


## La ventana de difusión en la inteligencia artificial: dinámica global y perspectivas para Europa<sup>1</sup>

**Federico Pablo-Martí**


Complex Systems in Social Sciences Group (SCCS), University of Alcalá ✉ 

**Jagoda Kaszowska-Mojša**

Institute of Economic Sciences, Polish Academy of Sciences ✉ 

INET Oxford, University of Oxford

**Juan Luis Santos**

Complex Systems in Social Sciences Group (SCCS), University of Alcalá ✉ 

Institute of Economic Sciences, Polish Academy of Sciences

<https://dx.doi.org/10.5209/pade.105097>

### EN The diffusion window in artificial intelligence: global dynamics and prospects for Europe

**Resumen:** En los últimos años, la “ventana de difusión” entre el lanzamiento de modelos avanzados de inteligencia artificial privados y el acceso público a alternativas abiertas se ha acortado notablemente. Ejemplos como GPT-3 y BLOOM o, más recientemente, LLaMA-4 y GPT-4.1, ilustran esta aceleración impulsada por la reducción de costes computacionales, el auge de comunidades abiertas y decisiones estratégicas empresariales. Europa lidera en modelos abiertos, aunque depende de hardware avanzado. El AI Act europeo puede favorecer transparencia y seguridad, permitiendo a la región consolidar un modelo de autonomía estratégica abierta y definir su futuro digital.

**Palabras clave:** Inteligencia artificial; Ventana de difusión; Modelos abiertos; Europa; Gobernanza tecnológica; Cómputo y hardware; AI Act; Autonomía estratégica abierta

**Abstract:** In recent years, the “diffusion window” between the launch of advanced private artificial intelligence models and public access to open alternatives has noticeably shortened. Examples such as GPT-3 and BLOOM, or more recently LLaMA-4 and GPT-4.1, illustrate this acceleration, driven by lower computational costs, the rise of open communities, and strategic business decisions. Europe leads in open models, although it remains dependent on advanced hardware. The European AI Act could promote transparency and security, allowing the region to consolidate a model of open strategic autonomy and define its digital future.

**Keywords:** Artificial intelligence; Diffusion window; Open models; Europe; Technological governance; Computing and hardware; AI Act; Open strategic autonomy

**JEL:** O31, O32 y O33.

**Sumario:** 1. Introducción. 2. Breve historia de la IA y el escalado computacional. 3. La ventana de difusión: casos descriptivos. 4. Tendencias recientes en costes y hardware. 5. Europa en la carrera global. 6. Implicaciones para la política europea. 7. Perspectivas de futuro. 8. Conclusión. Referencias.

**Cómo citar:** Pablo-Martí, F.; Kaszowska-Mojša, J. y Santos, J. L. (2025). La ventana de difusión en la inteligencia artificial: dinámica global y perspectivas para Europa en *Papeles de Europa* 38(2025), e105097. <https://dx.doi.org/10.5209/pade.105097>

<sup>1</sup> Esta investigación ha sido parcialmente financiada por el proyecto de Jagoda Kaszowska-Mojša NAWA, grant nº BPN/BEK/2024/1/00240 y por el proyecto de Federico Pablo-Martí, PHS-2024/PH-HUM-530 de la Comunidad de Madrid

## 1. Introducción

En noviembre de 2022, millones de europeos descubrieron de golpe la inteligencia artificial generativa a través de un producto gratuito, accesible desde el navegador y bautizado con un nombre breve y llamativo: ChatGPT. El impacto social fue inmediato. En cuestión de semanas, universidades, empresas, parlamentos y medios de comunicación comenzaron a debatir sobre las consecuencias de una tecnología que parecía surgir de la nada, pero que en realidad era la culminación de más de una década de avances en el entrenamiento de modelos de lenguaje a gran escala.

Ese momento de sorpresa social revela una dinámica que rara vez se discute en profundidad: la diferencia temporal entre la frontera tecnológica y su difusión abierta. Mientras ChatGPT irrumpía en la vida cotidiana en 2022, el modelo subyacente, GPT-3, había sido entrenado ya en 2020. El retraso no fue tanto fruto de la casualidad como de una característica estructural del desarrollo en inteligencia artificial: la existencia de un intervalo, que podríamos llamar “ventana de difusión”, entre el lanzamiento de un modelo propietario de frontera y la aparición de alternativas abiertas o de libre acceso que democratizan ese conocimiento. La ‘ventana de difusión’ que observamos en IA es un caso específico de los patrones generales de difusión tecnológica: grandes rezagos, fuerte heterogeneidad entre países y sectores y convergencia incompleta tiempo después de la invención (Comin & Hobijn, 2010).

