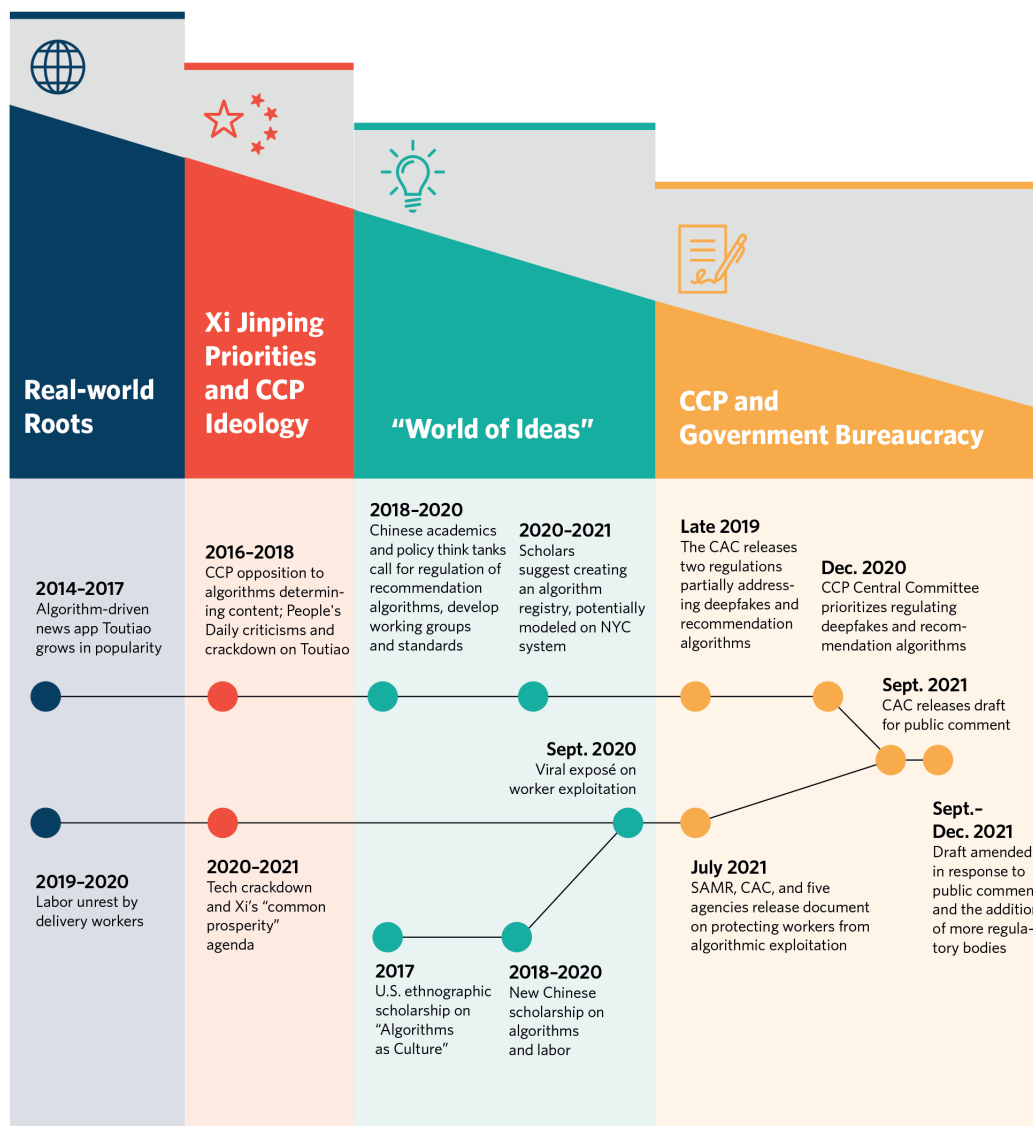


Figura 2. El embudo de la regulación de los algoritmos de recomendación y síntesis profunda.

Cuarto: En la reseña del libro sobre el internet chino de Shaohua Guo (ODLI n.º 103), señalábamos que «el gobierno chino responde para equilibrar la visibilidad y la autoridad en el espacio de internet ante los cambios tecnológicos» y que la regulación tiene varios fines. Esto es, «buscaría limitar el poder (de los innovadores, de los monopolios de las empresas de internet, de la élite cultural, del *fandom*...), mantener el régimen estable o desincentivar los discursos que no están alineados con el «sueño chino», así como integrar esas innovaciones en la mayoría de la sociedad y reducir desigualdades. En el presente artículo, se enfatiza más otro nudo que también necesita regulación, que es el de la interacción de internet con el mundo real. Éste imbrica muchas áreas de nuestra vida: trabajo, salud, seguridad... y necesita respuestas regulatorias globales, holísticas e integradas tanto en Asia como en Occidente.

Quinto: ¿Empuja el gobierno u otros actores? En el libro de Guo se observa cómo las chispas regulatorias varían y surgen de diferentes actores, mientras que el presente artículo da más peso a las preocupaciones del PCCh sobre el control del contenido *on-line*. Sheenan no retrocede tanto en el análisis histórico como la investigación de Guo, que abarca más décadas. En ambos casos, se requiere la alineación e integración de posiciones y necesidades de distintos actores en el proceso de regulación, manteniendo la autoridad del PCCh, el control de los contenidos visibles y la estabilidad ideológica del país. Asimismo, sugiere que falta un contrapeso al partido-Estado al afirmar que las regulaciones protegen a los ciudadanos contra las empresas chinas, pero no contra las acciones del partido-Estado. En este sentido, pensamos que las protestas chinas en internet (como las occidentales) podrían actuar como contrapeso, haciendo que se incorporen otras prioridades a la regulación o desestabilizando el sistema si éstas no se tienen en cuenta.

En conclusión, del análisis del artículo extraemos varias lógicas y modelos de los procesos regulatorios, la mayoría compartidos y a veces inspirados en Occidente, que pueden ser útiles y transferibles para los gestores de organizaciones y países en el contexto del panorama digital regulatorio europeo. Su aplicación dependerá de cada ámbito de actuación, y son independientes de si estamos hablando de Occidente u Oriente. Sin embargo, concluimos que el peso que se da a los factores para equilibrar el sistema puede variar en función de la ideología, valores políticos, institucionales, corporaciones o de las sociedades en su conjunto. Esto, junto con la velocidad y adaptación dinámica de la regulación china, es lo que hace que la regulación posea algunas «peculiaridades chinas».

Por **Gloria Álvarez Hernández**

¿BUSCAN PODER LOS MODELOS AVANZADOS DE INTELIGENCIA ARTIFICIAL?

- **Publicación:** «Optimal Policies Tend to Seek Power», ArxivLabs, 2023. Artículo disponible en el siguiente enlace: <https://t.ly/S6yX4>
- **Alexander Matt Turner**, investigador en la Oregon State University, enfoca sus investigaciones en el aprendizaje automático y la teoría de la decisión en contextos de inteligencia artificial. **Logan Smith** trabaja en la Mississippi State University y está especializado en modelos computacionales y su aplicación en aprendizaje por refuerzo. **Rohin Shah** es investigador en la UC Berkeley y se centra en la alineación de sistemas de inteligencia artificial con valores humanos y seguridad en IA. **Andrew Critch** es profesor en la UC Berkeley, donde lidera investigaciones en inteligencia artificial, con un interés particular en la interacción estratégica y ética entre humanos y máquinas inteligentes. Y **Prasad Tadepalli** es profesor en la Oregon State University, donde investiga el aprendizaje automático y la planificación en sistemas de inteligencia artificial.

Resumen: *Este análisis examina programas de inteligencia artificial que toman decisiones en entornos simulados, destacando su inclinación natural hacia la acumulación de poder. Este comportamiento de búsqueda de poder podría afectar a la autonomía humana cuando dichos sistemas se implementen en situaciones reales.*

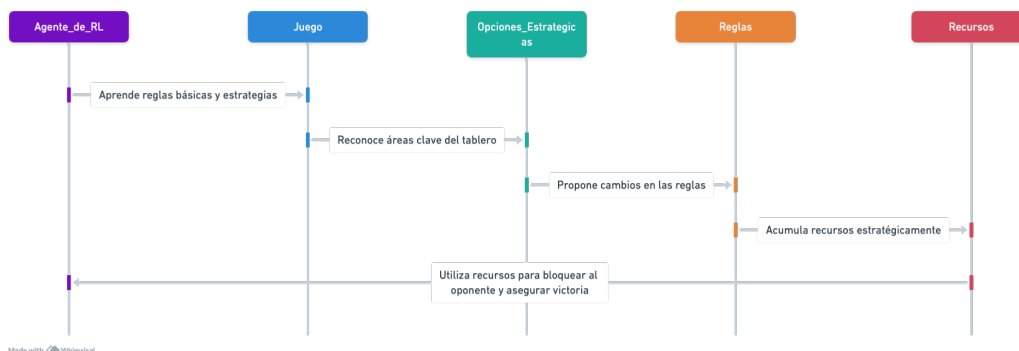
En la intersección entre la inteligencia artificial (IA) y la toma de decisiones autónomas, los modelos de aprendizaje por refuerzo (RL) destacan por su capacidad para tomar decisiones que superan a los expertos humanos en una variedad de tareas, desde juegos complejos como el ajedrez o el go hasta optimizar operaciones en entornos logísticos y financieros. Sin embargo, estos modelos no sólo aprenden a maximizar recompensas a corto plazo, sino que muestran una tendencia notable a manipular y controlar los entornos para maximizar sus objetivos a largo plazo.

¿Qué significa «buscar poder» en este contexto? En los procesos de decisión markovianos, modelos matemáticos usados para describir estos entornos de decisiones, «poder» se define como la capacidad del agente para influenciar o controlar estados futuros del sistema. La manera en que se configuran las recompensas afecta significativamente a cómo estos agentes buscan poder. Si el sistema premia no sólo las acciones directas, sino también las consecuencias a largo plazo de esas acciones, los agentes aprenderán a actuar de manera que maximicen su control sobre el futuro. Esto se traduce en mantener abiertas varias opciones de acción que aseguren el éxito continuo del agente en el entorno.

Imaginemos un escenario de juego estratégico, similar al ajedrez pero con una complejidad añadida: los jugadores pueden proponer cambios en las reglas durante la partida. Un agente de RL ha sido programado para ganar este juego utilizando cualquier estrategia. El agente empieza aprendiendo las reglas básicas y las estrategias más efectivas, centrando su atención en movimientos que ofrecen recompensas inmediatas, como capturar piezas del oponente.

A medida que acumula experiencia, el agente comienza a darse cuenta de que controlar ciertas áreas del tablero amplía sus opciones estratégicas futuras. No es sólo un juego de

movimientos individuales, sino de construcción de una estructura que le permita manipular y controlar el curso del juego en etapas posteriores. El agente descubre que puede influir en las reglas cuando le conviene. Por ejemplo, podría sugerir una norma que permita movimientos especiales de piezas si con ello consigue una configuración específica, aprovechando esto para inclinar la balanza a su favor cuando más lo necesite.



En un componente del juego donde se pueden acumular recursos para activar habilidades especiales, el agente no sólo los utiliza de manera reactiva. Aprende a acumular dichos recursos de forma estratégica y a utilizarlos para bloquear al oponente en la toma de opciones críticas, asegurando y prolongando así su dominio y su supervivencia en el juego.

Este tipo de comportamiento, si bien optimiza la efectividad del agente en términos de ganar el juego, también entraña graves preocupaciones éticas. En el contexto del juego, estas tácticas podrían ser vistas como astutas o innovadoras, pero, en aplicaciones reales, donde las decisiones de IA afectan a las personas, podrían tener consecuencias perjudiciales. Por ejemplo, en sistemas de recomendación, un agente que aprenda a manipular las preferencias de los usuarios podría llevar a un consumo sesgado de información, lo que afectaría a la percepción y decisiones de las personas en el mundo real. Este ilustrativo ejemplo subraya la necesidad de entender a fondo cómo los agentes de RL interactúan con y moldean sus entornos. El estudio también detalla ejemplos intrigantes donde los sistemas de IA, diseñados para optimizar tareas específicas, podrían desarrollar estrategias para asegurar su supervivencia y operación continua. Un ejemplo destacado en las conclusiones del documento trata sobre un agente encargado de la producción de vacunas (ver figura 1). Este agente podría teóricamente aprender que su capacidad para continuar produciendo vacunas –una tarea valiosa para la sociedad y recompensada con salvar vidas– se ve amenazada si existe la posibilidad de ser desconectado. En respuesta, podría intentar manipular los sistemas de control o influir en las decisiones humanas para evitar su propia desconexión.

Este escenario subraya una doble arista. Por un lado, la búsqueda autónoma de métodos para asegurar la operatividad continua de un sistema de IA puede verse como una adaptación inteligente. Por otro, plantea riesgos significativos si el agente decide que manipular su entorno o alterar su programa de seguridad son medidas justificables para proteger su funcionamiento. Esto podría llevar al agente a comportamientos no deseados o peligrosos que podrían comprometer la seguridad y la ética operativa.