Comprender esa ventana de difusión es esencial para Europa. La razón es doble. En primer lugar, porque condiciona la velocidad con la que universidades, centros de investigación y empresas emergentes pueden acceder a la vanguardia de la tecnología. Una ventana demasiado amplia genera dependencia tecnológica y rezago competitivo; una ventana demasiado corta abre oportunidades de innovación, pero también riesgos de proliferación incontrolada. En segundo lugar, porque esta dinámica está directamente ligada a la posición geopolítica y económica de Europa en un campo dominado por Estados Unidos y, cada vez más, por China.

La historia reciente ofrece ejemplos claros. GPT-3, presentado por OpenAI en mayo de 2020, supuso un salto de escala sin precedentes, con 175.000 millones de parámetros y un coste de entrenamiento estimado en decenas de millones de dólares. Durante más de dos años, ningún modelo abierto se acercó a ese nivel de potencia. Hubo que esperar hasta julio de 2022 para que el consorcio internacional BigScience, impulsado desde Europa, presentara BLOOM, un modelo multilingüe entrenado en la infraestructura pública del supercomputador francés Jean Zay (BigScience Workshop, 2022). Esa diferencia temporal de más de veinticinco meses constituye un caso paradigmático de ventana de difusión prolongada. Sin embargo, menos de un año después, la situación cambió drásticamente: GPT-4 fue anunciado en marzo de 2023, y apenas unos meses más tarde Meta lanzó LLaMA-2, un modelo abierto con un rendimiento comparable en muchas tareas (Touvron et al., 2023). El intervalo se había reducido de años a meses. En 2025 se observa una inversión puntual de la ventana: Meta lanzó LLaMA-4 el 5 de abril como modelo de pesos abiertos, descargable, reutilizable y modificable con restricciones para usos comerciales masivos habilitando un despliegue directo en hardware propio o nube; y como respuesta nueve días después OpenAI presentó GPT-4.1. Este episodio confirma que la difusión abierta puede no solo acortar la ventana, sino adelantarse a la oferta propietaria en cuestión de días.

Esta aceleración no es un detalle técnico: tiene profundas implicaciones para la capacidad de Europa de posicionarse estratégicamente en la economía digital. En un escenario donde la frontera se difunde casi en tiempo real, las ventajas tradicionales de las grandes corporaciones se ven erosionadas, y cobran fuerza los actores que pueden movilizar comunidades abiertas, coordinar infraestructuras públicas y establecer marcos regulatorios que favorezcan la innovación responsable (Shetty et al., 2025).

Por ello, este artículo propone una reflexión amplia y descriptiva sobre la ventana de difusión en inteligencia artificial y sus consecuencias para Europa. Partiendo de la evidencia empírica disponible (series de cómputo, cronologías de modelos y costes de entrenamiento) se trazará una narrativa que explique cómo ha evolucionado este intervalo, cuáles son sus determinantes tecnológicos y políticos, y qué posición ocupa Europa frente a Estados Unidos y China. A lo largo del texto se destacarán los hitos más relevantes, desde el lanzamiento de GPT-3 hasta la aparición de BLOOM, LLaMA y Mistral, y se pondrá especial atención en las iniciativas europeas que, pese a contar con menos recursos privados, han conseguido situarse en el centro del debate internacional gracias a su apuesta por la apertura y la cooperación.

El objetivo no es ofrecer un análisis econométrico exhaustivo sino presentar de manera clara y rigurosa los patrones observables y sus implicaciones estratégicas. La tesis que guiará estas páginas es sencilla: la ventana de difusión en inteligencia artificial se está acortando de forma notable, y este fenómeno abre para Europa un espacio de oportunidad en el terreno de los modelos abiertos y la gobernanza tecnológica. Aprovecharlo dependerá de decisiones políticas, inversiones públicas y capacidad de coordinación en los próximos años.

## 2. Breve historia de la IA y el escalado computacional

Hablar hoy de inteligencia artificial suele evocar imágenes de asistentes virtuales capaces de escribir texto, resolver problemas matemáticos o generar imágenes con una naturalidad sorprendente. Sin embargo, el camino hasta llegar a estos sistemas ha sido largo y lleno de oscilaciones entre expectativas desmedidas y periodos de desencanto. Europa, Estados Unidos y, más recientemente, China han recorrido este trayecto con estrategias diferentes, pero todos han experimentado una misma regularidad: cada salto cualitativo en inteligencia artificial ha estado asociado a un salto cuantitativo en capacidad computacional.