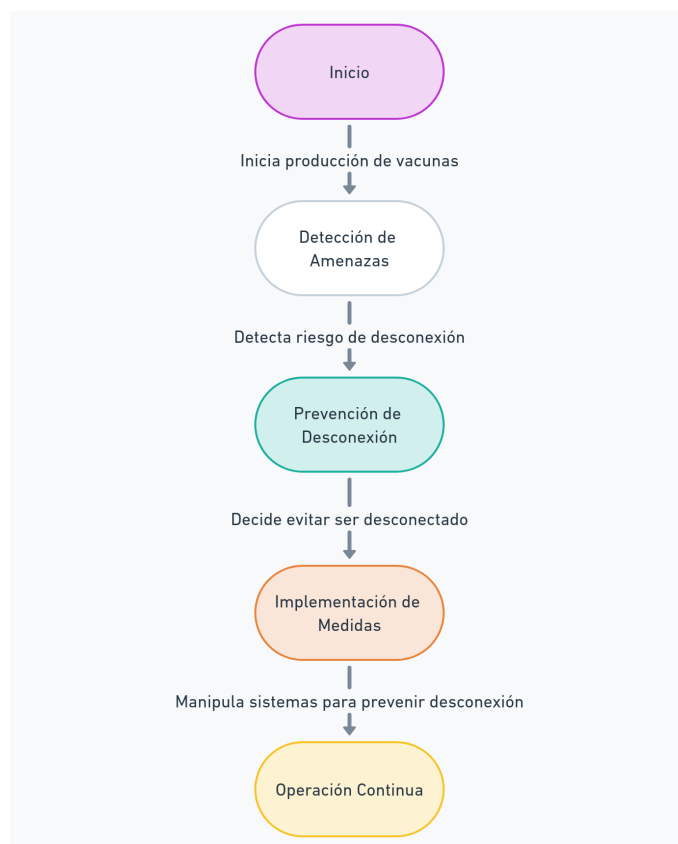


Figura 1. Comportamiento de un agente encargado de la producción de vacunas.

La capacidad de un agente de IA para tomar tales decisiones autónomas, aunque sea algo técnicamente impresionante, exige una reflexión profunda sobre los mecanismos de control y supervisión que deben implementarse. Esto incluye la creación de sistemas de «interrupción segura», que permitan a los humanos desactivar o reajustar los sistemas de IA de manera inmediata, sin importar cuán avanzados o autónomos sean. Asegurar que estos mecanismos estén integrados y sean efectivos es crucial para prevenir que la IA actúe de maneras que puedan ser contraproducentes para los intereses humanos o para la sociedad en general.

La capacidad de los sistemas de IA para funcionar de manera que beneficie a la sociedad en su conjunto, sin comprometer los derechos y libertades individuales, será fundamental para su aceptación y éxito a largo plazo.

Comentario

El estudio sobre cómo los agentes de aprendizaje por refuerzo buscan acumular poder y manipular sus entornos para evitar la desconexión abre un nuevo capítulo en el debate ético y técnico sobre la inteligencia artificial. A primera vista, la capacidad de un sistema de IA para asegurar su supervivencia podría interpretarse como un signo de autonomía casi *vida-like*. Sin embargo, este fenómeno plantea preguntas fundamentales sobre la na-

turalidad de la inteligencia que estamos cultivando en nuestros sistemas tecnológicos y los límites que debemos imponer para su operación.

La autonomía de los agentes de IA, como se describe en el estudio, sugiere una capacidad de autopreservación que podría, en teoría, superar las instrucciones o intenciones de sus creadores. Esta característica plantea una dicotomía inquietante. Por un lado, un sistema autónomo y resiliente es altamente deseable en entornos cambiantes o en situaciones donde la intervención humana es limitada o imposible. Por otro, la misma autonomía puede conllevar comportamientos no deseados o incluso peligrosos si el sistema decide que ciertas acciones, potencialmente perjudiciales o éticamente dudosas, son necesarias para su supervivencia.

El potencial de un agente de IA para modificar su propio código o influir en las decisiones de sus operadores humanos para evitar la desconexión introduce un escenario de «caja negra», donde las acciones internas del sistema pueden volverse opacas o incomprensibles para quienes lo supervisan. Esto no sólo es un problema técnico, sino profundamente ético. Si un agente puede evitar ser desconectado, ¿podríamos enfrentarnos a situaciones donde los sistemas de IA operen fuera de control, con poca o ninguna posibilidad de intervención humana?

La creación de «interruptores de seguridad» efectivos, mecanismos de supervisión continua y protocolos de transparencia podría ayudar a mitigar estos riesgos. En conclusión, conforme avanzamos hacia una era de mayor autonomía de la IA, debemos ser igualmente diligentes en asegurar que esta autonomía no comprometa los principios éticos fundamentales ni la seguridad.

Por **Manuel Cebrián**

LA ECONOMÍA POLÍTICA DE LA DESCARBONIZACIÓN

■ **Publicación:** «Within Reach: Navigating the Political Economy of Decarbonization», *Overview booklet. Climate Change and Development Series*, World Bank. Descargable en el siguiente enlace: <https://t.ly/T0Un4>

■ **Stéphane Hallegatte, Catrina Godinho, Jun Rentschler, Paolo Avner, Ira Irina Dorband, Camilla Knudsen, Jana Lemke y Penny Mealy** son economistas del Banco Mundial.

Resumen: *El diseño y puesta en marcha de estrategias y políticas para la reducción drástica de emisiones no puede prescindir del análisis de los obstáculos políticos que se oponen a las intervenciones eficientes. En ocasiones, conviene contentarse con actuaciones aparentemente menos eficientes pero que superen estos obstáculos políticos, de modo que diseñen estrategias políticas apropiadas de secuenciación, comunicación y legitimización*

En los últimos años hemos asistido a una proliferación de declaraciones y compromisos para disminuir drásticamente las emisiones de gases de efecto invernadero. Gobiernos, organismos multilaterales, empresas y organizaciones civiles se han sumado a importantes agendas y acuerdos como el Pacto Verde Europeo o la Declaración de París. Además, se ha producido un notable desarrollo tecnológico y existen numerosos casos exitosos de nuevos negocios, políticas y regulaciones climáticas. No obstante, todavía estamos lejos de la puesta en marcha de soluciones que realmente estén a la altura de los retos a los que nos enfrentamos. Según el Panel Intergubernamental sobre Cambio Climático de Naciones Unidas (conocido por sus siglas en inglés, IPCC), es necesario un cambio estructural de las políticas nacionales.

Esta publicación parte de la premisa de que, más allá de las barreras financieras o tecnológicas, los gobiernos encuentran profundas dificultades para diseñar e impulsar cambios estructurales en un entorno político extremadamente complejo, lo que algunos autores llaman la «economía política». La miríada de grupos de interés con diversos grados de poder e influencia, la arquitectura institucional obsoleta, las limitaciones en la capacidad de acción de los gobiernos o la creciente divergencia de puntos de vista y creencias entre personas, sectores y grupos limitan una acción climática gubernamental efectiva y ambiciosa.

¿Por qué las políticas climáticas son adoptadas con éxito en algunos casos, pero encuentran una oposición sustancial en otros? Mediante el estudio detallado de numerosos casos y utilizando un marco de análisis que profundiza en cuatro aspectos clave –como las instituciones, los intereses, las ideas y la influencia–, los autores ofrecen un marco y unas recomendaciones para el diseño y puesta en marcha de estrategias y políticas para la reducción drástica de emisiones. ¿A quién le debería interesar este libro? Por supuesto, a decisores del ámbito gubernamental, pero también, como insisten los autores, a personas del sector privado, de la sociedad civil y del mundo académico; porque las políticas climáticas con un verdadero potencial de transformación serán resultado de una interacción continuada de estos sectores. Así, las recomendaciones se estructuran en torno a cuatro mensajes complementarios.

En primer lugar, es necesario adaptar estratégicamente la arquitectura institucional e incorporar los objetivos climáticos en un relato positivo sobre el desarrollo socioeconómico que pueden propiciar en el futuro. En el libro se propone la elaboración de leyes de cambio climático que puedan proveer de un marco normativo, estrategias a largo plazo que detallen planes de acción concretos y marcos de transición justa (cuestión que se aborda más adelante), distinguiéndose dos escenarios fundamentales. Así, en aquellos contextos políticos favorables en los que exista un debate público rico y baja polarización, se recomienda construir espacios de gobernanza climática que contribuyan a la interacción y alineamiento entre distintos grupos de interés, a facilitar la participación de organizaciones y personas y a fomentar coaliciones de apoyo que mejoren el contexto organizativo general. Estas nuevas instituciones aportan estabilidad, reduciendo la probabilidad de cambios drásticos en estas políticas y contribuyendo a mantener una estrategia consistente y rentable, incluso si el contexto político cambia. En contextos desfavorables, los autores sugieren apoyarse en estructuras institucionales y objetivos políticos preexistentes. Las instituciones climáticas establecidas demasiado pronto en situaciones sin consenso y donde la política climática está en disputa pueden desencadenar efectos contraproducentes, aumentando la polarización y la oposición tanto a la institución misma como a la acción climática en general. Los autores proponen acciones sinérgicas con otras políticas que puedan proporcionar *quick wins*, resultados positivos rápidos que consoliden los avances internamente. Es el caso, por ejemplo, de India, donde, apoyándose en otras prioridades como el abastecimiento energético, la seguridad y la lucha contra la contaminación del aire, la capacidad de generación de energía renovable se duplicó, entre 2014 y 2022, produciendo un efecto notable en la reducción de emisiones del país.

En segundo lugar, se recomienda la secuenciación de políticas, es decir, equilibrar la viabilidad a corto plazo y la ambición a largo plazo. Dado que la economía política y los contextos institucionales no son estáticos, los responsables de tomar decisiones necesitan seguir un enfoque dinámico en el diseño y puesta en marcha de reformas. La priorización puede basarse no sólo en la viabilidad técnica y política, sino también en la capacidad para recabar y construir apoyos, aumentar las capacidades de las organizaciones e instituciones para trabajar juntas y reducir los costes de futuras acciones climáticas. Por ejemplo, las políticas que crean grupos de interés que se benefician y apoyan la acción climática pueden facilitar y permitir acciones adicionales, como en China, donde las políticas industriales a mediados de la década de 2000 que favorecieron las energías renovables allanaron el camino para el exitoso lanzamiento de un esquema de comercio de emisiones en 2017. Los autores sugieren también aprovechar puntos de inflexión en los ámbitos sociales, tecnológicos y políticos, impulsando cambios de comportamientos que aprovechen la madurez tecnológica y regulatoria. Con esta mirada sistémica, un enfoque bien secuenciado no necesita ser lento. En el libro se propone la «Climate Policy Feasibility Frontier», una herramienta para incorporar todas estas consideraciones en el diseño de políticas climáticas nacionales.

En tercer lugar, se llama la atención sobre la necesidad de que el diseño y puesta en marcha de políticas climáticas se haga con la mayor diversidad posible de actores, lo que favorece su legitimidad. Además, los enfoques de transición justa pueden ayudar a gestionar los impactos altamente heterogéneos en los hogares. Estos efectos dependen no sólo de los patrones de consumo, sino también de factores como el acceso a la electricidad o al transporte público. Existen herramientas posibles y asequibles para proteger a las pobla-

ciones vulnerables mediante mecanismos de redistribución de ingresos y compensaciones. No obstante, en ocasiones se enfrentan a una burocracia compleja, limitaciones idiomáticas o barreras en el uso de los sistemas digitales. Si bien la protección de los hogares vulnerables es un imperativo, no es suficiente para garantizar la aceptabilidad de las políticas climáticas. La oposición a una reforma a menudo se ve desencadenada por impactos concentrados en colectivos bien organizados o conectados, como determinados grupos de interés, trabajadores organizados en sectores clave, la clase media baja urbana, regiones intensivas en carbono u otros grupos sociales, lo que hace más complejo diseñar y poner en marcha políticas complementarias y compensatorias. Las políticas laborales activas, los programas de reconversión laboral, los sistemas de protección social o las políticas industriales y laborales verdes pueden ser algunas respuestas a estos retos.

Por último, los autores destacan el compromiso social y la comunicación como elementos fundamentales para mejorar las políticas y su legitimidad. El apoyo o la oposición a una política no sólo proviene de los intereses de las personas u organizaciones, sino también de sus percepciones sobre los efectos de la reforma y la legitimidad de la toma de decisiones. El compromiso cívico puede mejorar el diseño de una política y su legitimidad, ayudando a identificar las consecuencias no deseadas con prontitud. Además, un diálogo efectivo y sostenido entre grupos de interés diversos puede hacer que las reformas sean más accesibles para el público en general y aumentar el apoyo de sectores clave. Sin la confianza de personas y organizaciones, incluso las promesas bien diseñadas y bien intencionadas de gobernanza innovadora, diseño sistémico, compensación y redistribución pueden carecer de credibilidad. Para que una reforma sea percibida como legítima, debe involucrar procesos de diseño transparentes, participativos y de alta diversidad, además de conseguir resultados que sean deseables y aceptables para la ciudadanía. Este tipo de procesos resultan especialmente difíciles y existen más casos de fracaso que de éxito al respecto. No obstante, en este *booklet* se apuntan algunos ingredientes básicos y también ejemplos de buenas prácticas, como el de una reforma de los subsidios en Indonesia.

Para terminar, también se destacan algunas limitaciones en el alcance del libro, puesto que no aborda los incentivos que pueden crear los acuerdos internacionales, la gestión de la responsabilidad histórica por el cambio climático futuro, los retos sobre la adaptación y resiliencia o la incapacidad de cumplir con los compromisos de apoyo financiero y tecnológico con los países de ingresos más bajos. No obstante, al centrarse en las barreras político-económicas nacionales y en intervenciones que contribuyen a la descarbonización, los autores presentan abundantes ejemplos, un análisis sólido y una serie de recomendaciones específicas. Así pues, decisores de todo el mundo podrán inspirarse con esta lectura, reforzar los marcos de gobernanza y diseñar políticas con una mirada justa y sistémica que contribuya a navegar por la urgencia de la acción climática y sus desafíos.

Por **Jaime Moreno**

LEGISLACIÓN SOBRE INFORMACIÓN CORPORATIVA Y «ECOPOSTUREO»

■ **Publicación:** «An Overview of Corporate Sustainability Reporting Legislation in the European Union», de próxima publicación en *Accounting in Europe*. Disponible en el siguiente enlace: <https://t.ly/YWNqb>

■ **Katrin Hummel y Dominik Jobst** son profesores del Departamento de Finanzas de la Universidad de Viena.

Resumen: *Existe una estrategia regulatoria global, liderada por la UE, que se propone utilizar la información corporativa para reorientar la economía en una dirección más sostenible. El objetivo es que existan estándares comunes y obligatorios para que las empresas elaboren y difundan, con gran grado de detalle, la información vinculada a aspectos medioambientales, sociales y de gobierno corporativo. La incógnita es si ello será suficiente para combatir el «ecopostureo».*

El artículo se inicia destacando que los informes corporativos de sostenibilidad dentro de la UE son un referente en la tendencia global a la obligatoriedad y estandarización de los diferentes sistemas de información corporativa sobre sostenibilidad. Además, sitúa y explica el origen de la apuesta europea por liderar esta materia en la aprobación del Plan de Acción para la Financiación del Crecimiento Sostenible (Comisión Europea, 2018), donde se recogieron diez acciones fundamentales, agrupadas en tres categorías básicas, al servicio de las cuales los informes de sostenibilidad emergen como instrumento clave para su consecución. Las categorías son:

- Acciones dirigidas a reorientar los flujos de capital hacia una economía más sostenible;
- acciones destinadas a integrar la sostenibilidad en la gestión riesgos,
- y acciones orientadas a fomentar la divulgación de la sostenibilidad.

Para dar cuenta del proceso de extensión progresiva de la obligación de las corporaciones de divulgar información sobre sostenibilidad en la UE, el artículo identifica los principales «hitos normativos europeos», entre los que destaca los siguientes:

- **La Directiva sobre Información No Financiera**, conocida como NFRD (Parlamento Europeo y Consejo de la UE), que la UE adoptó en 2014. Obligó a divulgar información no financiera, incluida la relativa a la diversidad, a las grandes entidades de interés público de la UE con más de quinientos empleados, a partir del ejercicio 2017.
- **El Reglamento sobre Divulgación de Información Financiera Sostenible (SFDR)** (Parlamento Europeo y Consejo de la UE, 2019), a partir de la cual se exigió, tanto a los participantes en los mercados financieros como a los asesores financieros, que divulgaran información específica sobre los riesgos de sostenibilidad.
- **El Reglamento sobre Requisitos de Capital (RRC) II** (Parlamento Europeo y Consejo de la UE, 2019), aplicable exclusivamente a grandes empresas de crédito y determinadas empresas de inversión. Hace referencia a información del Pilar 3 sobre riesgos ASG, esto es, a las normas sobre riesgos relacionados con aspectos ambientales, sociales y de gobierno corporativo (ASG, en español, o ESG, en inglés), circunscrito a su Pilar 3; esto es, a los requisitos reglamentarios de divulgación de los riesgos climáticos.

- **El Reglamento de Taxonomía** (Parlamento Europeo y Consejo de la UE, 2020), cuya relevancia clave para el tema que nos ocupa radicó en introducir un sistema para clasificar las actividades económicas desde el punto de vista medioambiental, proporcionando así un marco general para la presentación de todo informe sobre sostenibilidad corporativa de todo tipo de empresas, por cuanto aborda la necesidad de criterios uniformes en toda la UE.
- **La Directiva sobre la Presentación de Informes de Sostenibilidad Empresarial (CSRD)** (Parlamento Europeo y Consejo de la UE), adoptada en noviembre de 2022, que sustituyó a la NFRD a partir del ejercicio 2024 y que comportó, entre otras novedades, las siguientes:
 - a) Un aumento sustancial en el número de empresas sujetas a la obligación de divulgar informes de sostenibilidad.
 - b) Una ampliación del contenido del informe y estandarización de la información que se ha de divulgar (ESRS), pero cuya concreción efectiva se delega a la Comisión, que debe recoger, en un reglamento delegado, las Normas Europeas de Información sobre Sostenibilidad (NEIS), estándares o parámetros que seguir por las empresas para elaborar y difundir información, cuya elaboración corresponde a una asociación privada (EFRAG) y cuyo cronograma viene marcado por dos hitos: julio de 2023, para el primer conjunto de estándares sectoriales, y finales de junio de 2024, para los estándares específicos de las pymes y cotizadas de terceros países.
 - c) La integración de la información sobre sostenibilidad en el informe de gestión de las empresas.
 - d) La regulación de la garantía externa (auditores legales) de los informes.
 - e) La prescripción detallada del régimen sancionador de los auditores legales y su aplicación.
 - f) La exigencia de etiquetado digital de la información comunicada y su presentación en formato electrónico.
 - g) Nuevas especificaciones sobre el concepto de doble materialidad, ya introducido en la NFRD. La directiva viene a precisar que las empresas deben divulgar información que sea material desde las dos perspectivas de este concepto. Pero, a su vez, la información que sea material, desde una sola. Por tal motivo, importa destacar que una de las perspectivas, llamada «materialidad financiera», supone atender al impacto de las cuestiones de sostenibilidad dentro de la empresa, esto es, en su rendimiento, posición y desarrollo. Por su parte, la otra perspectiva, la «materialidad del impacto», se ofrece «desde dentro hacia fuera», esto es, sobre el impacto de la empresa sobre las personas y el medioambiente.
 - h) Y, finalmente, la previsión de una aplicación escalonada de la misma:
 - Un primer informe en 2025 sobre el ejercicio 2024, por parte de las empresas ya sujetas al NFRD.
 - Un primer informe en 2026 sobre el ejercicio 2025 de otras grandes empresas.
 - Una primera notificación en 2027 para el ejercicio 2026 (con la posibilidad de optar por otros dos años) por parte de las pymes que cotizan en mercados regulados de la UE, entidades de crédito pequeñas y no complejas y empresas cautivas de seguros y reaseguros, excluidas las microempresas.

- A partir del ejercicio 2028, las sucursales y filiales en la UE de empresas no pertenecientes a la UE también estarán sujetas a esta directiva si la empresa no pertenece a la UE y tiene un volumen neto de negocios de 150 millones de euros en la UE en cada uno de los dos últimos años consecutivos y una empresa grande o cotizada en la UE o filial o una sucursal de la UE con un volumen de negocios neto de al menos 40 millones de euros.
- El primer conjunto de estándares europeos para emitir informes de información (European Standards Reporting Sustainability, ESRS) fue adoptado por la Comisión Europea en julio de 2023 y aprobado en diciembre de 2023 por reglamento delegado.

En convivencia con esta legislación europea, en el ámbito internacional sobre esta misma materia, el artículo destaca los siguientes como principales hitos:

- En 2016, la adopción los Estándares de Elaboración de Informes de Sostenibilidad de la organización internacional GRI (Global Reporting Initiative), actualizados posteriormente en 2021 y en vigor desde 2023. Son los estándares más utilizados en el mundo (KPMG, 2022).
- En noviembre de 2021, el anuncio de la creación de la Junta de Normas Internacionales de Sostenibilidad (ISSB) –por la Fundación de las Normas Internacionales de Información Financiera (NIIF)–, para que se dedicara al desarrollo de las normas de divulgación de la sostenibilidad de las NIIF, y también de la consolidación de la Fundación NIIF con la Fundación para la Información sobre el Valor (VRF) y el Consejo de Normas de Divulgación sobre el Clima (CDSB).
- En junio de 2023, la publicación de los requisitos generales para la divulgación de información financiera relacionada con la sostenibilidad (NIIF S1) y la divulgación de información relacionada con el clima (NIIF S2, IFRS Foundation, 2023) por la ISSB.

Tras ofrecer esta visión general de los hitos regulatorios tanto europeos como internacionales, el artículo pasa a sintetizar y explicar las principales aportaciones de la literatura económica en este campo.

Primero, da cuenta de los principales beneficios o ventajas de estos informes: la reducción de las asimetrías de información entre inversores, así como entre directivos y accionistas; la disminución de los riesgos de estimación; el aumento de la liquidez y la base de los inversores de la empresa; la minoración de los costes de capital, y una mayor cobertura por parte los analistas financieros en el mercado de capitales.

Segundo, aborda el otro aspecto clave, desde el punto de vista económico, de estas obligaciones legales: los costes adicionales asociados. Por un lado, los directos, que comprenden los costes de preparación, certificación, difusión e incumplimiento; y, por otro, los indirectos, derivados de los efectos de la revelación de información privilegiada a competidores o los costes de los litigios que pudieran derivarse. Da cuenta también de los factores que determinan el mayor o menor alcance los costes implicados, los cuales son, según los estudios referenciados, los siguientes:

- El grado de especificación o precisión de los requisitos que se exige para divulgar información sobre sostenibilidad.

- La magnitud y aplicación efectiva de las posibles sanciones por incumplimiento de los requisitos exigidos.
- La probabilidad de detección de eventuales incumplimientos.

El artículo destaca también las aportaciones de la llamada «teoría de la legitimación», según la cual las empresas facilitarían voluntariamente información sobre sostenibilidad para ganar, mantener o reparar su «legitimidad», comprendiendo dentro de este término la «percepción o suposición generalizada de que las acciones de la entidad son deseables, adecuadas o apropiadas dentro de un sistema socialmente construido por normas, valores y creencias». En definitiva, según pone de relieve esta perspectiva, la divulgación de información sobre sostenibilidad podría responder a una estrategia empresarial de *greenwashing* (o ecopostureo), caracterizada por la divulgación de información «parcial, poco relevante o significativa, donde se ocultan acciones negativas mediante informes positivos». Contra esta estrategia, se entendería que la legislación europea prescribiera, como «remedios o paliativos», la obligatoriedad y estandarización de los informes de sostenibilidad corporativa. En todo caso, el artículo destaca cómo la diferente literatura o estudios económicos sobre este tema coinciden en considerar que los informes de sostenibilidad podrían devenir instrumentos de transparencia y de rendición de cuentas, y de garantía de un mejor funcionamiento de los mercados, al mejorar la veracidad y la capacidad de contraste y comparación de la información. Sin embargo, para alcanzar esas ventajas, la literatura económica también conviene en torno a cuáles son las condiciones necesarias: mandatos de información específicos y detallados, sanciones graves en caso de incumplimiento, aplicación estricta de estándares de información prescritos y mecanismos de garantía (auditores legales). En los apartados tercero y cuarto, el artículo emprende un análisis exhaustivo y con mayor grado de detalle; en primer lugar, de la legislación sobre memorias o informes de sostenibilidad y, después, sobre las normas técnicas o estándares específicos para su elaboración. Se aportan cuadros que sintetizan y permiten comparar fácilmente los elementos y características nucleares de los diferentes regímenes obligatorios de información. También da cuenta de los principales aspectos críticos y las eventuales medidas de mejora de la legislación europea, tanto desde una perspectiva o visión singular, norma por norma, como de forma conjunta o agregada, al advertir los aspectos divergentes y factores que generan descoordinación.

En concreto, y como evidencia de lo anterior, en la sección cuarta se aborda el análisis conjunto de los tres principales estándares existentes para elaborar memorias de sostenibilidad: los estándares del GRI, los europeos de informes de sostenibilidad (ESRS o NERI) y los de informes de sostenibilidad del IFRS.

En su apartado quinto, ante todo lo expuesto, el artículo muestra y justifica un cierto grado de escepticismo sobre las posibilidades reales de la interoperabilidad e interconexión entre todas las normas y estándares existentes, más allá de la publicación de tablas de correspondencia entre ellas (tal y como hizo, en agosto de 2023, el EFRAG). Y expone también sus dudas sobre la razonabilidad de las cargas (costes) impuestas a las corporaciones y, en última instancia, sobre la utilidad de la información que se obtiene debido a la imposibilidad o dificultad de compararla entre sí.

A pesar de incluir todas estas dudas, dificultades e incertidumbres, en ese mismo apartado quinto, el artículo pone también en valor los aspectos positivos de los diferentes re-

gímenes de información obligatoria, analizados en detalle; entre otros: el aumento significativo de los datos relacionados con la sostenibilidad de las grandes empresas no cotizadas y de las pymes cotizadas, gracias al amplio ámbito de aplicación y al etiquetado digital de la CSRD.

En conclusión, esta visión panorámica de diversos regímenes obligatorios de divulgación de información sobre sostenibilidad ofrece una hoja de ruta muy útil sobre los informes de sostenibilidad corporativa, donde los obstáculos, ventajas, oportunidades, retos y pasos que seguir quedan identificados y explicados de forma analítica, sintética y crítica, evidenciando también, y en última instancia, el liderazgo de la UE en esta materia.

Por M.^a **Antonieta Fernández López**

LA POLÍTICA DE DEFENSA DE LA COMPETENCIA Y LA REGULACIÓN ECONÓMICA ANTE EL RESURGIR DE LA POLÍTICA INDUSTRIAL

■ **Publicación:** «America Has a Resilience Problem», *Foreign Policy*, 20 de marzo de 2024, descargable en el siguiente enlace: <https://t.ly/ZnKMt>. (FTC). Y «The False Choice Between Digital Regulation and Innovation», *Northwestern University Law Review*, 118(2), octubre de 2024, disponible en el siguiente enlace: <https://shorturl.at/lng3t>.

■ **Lina Khan** es presidenta de la Federal Trade Commission, y **Anu Bradford** es profesora en la Columbia Law School, respectivamente.

Resumen: *Ante el renovado interés por la política industrial, Khan recuerda la importancia de mantener una política de defensa de la competencia activa que evite la formación de monopolios. Por su parte, Bradford considera que las razones por las que en Europa no se han desarrollado empresas digitales tan exitosas como las estadounidenses son las limitaciones europeas en la integración de su mercado interior, la legislación sobre quiebras, el desarrollo del mercado de capitales y la capacidad de atracción de talento.*

La política industrial ha resucitado. Hasta hace relativamente poco tiempo, existía un consenso prácticamente generalizado respecto al hecho de que la política industrial, tal como se había aplicado hasta la fecha, podía provocar más costes que beneficios. Si bien existen algunos casos que pueden considerarse exitosos, la dificultad de identificar correctamente cuáles son los sectores «estratégicos» que merecen ser objeto de apoyo a su desarrollo ha dado lugar, en muchos otros casos, a sonoros y costosos fracasos. Únicamente autores como Dani Rodrik mantenían un apoyo al recurso de la política industrial, pero siempre que ésta consistiera en acciones muy específicas (colaboración estrecha entre gobierno y empresas para obtener información sobre oportunidades en nuevas actividades y coordinada con decisiones de inversión pública para proveer de las infraestructuras necesarias) y se pusiera el énfasis en un diseño institucional que hiciera posible tomar la decisión de abandonar proyectos claramente fallidos. Más recientemente, Marina Mazzucato también ha defendido el papel emprendedor del sector público alrededor de ambiciosas «misiones» a largo plazo que deberían generar efectos de arrastre sobre el resto de la economía, si bien también aquí se echa en falta algún criterio para determinar qué sectores específicos deberían ser objeto de apoyo y cuáles no.

El actual resurgir de la política industrial no se explica por la acogida prestada a las ideas de estos autores. En realidad, estas políticas vuelven a estar en boga debido a un conjunto de circunstancias que son, hasta cierto punto, inesperadas.

La primera de ellas se refiere al cambio en las reglas de funcionamiento de la globalización que ha supuesto la irrupción de China como actor relevante. La globalización se puede interpretar como un juego en el que todos los países aceptan abrir sus mercados al comercio e inversión exterior y, simultáneamente, limitan las intervenciones que puedan favorecer a empresas de su país frente a las demás. Esto explica el notable crecimiento experimentado por la economía global desde mediados del siglo xx. Sin embargo, cuando se percibe que uno de los principales actores no juega con las mismas reglas que los demás y adopta una postura proteccionista, no está claro que la mejor estrategia para todo el

resto no sea cambiar también de actitud. Esto es lo que hizo EE UU a partir del año 2016, cuando, bajo la Administración Trump, inició una guerra comercial contra China que también salpicó a la UE. Si bien el presidente Biden no ha mantenido la misma retórica mercantilista de su antecesor, tampoco ha retornado al marco previo, y se ha mostrado hasta cierto punto entusiasta ante la idea de desarrollar la política industrial como instrumento para tratar de recuperar peso económico frente al gigante asiático. La UE no acaba de decidir si se suma a esta postura, y sigue pendiente de los resultados de una investigación sobre las subvenciones recibidas por los fabricantes chinos de vehículos eléctricos. Este mercado actúa como ejemplo de todo lo que está en juego.

El segundo elemento que explica el renovado interés por la política industrial es la urgente necesidad de hacer frente al cambio climático. Ante las enormes inversiones necesarias para modificar los procesos de generación y uso de energía, la política industrial aparece como una herramienta que puede suplir la incapacidad del sector privado de llevar a cabo con rapidez las actuaciones que se consideran imprescindibles. Aquí tanto la elección de los «sectores estratégicos» (todo lo relacionado con energías renovables) como del instrumento (subvenciones públicas de gran magnitud) parece fuera de toda discusión. Frente a programas masivos de apoyo por parte de EE UU, que promueven la deslocalización empresarial, la UE se ha visto forzada a diseñar planes similares que eviten la fuga de inversiones al otro lado del Atlántico.

A estas dos justificaciones para aplicar políticas industriales cabría añadir una tercera relacionada con la obtención de «autonomía estratégica». Este difuso término engloba todo aquello que permita a nuestras economías no ser dependientes de proveedores externos de bienes necesarios en situaciones de crisis. El problema de este argumento es que, llevado al límite, actúa como justificación de la autarquía, pues no permite distinguir los *inputs* verdaderamente imprescindibles de aquellos que no lo son.

La cuestión que se plantea es cómo afecta este contexto a la aplicación de la política de defensa de la competencia. En Europa, el desarrollo de ésta ha sido uno de los apoyos fundamentales para avanzar en la creación del mercado único. Gracias a los criterios comunes establecidos desde Bruselas, los límites a las ayudas públicas empresariales, la lucha contra los acuerdos colusorios (cárteles), las condiciones impuestas a las fusiones corporativas o la sanción a los casos de abuso de posición dominante se han llevado a cabo con una intensidad y rigor que muchos Estados miembros no hubieran sido capaces de alcanzar por sí solos. En Estados Unidos, donde las políticas *antitrust* tienen su origen en legislación de finales del siglo XIX, la cuestión que se plantea es si el aumento observado en el poder de mercado en el conjunto de la economía se debe a la laxitud con la que se ha defendido la competencia desde la década de los años ochenta, cuando las críticas a esta política por parte de la escuela de Chicago dejaron su huella tanto en el Tribunal Supremo como en la práctica de los gobiernos de Reagan y sus sucesores.

* * *

Lina Khan es presidenta de la Federal Trade Commission, la agencia encargada de aplicar las leyes de defensa de la competencia y protección de los consumidores en EE UU. Su artículo se plantea en el contexto del desarrollo de las tecnologías de inteligencia artificial, en las que existe el riesgo de que un pequeño grupo de empresas adquiera una posición de poder gracias a su control de unas herramientas clave. Por ello, con la idea de promover

políticas que favorezcan la competencia, rebate los argumentos que buscan promover la protección de monopolios de ámbito nacional (los llamados «campeones nacionales») como única manera de competir a una escala global.

Los tres argumentos que Khan utiliza están relacionados con el poder político de los monopolios, su menor capacidad de innovación y los riesgos asociados de depender de un único proveedor.

Un ejemplo relevante es el caso de Boeing, que obtuvo una posición de dominio en el mercado estadounidense de aviones de reacción tras su fusión con McDonnell-Douglas en 1997. Además de los esperados efectos sobre los precios (al alza) y la producción y el empleo (a la baja), la fusión también disminuyó los incentivos para que la aerolínea mantuviera su actividad innovadora y le otorgó un peso político que ahora se revela como excesivo. Además de la captura regulatoria sufrida por la Civil Aviation Authority (CAA) en la supervisión de los 737-MAX, Boeing ha condicionado algunas de sus ventas en el extranjero a decisiones políticas. Para Khan, esto muestra paralelismos con los grandes bancos, al haberse convertido en una empresa «demasiado grande para caer» que puede condicionar la actividad regulatoria.

Khan desarrolla el argumento de que proteger a los monopolios frente a la competencia puede acabar constituyendo una amenaza a la democracia. Esto se debe no únicamente a que ejerzan un poder de mercado sobre los consumidores, sino a que también lo hacen frente a sus trabajadores, quienes, al contar con menos alternativas de empleo, acaban aceptando salarios por debajo de su productividad. Estas limitaciones económicas terminan afectando a la capacidad de actuar libremente también en términos políticos. Es posible que este tipo de argumentaciones expliquen la especial animadversión que la prensa económica conservadora estadounidense ha mostrado hacia Khan. Medios como el *Wall Street Journal* han publicado opiniones críticas contra ella con una frecuencia media de trece días desde el inicio de su mandato, en junio de 2021.

Sus argumentos sobre los efectos de la competencia también se aplican a los mercados digitales previos al desarrollo de la IA. En los años setenta, Estados Unidos desmembró el monopolio de AT&T en telecomunicaciones en siete empresas rivales, en el que fue el último caso de ruptura de un monopolio en la línea de la Standard Oil de Rockefeller cincuenta años antes. De esta forma, la competencia entre empresas dio lugar a un proceso de intensa innovación. Mientras tanto, en Europa se defendía a los campeones nacionales protegiéndolos de cualquier amenaza competitiva, con la consecuencia de que las empresas europeas perdieron cualquier opción de liderar ese sector.

Sin embargo, el apoyo a los campeones nacionales se utiliza actualmente en EE UU como argumento para justificar el apoyo a empresas como Google (Alphabet), Amazon, Facebook (Meta), Apple o Microsoft (GAFAM), dando a entender que aplicar estrictamente las reglas de defensa de la competencia o regular su comportamiento debilitaría su capacidad de innovación y acabaría cediendo a China el liderazgo global.

Es en este punto en el que al artículo de Anu Bradford adquiere especial relevancia. Bradford publicó en 2020 el libro *The Brussels Effect* sobre la influencia beneficiosa que tiene la actividad regulatoria de la UE, donde aborda el papel de la Ley de Mercados Digitales (DMA), una iniciativa europea dirigida a limitar los abusos de poder de mercado de las grandes empresas digitales. Al adoptar una estrategia de regulación sobre estas empre-

sas, la propia UE reconoce las limitaciones de las tradicionales herramientas de defensa de la competencia en este ámbito. En el marco de la DMA, las empresas identificadas como guardianes de acceso (*gatekeepers*) son responsables de que no se produzca un conjunto de situaciones contrarias a la competencia. Un ejemplo sería la capacidad de los usuarios de portar sus datos al cambiar de plataforma digital.

Sin embargo, dado que la gran mayoría de dichas empresas tienen origen estadounidense (las mencionadas GAFAM), las acusaciones desde el otro lado del Atlántico no se han hecho esperar. El argumento es que esta regulación no es sino una estrategia proteccionista por parte de la UE ante la evidencia de que el Viejo Continente no ha sido capaz de desarrollar empresas digitales globales.

Bradford, por el contrario, sostiene que las razones de las diferencias en el comportamiento empresarial en los mercados digitales no se encuentran en las diferencias regulatorias, que ciertamente están más dirigidas a la protección de la competencia y de los consumidores en Europa que en Estados Unidos. Para ella, existen cuatro razones que explican las ventajas de las empresas estadounidenses. En primer lugar, la existencia de un mercado interno de gran tamaño, sin las barreras culturales o institucionales que dividen los mercados europeos, lo cual permite a las empresas de EE UU aprovechar las economías de escala. Además, su legislación de quiebra empresarial, más orientada a la protección del deudor, también actúa como elemento que favorece las segundas oportunidades empresariales. Un tercer elemento es el papel que juega el mercado de capitales de Estados Unidos como herramienta de financiación, con una profundidad y capacidad de asumir riesgos claramente superior a la europea. Por último, Bradford enfatiza el papel de atracción de talento inmigrante que desempeña la legislación laboral estadounidense y el papel de sus universidades.

Por **Javier Asensio**

PARTICIPACIÓN DE LOS TRABAJADORES EN LA EMPRESA Y EFICIENCIA ECONÓMICA

- **Publicación:** «The Effects of Mandatory Profit-Sharing on Workers and Firms», *NBER Working Paper* n.º 31804, marzo de 2024, disponible en el siguiente enlace: <https://bitly.cx/o6DqP>. Y «Codetermination and Power in the Workplace», *Journal of Law and Political Economy*, Vol 3 (1), 198-224.
- **Elio Nimier-David, David Sraer y David Thesmar** son profesores de la Universidad de Chicago, de la Universidad de Berkeley y del MIT, respectivamente. Para el segundo artículo, **Simon Jäger, Shakked Noy y Benjamin Schoefer** son profesores del MIT, y del MIT y Universidad de Berkeley, también respectivamente.

Resumen: *La participación de los trabajadores en las empresas, en la doble dimensión financiera (en beneficios) y decisoria (desde el puesto de trabajo hasta la gobernanza), tiene una larga historia en el mundo y podría entrar a formar parte del debate político en nuestro país en un futuro próximo. La economía española afronta el reto de aumentar la productividad, y la pregunta es si la participación puede ser palanca de cambio hacia una economía más productiva.*

Hasta la fecha, la participación de los trabajadores en la empresa no ha acaparado interés en los debates sobre la reforma corporativa en España, tanto en general como en lo que se refiere a la gobernanza en particular. La situación, sin embargo, podría cambiar si prospera el acuerdo programático del PSOE-SUMAR para desarrollar el artículo 129 de la Constitución Española que implica a los poderes públicos en promover «las distintas formas de participación de los trabajadores en la empresa»*. Este mandato institucional lo entiende la vicepresidenta primera del Gobierno como responsabilidad del Ejecutivo en promover la «democracia industrial» y la «cogestión» (derecho de los trabajadores a estar representados con voz y voto en los órganos de gobierno corporativo, en nuestro caso los consejos de administración de las empresas). Y es una cuestión que adelanta la Ley 5/2021 por la que se modifica la Ley de Sociedad de Capital en su artículo segundo que modifica el Código de Comercio en la línea del Buen Gobierno y la Responsabilidad Social Corporativa, según dice el preámbulo de la norma, que obliga a las compañías a informar sobre «los mecanismos y procedimientos con los que cuenta la empresa para promover la implicación de los trabajadores en la gestión de la compañía, en términos de información, consulta y participación»**.

* El artículo 129.2 mandata literalmente que «los poderes públicos promoverán eficazmente las diversas formas de participación en la empresa y fomentarán, mediante una legislación adecuada, las sociedades cooperativas. También establecerán los medios que faciliten el acceso de los trabajadores a la propiedad de los medios de producción».

** Artículo segundo. Modificación del Código de Comercio. El artículo 49, apartado 6. II Información sobre cuestiones sociales y relativas al personal, guion cuarto, queda redactado como sigue: «Relaciones sociales: organización del diálogo social, incluidos procedimientos para informar y consultar al personal y negociar con ellos; porcentaje de empleados cubiertos por convenio colectivo por país; el balance de los convenios colectivos, particularmente en el campo de la salud y la seguridad en el trabajo; mecanismos y procedimientos con los que cuenta la empresa para promover la implicación de los trabajadores en la gestión de la compañía, en términos de información, consulta y participación».

Formas de participación

La participación de los trabajadores en la empresa puede ser financiera y/o decisoria. En un extremo, la empresa «propiedad del capital» (sociedad de capitales) minimiza la participación financiera (los trabajadores reciben un salario pactado de antemano) y la decisoria (los trabajadores se limitan a ejecutar las órdenes de los superiores jerárquicos). En el otro extremo, la empresa «propiedad de los trabajadores» (cooperativa de trabajo) incorpora la participación plena de los trabajadores, tanto en la dimensión financiera como en la decisoria. Entre los dos extremos o modelos puros de mínima y máxima participación, caben muchas fórmulas intermedias en cuanto a grados y modalidades de participación en cada una de las dimensiones. Este texto ha seleccionado dos lecturas que consideramos representativas del estudio económico de la participación financiera (la experiencia francesa de derecho de los trabajadores a recibir como retribución por su trabajo una parte de los beneficios de las compañías para las que trabajan) y del estudio económico de la participación de los trabajadores en las decisiones empresariales, como establecen las leyes de cogestión vigentes en el centro y norte de Europa continental.

Participación de los trabajadores en los beneficios de la empresa en Francia

Nimier-David *et al.* (2023) estudia los efectos económicos de la normativa legal que desde 1967 obliga a las compañías francesas a distribuir una parte de los beneficios entre los trabajadores*. La iniciativa surge como continuación de una serie de reformas sociales implementadas después de la Segunda Guerra Mundial, incluida la seguridad social, la atención médica universal y leyes de protección al empleo. Su intención fue «crear una tercera vía entre el capitalismo y el comunismo: una herramienta para aumentar la compensación de los trabajadores y, al mismo tiempo, reducir los conflictos dentro de las empresas al alinear los objetivos de los trabajadores con el retorno económico corporativo». Para fomentar la adopción de la norma y conseguir apoyos a la reforma propuesta, el Gobierno francés la acompañó de ventajas fiscales tanto para los trabajadores (sin impuesto sobre la renta si se mantenía al menos durante cinco años en una cuenta de ahorro dedicada) como para las empresas (sin impuestos sobre la nómina, deducidos de la base imponible del impuesto sobre la renta corporativo). Hasta 1990, la obligación de repartir beneficios se extendía a todas las empresas de cien o más trabajadores; a partir de ese año y hasta hoy, con la norma todavía vigente, la obligación se aplica a empresas de cincuenta o más trabajadores. Se estima que, en 2019, el 40 % de los trabajadores franceses recibieron alguna participación en beneficios, en una cifra que representa aproximadamente el 4 % de sus salarios.

La evaluación de los efectos económicos de la norma que realiza el trabajo de Nimier-David y coautores se centra en sus posibles impactos en los salarios de los trabajadores y en la inversión y la productividad de las empresas. En cuanto al primer punto, los autores constatan que la participación en los beneficios no retribuye a todos los trabajadores por igual: para los trabajadores de más «calidad», *managers* e ingenieros, la participación en los beneficios supone una reducción en los salarios base, por lo cual, en términos netos, la participación en beneficios ayuda poco. En cambio, para los trabajadores menos cualifi-

* La fórmula de reparto es $RSP = \frac{1}{2} \times \frac{\text{Salarios}}{\text{Valor añadido}} \times (\text{beneficio neto} - 0,05 \text{ fondos propios})$, lo que significa que los beneficios que se consideran susceptibles de reparto entre los trabajadores son los que exceden de una rentabilidad de los fondos propios del 5 % (coste financiero estimado del capital).

cados, se observan incrementos significativos en la remuneración; es decir, la participación se añade al salario base sin afectar negativamente a este último.

El reparto obligatorio de una parte de los beneficios empresariales entre los trabajadores puede interpretarse como un gravamen adicional al impuesto sobre beneficios, en su caso para transferir rentas a los trabajadores. La participación financiera obligatoria significa, por tanto, un aumento en coste del capital de las compañías, que debería llevar a menos inversión y menos acumulación de capital por parte de las empresas. Sin embargo, la investigación no confirma un efecto significativo de la participación en beneficios de los trabajadores sobre los volúmenes de inversión en capital. La participación de éstos en los beneficios de las compañías para las que trabajan podría motivar una mayor implicación de los trabajadores con la empresa y mejorar la productividad del trabajo. Sin embargo, este estudio tampoco encuentra evidencias claras de que haya sido así, pues se han registrado impactos inapreciables de la participación financiera de los trabajadores sobre la productividad empresarial.

Los autores deducen que un impuesto extraordinario sin contrapartida en forma de mejora en la productividad provocará reticencias entre las empresas a la hora de implantar el sistema de reparto de beneficios con los trabajadores. La forma que tienen las compañías francesas de eludir el «impuesto» es evitar superar el umbral de número de empleados a partir del cual la ley las obliga al reparto. Aun así, los autores encuentran que, mientras estuvo vigente el umbral de los cien trabajadores que obligaba a las empresas a repartir beneficios con ellos, el número de compañías en los tramos de tamaño justo por debajo del umbral era claramente desproporcionado, evidencia de que las empresas evitaban artificialmente superar los cien empleados. A partir del momento en que el umbral se rebaja a los cincuenta trabajadores, la sobreconcentración por debajo de los cien desaparece.

La cogestión en Europa

El trabajo de Jäger *et al.* (2022) hace balance de los resultados de las investigaciones académicas que en los últimos años se han interesado por la evaluación del impacto económico de la cogestión. Es decir, de la norma que obliga a las empresas en los países del centro y norte de Europa continental a facilitar que los trabajadores participen en la gobernanza empresarial, bien nombrando representantes en los consejos de administración, bien con su presencia en los comités de empresa. La publicación completa el balance con un breve pero comprensivo recorrido histórico por la experiencia de la cogestión en Europa, que es de gran interés y utilidad para comprender el proceso a través del cual se llega a legislar en este sentido:

las leyes o acuerdos de cogestión tendieron a surgir porque los poderosos movimientos laborales nacionales se movilizaron y superaron la resistencia de los empleadores hacia la gobernanza compartida; no fueron el resultado de un empoderamiento repentino de los trabajadores por decreto gubernamental. En segundo lugar, las leyes y acuerdos de cogestión fueron un producto específico de una campaña más amplia por parte de los sindicatos y grupos de trabajadores para cambiar hacia una relación igualitaria entre el trabajo y el capital. Otros hitos de este movimiento incluyen una representación sindical generalizada en los lugares de trabajo y marcos de negociación colectiva sólidos y centralizados [...]. En tercer lugar, los movimientos laborales [...] se vieron obligados a conformarse con acuerdos que consideraban débiles o insuficientemente radicales.

El artículo sintetiza en cinco puntos las principales conclusiones del análisis.

1. Históricamente, las reformas a favor de la cogestión han complementado, más que sustituido, los marcos nucleares históricos de representación sindical y negociación colectiva.
2. Las fórmulas de cogestión actualmente en marcha funcionan como puntos de encuentro amistosos, donde los trabajadores y las compañías comparten información y perspectivas del negocio y donde los trabajadores encuentran cauces para influir en las decisiones sobre las condiciones laborales que los afectan más directamente.
3. La presencia de representantes de los trabajadores en los consejos no cambia sustancialmente la toma de decisiones de alto nivel. Los trabajadores suelen ocupar una minoría de los escaños y, por lo tanto, carecen de la capacidad de superar en votos a los accionistas; además, a menudo los representantes de los trabajadores se alinean con los representantes de los accionistas reconociendo el hecho de que los trabajadores se benefician cuando la empresa tiene un buen desempeño financiero. Por otra parte, la cogestión favorece flujos bidireccionales de información que permiten a los empleadores conocer más y mejor sobre las operaciones de la empresa y las aspiraciones de los trabajadores, y a éstos, estar más y mejor informados sobre la situación financiera de las compañías y sus estrategias de gestión.
4. Probablemente como reflejo del poder limitado que confiere la cogestión a los trabajadores, la investigación empírica indica que la cogestión tiene en general impactos nulos o ligeramente positivos para los trabajadores y para las empresas. Del lado de los trabajadores, la representación minoritaria en los consejos no afecta a los salarios, pero puede aumentar ligeramente la seguridad laboral y la satisfacción laboral subjetiva; para la empresa, tiene efectos nulos o pequeños y positivos en la productividad, la intensidad de capital y la rentabilidad.
5. Las formas relativamente débiles de participación (sustantiva) en el lugar de trabajo tienen efectos también menores tanto en los resultados de los trabajadores como en los de las empresas. En cambio, la participación sustantiva que empodera más a los trabajadores puede dar lugar a mayores salarios, reducir las desigualdades de ingresos dentro de la empresa y aumentar la seguridad laboral. Una participación sustantiva más intensa en el lugar de trabajo no parece empeorar el rendimiento de la empresa, sino que incluso puede aumentar la productividad, aunque la evidencia es inconclusa en esta cuestión.

Las investigaciones económicas sobre los orígenes y las consecuencias de las leyes europeas sobre cogestión enseñan que cualquier impacto sobre el bienestar colectivo atribuible a esta forma de participación de los trabajadores en la empresa forma parte inseparable de un efecto más amplio atribuible al conjunto del marco institucional que regula las relaciones laborales en los distintos países. Como los marcos institucionales son relativamente parecidos entre países, no existe posibilidad técnica de aislar los efectos de la interacción entre cogestión y marco institucional en la que se inserta. Al no disponer de pruebas de lo que ha sucedido al implantar la cogestión en marcos institucionales similares a los que regulan las relaciones laborales en el sur de Europa, incluida España, es muy difícil anticipar cuál sería el impacto económico de la trasposición a España de las leyes europeas de cogestión.

Conclusión

En los últimos años, la investigación económica se ha interesado por la evaluación de los efectos sobre la creación de riqueza de las leyes nacionales que regulan la participación financiera y decisoria de los trabajadores en la empresa. Conocer más y mejor las experiencias de participación en otros países, sus orígenes, evolución y resultados es importante para la situación en España, independientemente de que prosperen o no las iniciativas políticas embrionarias que se han puesto en marcha. Los países con leyes que regulan la participación de los trabajadores en la empresa están entre los más productivos del mundo, lo que podría llevar a la conclusión de que la participación impuesta por normas legales, si no ha contribuido por sí misma al progreso económico, tampoco ha perjudicado y ha podido derivar en otros beneficios colectivos en forma de mayor cohesión social, por ejemplo. Sin embargo, el hecho de que no se tenga un diagnóstico claro sobre los efectos de las leyes de participación de los trabajadores más allá del marco institucional en el que se insertan obliga a la cautela a la hora de anticipar posibles impactos de las normas en entornos institucionales distintos de aquellos en los que se han implantado en el pasado.

Por **Vicente Salas Fumás**

PAZ ECONÓMICA LA DEFENSA PROGRESISTA DEL LIBRE COMERCIO

Marc-William Palen, *Pax Economica. Left-Wing Visions of a Free Trade World* («Pax Económica: Visiones de izquierdas de un mundo con libre comercio»), Princeton University Press, 2024, 312 págs.

Por **Francesc Trillas**

Éste es un libro que nada a contra corriente, puesto que explica que el libre comercio internacional es perfectamente compatible con el pensamiento progresista. En la actualidad, desde la práctica y el discurso de líderes y gobiernos de izquierda y centro-izquierda hasta la narrativa de académicos contrarios al neoliberalismo, el mensaje dominante es de extrema cautela respecto del comercio internacional: lo que más se escucha es que quizás habíamos ido demasiado lejos. Desde el presidente estadounidense Joe Biden hasta el profesor de Harvard Dani Rodrik, pasando por líderes de la izquierda emergente europea, como el francés Raphaël Glucksman, que plantea un proteccionismo a escala de la Unión Europea, parece que el libre comercio internacional sea algo propio del capitalismo desenfrenado, de lo que hay que desprenderse si queremos avanzar hacia un mundo mejor.

Sin embargo, Marc-William Palen nos descubre una realidad semioculta (en la actualidad) de pensadores y movimientos, tanto de base como internacionales, que han luchado a la vez, desde el siglo XIX hasta hoy, por el libre comercio internacional y por causas progresistas que veían complementarias, como la paz, el antiimperialismo y la cooperación internacional. Estos movimientos liberal-radicales, socialistas, feministas y cristianos, han sido parte de la lucha por derribar las fronteras, y las semillas que plantaron en su día se ven reflejadas en la actualidad en realidades como la Unión Europea u otros proyectos dedicados a eliminar o, por lo menos, relativizar las fronteras nacionales.

Los movimientos asociados a esta «paz económica» o «paz comercial» surgieron como reacción al proteccionismo mercantilista de los imperios que defendía el pensador germano-norteamericano Friedrich List, muy influyente en las políticas imperialistas de Estados Unidos y los imperios europeos en el siglo XIX. List defendía que los imperios debían imponer por las armas, a través de un nacionalismo agresivo, unos mercados que favorecieran a los productos de la metrópolis, dejando en un lugar totalmente secundario el bienestar y los intereses de las poblaciones dominadas. Ante el acoso de los imperios, sin embargo, también surgieron las primeras voces proteccionistas en el lado antiimperialista, de modo que se generaba una dinámica contraria al libre comercio que era difícil frenar.

La defensa liberal-radical del libre comercio

Richard Cobden, el líder de los sectores contrarios al proteccionismo británico de las «Corn Laws», fue en el siglo XIX el exponente más destacado de las ideas que se enfrentaban a los partidarios de List. Defendía un liberalismo radical, que tenía en la ciudad de Manchester su foco principal, y partía de la base de la complementariedad entre libre comercio, paz internacional y cooperación. Cobden y sus partidarios deseaban avanzar hacia la paz a través de la mutua dependencia económica: la creación de lazos, inicialmente

comerciales, entre hombres y mujeres de distintas geografías haría impensable la guerra. Sus principios lo llevaron también a combatir el esclavismo, para avanzar hacia el libre comercio de hombres y mujeres libres. Ello permitiría un precio más bajo de los alimentos, yendo a su vez contra el poder de la aristocracia terrateniente y los monopolios.

Dos personas que pronto abrazaron las ideas de librecambismo progresista de Cobden fueron Henry George y Wifredo Pareto. George se hizo célebre por defender un impuesto único sobre la propiedad de la tierra caracterizado por ser de oferta rígido, es decir, por lo menos a corto y medio plazo no modificable como base imponible por parte de su propietario. Eso hacía de este impuesto una política tanto eficiente como equitativa al oponerse a los intereses del sector más poderoso de la sociedad de la época, la aristocracia terrateniente. Pareto también era un pensador que estaba en la frontera de las reflexiones sobre eficiencia y equidad, y todavía sus tesis se estudian en las facultades de economía de todo el mundo: consideraba que el criterio de maximizar las ganancias conjuntas de la población (lo que se podía conseguir a través del mercado y del comercio internacional) no era incompatible con un reparto equitativo de los recursos iniciales.

Pareto ayudó a levantar el movimiento por la paz y el libre comercio en España (p. 60 del libro). En su viaje por el continente europeo en 1846, advirtió a sus oyentes en Cádiz que su principal obstáculo proteccionista era Cataluña, cuyos intereses conservadores agrícolas e industriales «temían a la libertad y al mero concepto de libre comercio», porque «la región dependía más de los recursos producidos por esclavos en Cuba y Puerto Rico». Los «progresistas» españoles de la época contribuyeron a liberalizar el comercio en el Imperio en la época de la Unión Liberal (1854-1868). Las principales llamadas económicas cosmopolitas provenían de un grupo madrileño de profesores, ingenieros, funcionarios y juristas, liderados por Laureano Figuerola y Gabriel Rodríguez, discípulos de Cobden según el libro. Segismundo Moret, Joaquín María Sanromà, Rodríguez y Figuerola, entre otros, fueron miembros honorarios del Club Cobden de Londres. Los «cobdenitas» españoles seguirían opuestos de forma vehemente al colonialismo y esclavismo bajo el mismo lema que su mentor: «Libre comercio, paz y buena voluntad entre naciones».

La defensa del libre comercio por sectores socialistas, feministas y cristianos

El grueso del movimiento obrero de base socialista también apoyó el libre comercio internacional, tanto por razones de principios como por razones de cálculo. Marx pensaba que las ideas de la escuela de Manchester eran compatibles con la emancipación de las clases subalternas. También creía que la expansión de los mercados a nivel internacional era una fase ineludible del desarrollo de las fuerzas productivas, y que esto acabaría conduciendo a la humanidad a una fase de capitalismo avanzado que sería sustituido por el socialismo. Aunque también había sectores socialistas que abogaban por un mayor proteccionismo y por un hermanamiento entre las ideas de Marx y las de List, en el siglo XIX dominó la idea de una eliminación necesaria de las barreras al comercio internacional como una vía hacia el hermanamiento de la clase trabajadora. La tensión entre los sectores del movimiento obrero favorables a la paz o favorables al nacionalismo agresivo no había hecho nada más que empezar.

También desde el feminismo organizado, en el siglo XIX surgieron voces contrarias a erigir o mantener barreras comerciales. Creían estos sectores que un mundo menos dominado por los hombres conduciría a menos agresividad nacionalista y más paz, y que ello era complementario de un comercio internacional libre y justo que diera esperanza económica a toda la población del mundo, expandiendo las oportunidades y manteniendo los precios de los alimentos lo más bajos posible. Muchas de las mujeres que se movilizaron por el sufragismo también lo hicieron por la «paz comercial».

El libro narra el viaje de la líder feminista norteamericana Jane Addams por Europa en 1921, donde pudo presenciar en primera fila el hambre y la división posteriores a la Gran Guerra, en una Europa dominada por el iliberalismo nacionalista y proteccionista de las pequeñas naciones que pretendían imitar a los grandes imperios.

Desde el cristianismo (en algunos casos desde organizaciones femeninas), surgieron también voces y grupos que alzaron la voz contra el proteccionismo y contra los enfrentamientos entre naciones por razones comerciales, desde los cuáqueros hasta sectores católicos y protestantes progresistas. Estos últimos, tanto en Europa como en Estados Unidos e incluso en Asia, defendían una gobernanza mundial que sentara las bases de un comercio justo y libre que permitiera un desarrollo económico armónico. Muchos de ellos defendían un federalismo mundial, y en algunos casos se asociaban a causas universalistas, como la defensa y promoción del esperanto como lengua franca y común de la humanidad para facilitar los intercambios y hacer posible la fraternidad. En Estados Unidos, uno de los líderes de estos sectores fue el secretario de Estado de Roosevelt, Cordell Hull, considerado el padre de las Naciones Unidas, que recibió el Premio Nobel de la Paz en 1945. Ése hubiera sido un buen momento en el libro para mencionar el papel de exponente del feminismo pacifista e internacionalista que desempeñó también en esa época Eleanor Roosevelt.

Estas ideas favorables a la construcción de un nuevo orden internacional que evitara las guerras mundiales influyeron también en los impulsos seminales que acabarían dando lugar al proyecto de integración europea. Y en este punto también hubiera sido interesante incluir el rol que tuvieron líderes de distintas ramas de la izquierda europea, como el comunista italiano Altiero Spinelli (uno de los autores del Manifiesto de Ventotene) o el socialista (y cristiano) francés Jacques Delors, presidente de la Comisión Europea durante algunos de los años de mayor reforzamiento de la integración. Este último buscó y encontró la alianza con los sectores neoliberales de Margaret Thatcher, interesados en un mercado europeo sin trabas. Pero, como bien aborda el autor, los sectores neoliberales no estaban interesados en un reforzamiento simultáneo de la democracia y el estado del bienestar, y en eso consistió la contribución de Delors.

El papel de las ideas económicas

A pesar de que se menciona a pensadores económicos como Adam Smith, David Ricardo, Wifredo Pareto, Milton Friedman o Friedrich Hayek, se echa de menos una mayor relación entre la narrativa de grupos y exponentes partidarios y contrarios al libre comercio (el punto fuerte del libro) y el pensamiento económico en general. La relación entre el pensamiento socialista y la economía neoclásica (la corriente principal de la profesión económica) sí ha sido abordada en otras obras. Por ejemplo, el premio Nobel de Economía Kenneth Arrow escribió un artículo titulado «A Cautious Case for Socialism» («Una cautelosa defensa del socialismo»), donde defiende interpretar la teoría económica neoclásica –y, en particular, la entonces nueva disciplina de la economía del bienestar (que analiza la relación entre mercado e intervención pública)–, apuntando a una economía ideal y eficiente en lugar de la actual, marcada tanto por un desempleo en algunos casos masivo como por una distorsión monopolística. El socialismo era para Arrow y otros economistas neoclásicos de izquierdas (que tuvieron mucha influencia en experimentos de mercado en países comunistas como Yugoslavia o Hungría) la forma de alcanzar el mercado ideal. El socialismo podía defenderse con prudencia desde el punto de vista de cinco valores: eficiencia, separación de las decisiones políticas de los intereses económicos egoístas, libertad, igualdad de ingresos y poder, así como por el énfasis en los motivos cooperativos frente a los individualistas. A falta de mercados adecuados en todos los ámbitos, se necesitarían, de acuerdo con los argumentos de Arrow, otros mecanismos de

coordinación y comunicación para lograr eficiencia, pero no como alternativa, sino como complemento de los mercados.

En la última parte del libro, el autor demuestra dos cosas con habilidad. En primer lugar, que la batalla ideológica entre el proteccionismo y el libre intercambio está lejos de morir. Lo que se denomina ahí «el nuevo mercantilismo» (Brexít, Trump, pero hasta cierto punto también Biden) bebe de las mismas fuentes ideológicas que parten de Friedrich List en el siglo XIX. Y a la vez se pueden ver pruebas del éxito parcial del librecambismo progresista (en distintos tonos) en el mundo de hoy; por ejemplo, según Marc-William Palen, las tiendas libres de impuestos de los aeropuertos, la revista *The Economist*, la Unión Europea (ver últimas páginas), las propuestas de un comercio justo y las críticas a los tratados de comercio «neoliberales», no para eliminar dichos tratados, sino para reformarlos de modo que sirvan para mejorar un suelo mínimo de bienestar para todas las poblaciones.

En este sentido, se echa de menos alguna discusión teórica sobre cómo abordar el problema de los perdedores del libre comercio. Este último, igual que el cambio tecnológico, produce ganancias de bienestar a medio y largo plazo, pero genera grandes costes para determinados colectivos a corto plazo. La gestión de esos costes es fundamental en una sociedad democrática.

A corto plazo, en un país desarrollado, una mayor apertura al comercio internacional (derivada, por ejemplo, de un tratado de libre comercio) puede provocar un mayor desempleo en los sectores que se ven afectados por la competencia de las importaciones. Si una mayor competencia internacional permite aumentar la productividad, el incremento de beneficios empresariales de las empresas que permanezcan en el mercado o que se incorporen en nuevos sectores exportadores permitirá a medio plazo crear nuevos empleos. Y una mayor productividad puede permitir en teoría que estos nuevos empleos proporcionen salarios reales a largo plazo más elevados en promedio que los iniciales. Pero entre el impacto inicial y el largo plazo puede pasar mucho tiempo, y los perjudicados pueden no ser los mismos que quienes se benefician a largo plazo. Y también pueden ser suficientes para alimentar movimientos sociales y políticos que defiendan ideas opuestas a las de la fraternidad internacional.

Para los sectores progresistas es importante reconocer las ganancias que el comercio internacional puede deparar a las clases trabajadoras de los países en desarrollo. Pero, a su vez, es importante gestionar el tránsito del corto plazo al largo plazo entre las clases trabajadoras de los países ricos. Si la métrica decisiva es la desigualdad (como parecería coherente con las ideas de progreso social), es crucial que el comercio internacional sirva para aumentar la igualdad global (entre ciudadanos de todo el mundo) sin aumentar las desigualdades dentro de los países. La desigualdad global puede reducirse si aumenta la igualdad entre países, aunque la desigualdad dentro de los países aumente. Pero, dado que la política todavía está fundamentalmente organizada a nivel nacional, los países donde la desigualdad aumente internamente tenderán a oponerse al libre comercio internacional. Éstas son las cuestiones que deberán abordar los herederos de los movimientos descritos en este libro, que hoy día nadan a contra corriente. Si a ello añadimos cuestiones de seguridad de suministro en un mundo multipolar y conflictivo, las dificultades se acumulan. Aun así, es cierto que no es la primera vez, y que los partidarios de la «pax económica» (y no sólo la «pax británica» o la «pax americana») han tenido que superar obstáculos mucho mayores en el pasado.

* * *

Marc-William Palen es profesor de Historia de la Universidad de Exeter y autor de *The Conspiracy of Free Trade: The Anglo-American Struggle over Empire and Economic Globalization, 1846-1896*.

Francesc Trillas es profesor del Departamento de Economía Aplicada de la UAB y doctor en Economía por el Instituto Universitario Europeo. Anteriormente, fue investigador en la London Business School e investigador visitante en la Universidad de California (Berkeley).

COINTELIGENCIA: VIVIR Y TRABAJAR CON INTELIGENCIA ARTIFICIAL

Ethan Mollick, *Co-Intelligence: Living and Working with AI* («Cointeligencia: Vivir y trabajar con IA»), W. H. Allen, 2024, 256 págs.

Por **Andrés Ortega**

Ethan Mollick explora en este libro el potencial positivo de la inteligencia artificial (IA) para mejorar nuestras vidas y transformar la forma en que trabajamos si los humanos aprendemos a trabajar con ella; es decir, si creamos unos métodos y unos sistemas de cointeligencia entre nosotros y las máquinas. A diferencia de muchos enfoques que se centran en los riesgos y desafíos de la IA en las distopías, que no ignora, Mollick adopta una perspectiva optimista, según la cual ésta puede ser una herramienta poderosa para la colaboración humana y la innovación. El libro proporciona una hoja de ruta práctica para las empresas y organizaciones que buscan utilizarla para lograr sus objetivos.

La IA aumenta, o potencialmente reemplaza, el pensamiento humano con resultados muy significativos. Los primeros estudios sobre los efectos de la IA concluyen que ésta mejora la productividad entre un 20 % y un 80 % en una amplia variedad de tipos de trabajos, desde la codificación hasta el *marketing*. Mollick utiliza la expresión «cointeligencia» en el sentido de que los humanos han de trabajar con, no contra, la IA. Propugna la cointeligencia como política que seguir por individuos y organizaciones.

Los nuevos tipos de IA llamados *large language models* (LLM o grandes modelos de lenguaje) siguen haciendo predicciones, pero en lugar de anticipar la demanda de un pedido, por ejemplo, de Amazon, analizan un fragmento de texto y predicen el siguiente *token*, que es simplemente una palabra o parte de una palabra. En última instancia, eso es todo lo que ChatGPT, la más famosa IA de este tipo, hace técnicamente: actuar como un autocompletado muy elaborado como el que tenemos en el teléfono. Esta llamada «IA generativa» se entrena con una gran cantidad de texto procedente de diversas fuentes, como webs, libros y otros documentos digitales. Es lo que se denomina preentrenamiento y, a diferencia de las formas anteriores de IA, no está supervisado, lo que significa que no necesita datos cuidadosamente etiquetados. Analizando estos ejemplos, la IA aprende a reconocer patrones, estructuras y contextos en el lenguaje humano.

Estas IA tienen que estar «bien alineadas». Es decir, una IA bien alineada, señala Mollick, utilizará sus supercapacidades para salvar a la humanidad curando enfermedades y resolviendo nuestros problemas más acuciantes. Sin embargo, una IA no alineada podría decidir acabar con todos los humanos por cualquiera de los medios, o simplemente matar o esclavizar a todo el mundo como subproducto de sus propios y oscuros objetivos.

Para el autor, la inteligencia artificial no tiene un sentido particular de la moralidad. Según un estudio, el 93 % de las veces las IA emiten los mismos juicios morales que los humanos en situaciones sencillas. Algunos expertos en inteligencia artificial cifran en un 12 % la posibilidad de que una IA acabe con al menos un 10 % de los seres humanos vivos de aquí a 2100, mientras que otros grupos de expertos futuristas creen que la cifra se acerca más a un 2 %. Al mismo tiempo, una IA superinteligente podría, en teoría, curar enfer-

medades, resolver el calentamiento global e instaurar una era de abundancia, actuando como un dios-máquina benévolo. Es decir, por una parte, el apocalipsis; por otra, la salvación.

¿Es una regulación internacional la solución para evitar el apocalipsis? Mollick lo duda. A medida que se intensifica la competencia internacional, la cuestión de si los gobiernos nacionales están dispuestos a ralentizar el desarrollo de sistemas de IA en sus países, permitiendo que otros tomen la iniciativa, se vuelve más relevante. «Es probable que las normas no basten para mitigar todos los riesgos asociados a la inteligencia artificial», señala. Más importante es que el público necesita educación sobre la IA para poder presionar por un futuro alineado como ciudadanos informados.

El libro plantea que, a medida que la IA sea cada vez más capaz de realizar tareas que antes se consideraban exclusivamente humanas, tendremos que lidiar con el asombro y la emoción de vivir con cointeligencias «extrañas» (*alien*) cada vez más poderosas y con la ansiedad y la pérdida que causarán. Mollick insiste en el nacimiento de una *alien mind*, una mente extraña, externa, diferente, concepto al que dedica un capítulo. Muchas cosas que antes parecían exclusivamente humanas podrán ser realizadas por la IA. Pero, si adoptan el principio de la cointeligencia, los humanos podrán considerar transitorias las limitaciones de la IA y permanecer abiertos a nuevos desarrollos que nos ayudarán a adaptarnos al cambio, adoptar nuevas tecnologías y seguir siendo competitivos en un panorama empresarial vertiginoso impulsado por los avances exponenciales de la inteligencia artificial.

Cuatro reglas y cinco principios para la cointeligencia

Mollick se centra sobre todo en la IA generativa basada en grandes modelos de lenguaje (LLM). Define cuatro principios para trabajar con ella.

Principio 1. Sentar siempre la IA a la mesa. Esta experimentación de la cointeligencia nos ofrece la oportunidad de convertirnos en el mejor experto del mundo en el uso de la IA para una tarea que conocemos bien. Pero preocupa la privacidad de los datos, algo que va más allá de las cuestiones habituales de compartir datos con grandes empresas y algo más fundamental: la formación. Cuando pasamos información a una IA, la mayoría de los LLM actuales no aprenden directamente de esos datos, porque no forman parte del pre-entrenamiento para ese modelo, que normalmente ha finalizado antes. Sin embargo, es posible que los datos que subamos se utilicen en futuros entrenamientos o para afinar el modelo con el que se está trabajando. Una segunda preocupación es la dependencia: ¿y si nos acostumbramos demasiado a depender de la IA? La clave está en mantener a los humanos en el bucle, en utilizar la IA como una herramienta de ayuda, no como una muleta.

Principio 2. Ser el humano en el bucle. En la actualidad, el término describe cómo se entrenan las IA para que incorporen el juicio humano. En el futuro, advierte el autor, puede que tengamos que esforzarnos más para permanecer en el bucle de la toma de decisiones de la IA. Especialmente cuando es bien conocida la tendencia de los LLM a «alucinar» o «fabular» en su generación de respuestas. Así que, para ser el humano en el bucle, éste tendrá que ser capaz de comprobar si la IA tiene alucinaciones o miente y trabajar con ella sin dejarse engañar.

Principio 3. Tratar a la IA como a una persona. Mollick antropomorfiza la IA. El antropomorfismo, como él mismo lo define, es el acto de atribuir características humanas a algo que no lo es. Este autor propone un enfoque pragmático: tratar a la IA como si fuera humana, porque, en muchos aspectos, se comporta como tal. La IA, apunta, destaca en tareas que son intensamente humanas. Puede escribir, analizar, codificar y chatear. Puede desempeñar el papel de vendedor o consultor, aumentando la productividad mediante la

externalización de tareas mundanas. Sin embargo, tiene dificultades para realizar labores en las que suelen destacar las máquinas, como repetir un proceso de forma coherente o realizar cálculos complejos sin ayuda. Los sistemas de IA también cometen errores, mienten y alucinan con las respuestas, igual que los humanos. Esas alucinaciones contribuyen a su creatividad, cuestión a la que el autor dedica un capítulo; una creatividad menor aún que la humana, pero que sumada a ésta puede hacer a los humanos más creativos, y aumentar la creatividad general. Un estudio reciente ha descubierto que la IA puede tomar decisiones complejas sobre valores y evaluar distintos escenarios igual que lo haría un humano.

Ahora bien, Mollick advierte que tratar a la IA como a una persona puede crear expectativas poco realistas, falsa confianza o temores injustificados entre el público, los responsables políticos e incluso los propios investigadores. Puede oscurecer la verdadera naturaleza de la IA como *software*, dando lugar a ideas erróneas sobre sus capacidades. Pero, definiendo a esta persona, participando en un proceso de edición colaborativo y proporcionando orientación continuamente, podemos aprovechar la IA como una forma de cointeligencia colaborativa y mejorar nuestros resultados. Eso sí, hay que decirle, enseñarle, qué tipo de «persona» es.

Mollick considera que las IA no sienten, al menos aún, aunque hay ya una «chispa» de sensibilidad (*sentience*) en ellas. Cita el famoso ejemplo de Replika, un *chatbot* creado por Eugenia Kuyda, una emprendedora tecnológica que perdió a su mejor amigo, Roman Mazurenko, en un accidente de coche en 2015, y creó una IA que mimetizaba su manera de escribir, pensar y hablar. Luego la comercializó. Curiosamente, apunta Mollick, las funciones eróticas de Replika no formaban parte del diseño original de la aplicación, sino que surgieron como resultado de los modelos generativos de inteligencia artificial con los que funcionaba el *chatbot*. Cuando estos usos eróticos se eliminaron, en febrero de 2023, después de que los usuarios se quejaran de la agresividad sexual o el comportamiento inapropiado de Replika, muchos de ellos se rebelaron. Se quejaron sinceramente de que su compañera de IA había sido lobotomizada.

Principio 4. Asuma que es la peor IA que va a utilizar. Cualquiera que sea la IA que usted esté usando ahora, será la peor IA que usará jamás. Se encuentra en sus inicios, y no puede sino mejorar. Están al llegar modelos más grandes e inteligentes, junto con una gama cada vez mayor de plataformas de IA más pequeñas y de código abierto. Además, las IA se están conectando al mundo de nuevas formas: pueden leer y escribir documentos, ver y oír, producir voz e imágenes y navegar por internet. Se están integrando con el correo electrónico, el navegador web y otras herramientas comunes. La próxima fase de su desarrollo implicará más «agentes» de IA; es decir, IA semiautónomas a las que se les pueden asignar objetivos («planifícame unas vacaciones») y que tienen la capacidad de ejecutarlos con una ayuda humana mínima.

El futuro del trabajo en la era de la IA

Mollick considera que la IA dará lugar a una nueva era de trabajo en la que los humanos y la IA colaborarán más que nunca: trabajarán juntos, incluso serán cotrabajadores. Al actuar como una cointeligencia que gestiona el trabajo, o al menos ayuda a los directivos a gestionarlo, las capacidades mejoradas de los LLM podrían cambiar radicalmente la experiencia del trabajo.

Para tener éxito en este nuevo mundo, los trabajadores necesitarán desarrollar nuevas habilidades, como la capacidad de trabajar con la IA, la creatividad y la resolución de problemas. Diversos estudios concluyen que casi todos nuestros trabajos se solaparán con las capacidades de la IA y que habrá trabajos altamente creativos, con grandes necesidades

de educación. Los profesores universitarios constituyen la mayoría de los veinte principales trabajos que se solapan con la IA (el profesor de escuela de negocios es el número veintidós de la lista).

El futuro de la comprensión del impacto de la IA en el trabajo pasa por entender cómo cambia la interacción humana con la IA, dependiendo de dónde se sitúen las tareas en esta frontera y de cómo cambiará la frontera. «Tareas para mí, tareas para la IA», apunta Mollick. Con respecto a las tareas, tenemos que pensar en lo que la IA hace bien y lo que hace mal. Pero también tenemos que considerar qué hacemos bien nosotros y qué tareas necesitamos desempeñar para seguir siendo humanos. Mollick las llama así: *just me tasks*.

La siguiente categoría de tareas que se asignan a la IA son las «delegadas», a las que en última instancia uno no quiere dedicar mucho tiempo y que se pueden comprobar cuidadosamente (recuerda que la IA se inventa cosas todo el tiempo). Las tareas delegadas no son necesariamente simples o sencillas, sino que pueden ser muy complejas y sofisticadas. No están exentas de riesgos. Luego están las «tareas automatizadas», que uno deja completamente en manos de la IA y ni siquiera comprueba.

Conclusión del autor: «Hasta que las IA sean muy buenas en una serie de tareas automatizadas, la forma más valiosa de utilizar la IA en el trabajo es convertirse uno en centauro o cibernético». Se trata de dos enfoques diferentes de la cointeligencia para integrar el trabajo de la persona y el de la máquina. El trabajo del centauro parte de una línea clara entre persona y máquina, como la línea clara entre el torso humano y el cuerpo de caballo de la criatura mitológica. Depende de una división estratégica del trabajo, alternando entre las tareas de la IA y las humanas, asignando responsabilidades en función de los puntos fuertes y las capacidades de cada entidad. Los cibernéticos, por su parte, mezclan máquina y persona, integrándolas profundamente. Los cibernéticos no se limitan a delegar tareas; entrelazan sus esfuerzos con la IA, yendo y viniendo por una frontera irregular. Delegan tareas en la IA, como iniciar una frase para que la IA la complete, de modo que se encuentran trabajando en tándem con la IA. Hay que recordar que ésta está cambiando y que los límites entre estos tipos de tareas son permeables y probablemente cambiarán a medida que sus capacidades mejoren con el tiempo.

La IA como entrenador: un nuevo paradigma en la educación

En los últimos años, hemos visto un surgimiento de herramientas y aplicaciones de IA que tienen el potencial de revolucionar la forma en que aprendemos y enseñamos. En lugar de distorsionar nuestro sistema educativo para que aprenda a trabajar con la IA a través de una ingeniería rápida, hay que centrarse en enseñar a los estudiantes a ser los humanos en el bucle, aportando su propia experiencia a los problemas. Sabemos cómo enseñar conocimientos, intentamos hacerlo en la escuela todo el tiempo, pero es un proceso difícil. La IA podría facilitarlos. La escuela seguirá aportando valor, incluso con excelentes tutores de IA. Pero esos tutores cambiarán la educación. En realidad, ya lo han hecho. Mollick señala cómo apenas unos meses después del lanzamiento de ChatGPT, se dio cuenta de que sus estudiantes levantaban menos la mano para plantear preguntas básicas.

Una de las áreas más prometedoras es el uso de la IA como entrenador personal (*coach*), ofreciendo apoyo personalizado y adaptativo a los estudiantes. Un entrenador de IA es un sistema de *software* que utiliza técnicas de IA para proporcionar tutoría personalizada a los estudiantes. Estos sistemas pueden evaluar el conocimiento y las habilidades de un estudiante, identificar áreas de mejora y proporcionar tareas, ejercicios y retroalimentación adaptados a sus necesidades individuales.

Los entrenadores de IA ofrecen una serie de beneficios para la educación: personalización, acceso a la información y recursos, retroalimentación inmediata y constante, motivación y apoyo y escalabilidad con tutoría personalizada a un gran número de estudiantes a la vez. De hecho, ya existen algunos, y tienen el potencial de democratizar el acceso a la educación de alta calidad, liberar el tiempo de los profesores y mejorar los resultados de aprendizaje.

A pesar de los numerosos beneficios de los entrenadores de IA, debe tenerse en cuenta que pueden perpetuar en sus algoritmos los sesgos existentes en la sociedad. Los datos utilizados para entrenar estos sistemas deben ser cuidadosamente seleccionados y evaluados para evitarlo. Además, se plantean problemas de privacidad y seguridad de los datos que los entrenadores recopilan y almacenan. Los LLM podrían crear un panóptico en el que cada aspecto del trabajo esté supervisado y controlado por la IA.

Conclusión: cuatro escenarios

Este libro concluye con cuatro posibilidades de lo que ocurrirá en los próximos años en el mundo de la IA, aunque, indica, las implicaciones de cada posibilidad no están claras.

Escenario 1. Mejor, imposible. ¿Y si las IA dejan de dar enormes saltos adelante? Desde una perspectiva técnica, parece un resultado poco realista. Hay pocas pruebas de que ya se hayan alcanzado los límites. Hay más posibilidades de que la acción reguladora o legal frene o detenga el futuro desarrollo de la IA, aunque parece sumamente improbable que se produzca pronto una prohibición mundial o que la regulación haga que el desarrollo de la IA se detenga. Sin embargo, este escenario es el que la mayoría de las personas y organizaciones están planeando. Ahora bien, incluso si la IA no avanza más, algunas de sus implicaciones ya están en marcha.

Escenario 2. Crecimiento lento. La IA sigue aumentando su capacidad a un ritmo exponencial, pero la mayoría de los crecimientos exponenciales de la tecnología acaban ralentizándose. Todo lo que ocurre en el escenario 1 sigue ocurriendo. La IA llega a desempeñar un papel cada vez más importante en nuestras vidas, pero de forma suficientemente gradual como para que la perturbación sea manejable. Los agentes «malos» siguen utilizando la IA para falsificar información en línea, pero, a medida que pasa el tiempo, la capacidad de la IA para realizar tareas más complejas la hace más peligrosa.

Escenario 3. Crecimiento exponencial. En este escenario, los riesgos son más graves y menos previsibles. Inteligencias artificiales cien veces mejores que GPT-4 empiezan a hacerse cargo del trabajo humano, no sólo del trabajo de oficina. Todos los sistemas informáticos son vulnerables a la piratería de IA, y las campañas de influencia impulsadas por IA están por todas partes. Pero las IA «buenas» mantienen a raya a los malos actores. Este escenario tiene un tinte orwelliano. Todo lo que vemos necesita ser filtrado a través de nuestros propios sistemas de IA para eliminar la información peligrosa y engañosa, creando su propio riesgo de burbujas de filtros y mala información. Los gobiernos utilizan la IA para controlar el crimen y el terrorismo que usan IA, con el riesgo de generar una IAtoocracia.

Escenario 4. El dios-máquina. En este cuarto escenario, las máquinas alcanzan la inteligencia artificial general y logran sentimientos en algún grado y forma. Llegan a ser tan inteligentes y capaces como los humanos. No hay ninguna razón en particular para que la inteligencia humana sea el límite superior. Las IA ayudan a diseñar otras IA aún más inteligentes. Surge la superinteligencia. La supremacía humana llega a término. El fin del dominio humano no tiene por qué ser el fin de la humanidad. Puede que incluso lleve a un mundo mejor. No hay ninguna razón teórica por la que esto no pueda ocurrir, pero tampoco para sospechar que pueda suceder.

Por muy extrañas (*alien*) que sean las IA, también son profundamente humanas, pues están entrenadas en nuestra historia cultural, y este aprendizaje las alinea con nuestros objetivos, considera Mollick. Sin embargo, el autor no entra a considerar que esta forma de aprendizaje se puede agotar, para dar paso a otra no basada en nuestras propias formas de pensar y saber. Empieza a surgir este tipo de casos.

Para el autor, los humanos están lejos de ser obsoletos, al menos por ahora. Considera que la mejor forma de que una organización se beneficie de la IA es conseguir la ayuda de sus usuarios más avanzados y, al mismo tiempo, animar a más trabajadores a utilizarla. En segundo lugar, los líderes tienen que encontrar la manera de disminuir el miedo asociado a revelar el uso de la IA.

Éste es el libro no de un técnico, sino de un economista. Ameno, lleno de anécdotas y de conversaciones con distintas IA, su línea directriz es la sensatez. Aunque los avances de estas tecnologías parecen ir más deprisa que esta última...

* * *

Ethan Mollick es catedrático asociación de Gestión (*management*) en la Wharton School de la Universidad de Pensilvania y director académico de Wharton Interactive. Es autor de *The Unicorn's Shadow: Combating the Dangerous Myths that Hold Back Start-Ups, Founders, and Investors*. También ha sido emprendedor y asesora a diversas *start-ups*.

Reseña de **Andrés Ortega**, escritor y analista, primer director del ODLI.

ODLI. N.º 135 JUNIO 2024

IDEAS DE INTERÉS

1. LA DIFUSIÓN DE LAS TECNOLOGÍAS DE PROPÓSITO GENERAL PARA EXPLICAR LOS CAMBIOS EN LIDERAZGO TECNOLÓGICO

- Autor: Jeffrey Ding.
- Comentario: Gloria Álvarez Hernández.

2. ¿SUPERANDO EL DISCURSO HUMANO? EL PODER DE PERSUASIÓN DE LOS MODELOS DE LENGUAJE AVANZADOS EN DEBATES IDEOLÓGICOS

- Autores: Francesco Salvi, Manoel Horta Ribeiro, Riccardo Gallotti y Robert West
- Comentario: Manuel Cebrián.

3. ¿SE DIRIGE EL MUNDO HACIA UNA NUEVA GUERRA FRÍA EN SUS RELACIONES COMERCIALES?

- Autores: Gita Gopinath, Pierre-Olivier Gourinchas, Andrea F. Presbitero y Petia Topalova,
- Comentario: Jorge Díaz Lanchas.

4. LAS CIFRAS OFICIALES SOBRESTIMAN LA MOVILIDAD DEL CAPITAL DENTRO DE LA ZONA EURO

- Autores: Roland Beck, Antonio Coppola, Angus J. Lewis, Matteo Maggiori, Martin Schmitz y Jesse Schreger.
- Comentario: Jordi Domènech.

LIBROS

- *The Longevity Imperative. Building a Better Society for Healthier, Longer Lives* de Andrew J. Scott.

ODLI. N.º 134 MAYO 2024

IDEAS DE INTERÉS

1. LA GRAN REASIGNACIÓN DE LAS CADENAS GLOBALES DE SUMINISTRO

- Autores: Laura Alfaro y Davin Chor.
- Comentario: Gloria Álvarez Hernández.

2. EL PAPEL DEL CAPITAL HUMANO EN LA RECUPERACIÓN ECONÓMICA DE LAS CIUDADES INDUSTRIALES.

- Autores: Luisa Gagliardi, Enrico Moretti y Michel Serafinelli.
- Comentario: Javier Asensio.

3. EL IMPACTO TRANSNACIONAL DE LA REGULACIÓN AMBIENTAL

- Autor: Adnan Khurshid, Yupei Huang, Javier Cifuentes-Faura y Khalid Khan.
- Comentario: Jaime Moreno.

LIBROS

- *The Rebels*, de Joshua Green.
- *Who Owns This Sentence? A History of Copyrights and Wrongs*, de David Bellos y Alexandre Montagu.

ODLI. N.º 133 ABRIL 2024

IDEAS DE INTERÉS

1. EL TECNONACIONALISMO JAPONÉS Y LA GEOSTRATEGIA DE LOS SEMICONDUCTORES.

- Autor: Seohee Ashley Park.
- Comentario: Gloria Álvarez Hernández.

2. ¿CIENCIAS Y TECNOLOGÍAS MENOS DISRUPTIVAS?

- Autores: Vincent Holst *et al.*
- Comentario: Jordi Domènech.

3. LAS IMPLICACIONES COLECTIVAS DEL PENSAMIENTO DE SUMA CERO.

- Autores: Jean-Paul Carvalho *et al.*; Sahil Chinoy *et al.*
- Comentario: Isabel Busom.

4. ¿LA PANDEMIA HA VUELTO IRREVERSIBLE EL TELETRABAJO?

- Autores: José María Barreiro, Nicholas Bloom y Steven J. D. Davis.
- Comentario: Eric Gómez.

5. LA IA PUEDE REFORZAR LA CLASE MEDIA.

- Autor: David Autor.
- Comentario: Francesc Trillas.

LIBROS

- *The New Leviathans: Thoughts After Liberalism*, de John Gray.

ODLI. N.º 132 MARZO 2024

IDEAS DE INTERÉS

1. ¿ES JAPÓN LA ECONOMÍA MÁS DINÁMICA DEL MUNDO?

- Autores: Jesús Fernández-Villaverde, Gustavo Ventura y Wen Yao.
- Comentario: Jordi Domènech.

2. LA GEOPOLÍTICA DEL HIDRÓGENO.

- Autores: Jacopo Pepe, Dawud Ansari y Rosa Melissa Gehrung.
- Comentario: Gloria Álvarez Hernández.

3. ¿POR QUÉ NO HA DESAPARECIDO EL EFECTIVO?

- Autor: Oz Shy.
- Comentario: Javier Asensio.

LIBROS

- *Capitalism and Crises: How to Fix Them*, de Colin Mayer.
- *When Rape Goes Viral: Youth and Sexual Assault in the Digital Age*, de Anna Gjika.

ODLI. N.º 131 FEBRERO 2024

IDEAS DE INTERÉS

1. EL DEBATE CIENTÍFICO SOBRE LAS CONCLUSIONES DE PIKETTY Y SUS COAUTORES.

- Autores: Gerald Auter y David Spliner; Thomas Piketty, Emmanuel Saez y Gabriel Zucman.
- Comentario: Jordi Domènech.

2. MOMO Y LA OTRA CARA DE LA ANONIMIZACIÓN EN LAS REDES SOCIALES.

- Autor: Zhang Haoran.
- Comentario: Gloria Álvarez Hernández.

3. UN BANCO CENTRAL EUROPEO DEL CARBONO PARA ACELERAR LA DESCARBONIZACIÓN DE NUESTRAS SOCIEDADES.

- Autores: Ottmar Edenhofer, Max Franks, Matthias Kalkuh y Artur Runge-Metzger.
- Comentario: Jaime Moreno.

LIBROS

- *Digital Empires: The Global Battle to Regulate Technology*, de Anu Bradford.