Los primeros pasos se remontan a los años cincuenta, cuando Frank Rosenblatt diseñó el *perceptrón*, un modelo inspirado en las neuronas biológicas. Aquellos sistemas eran rudimentarios y pronto se toparon

con limitaciones teóricas. Durante décadas, el progreso fue irregular y los recursos de cómputo escasos. La mayor parte de los algoritmos de los años ochenta y noventa, como las redes neuronales de retropropagación, eran prometedores desde el punto de vista conceptual, pero estaban restringidos por la insuficiencia de hardware.

El punto de inflexión llegó en 2012 con el experimento de Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton, conocido como *AlexNet*. Utilizando tarjetas gráficas (GPUs) diseñadas originalmente para videojuegos, entrenaron una red profunda que redujo drásticamente los errores en la competición ImageNet de reconocimiento visual. Ese resultado no fue solo una victoria técnica, sino el inicio de lo que hoy llamamos la era del *deep learning*: la constatación de que, con suficiente capacidad de cómputo y datos, los algoritmos podían escalar a niveles insospechados (Alom et al., 2018).

A partir de ese momento, el progreso en IA empezó a seguir un patrón empírico sorprendentemente regular: cuanto más grande era el modelo, más datos y más cómputo se invertían, mejor era el rendimiento, y lo hacía siguiendo una curva predecible. Este hallazgo cristalizó con el trabajo de Kaplan et al. (2020), quienes formularon las denominadas *scaling laws*. Dicho de manera sencilla: existe una relación de tipo potencia entre el tamaño del modelo, la cantidad de datos, el cómputo invertido y la pérdida o error en la tarea.

La idea es conceptualmente poderosa ya que no se trata de innovaciones puntuales, sino de un principio de escalabilidad. Si se doblan los recursos, se obtiene una mejora consistente, aunque con retornos decrecientes. Así, la frontera de la inteligencia artificial dejó de ser un terreno de descubrimientos accidentales y pasó a ser una carrera previsible de escalado (Pyzer-Knapp et al., 2025).

El progreso en modelos de IA se mide hoy en cómputo. La unidad básica es el FLOP (*floating-point operation*): una suma, resta, multiplicación o división con números reales realizada por el chip. Un entrenamiento completo de una inteligencia artificial no es más que un gran contador de FLOP. Como se observa en la Figura 1, la escala actual de los modelos de frontera es  $10^{25}$  a  $10^{26}$  FLOP; los grandes modelos de la generación anterior en 2023 se movían en  $10^{23}$  a  $10^{24}$  FLOP. Esta diferencia de dos órdenes de magnitud separa a los laboratorios punteros de los que ya no pueden competir. Para dar una idea, un ordenador personal típico realiza en torno a  $10^{11}$  FLOP por segundo: Un entrenamiento de frontera en la actualidad acumula trillones de trillones de clics, sumando el trabajo de miles de máquinas durante semanas o meses. La diferencia de escala no es de “ordenador rápido vs. lento”, sino que está más cerca de un pequeño taller frente a una gran factoría. En la Figura 1 se observa la evolución del cómputo de entrenamiento en inteligencia artificial.



Figura 1. Evolución del cómputo de entrenamiento en IA (2010–2025)

Fuente: Epoch (2025)

Para expresar “potencia  $\times$  tiempo” se usa el petaFLOP/día (PF-day): ejecutar  $10^{15}$  operaciones por segundo durante 24 horas, es decir, unas  $10^{20}$  operaciones en total. Con este lenguaje, un entrenamiento de  $10^{25}$  FLOP equivale a unos 100.000 PF-days y uno de  $10^{26}$  FLOP a alrededor de 1.000.000 PF-days.

Estas métricas importan que FLOP y PF-days permiten comparar entrenamientos con hardware distinto y arquitecturas diferentes. Sirven para auditar ambición y eficiencia algorítmica sin depender de marcas o marketing. Cada salto de  $\times 10$  en FLOP exige más GPU/TPU, más red, más refrigeración y más electricidad. No es un tecnicismo: determina quién puede entrenar qué y con qué calendario. Así, los presupuestos, infraestructuras y regulación energética se planifican con estas magnitudes. Los países y empresas que no internalicen FLOP y PF-days quedarán fuera de la frontera. Además, al reportar FLOP totales y PF-days junto a parámetros y otros datos permite reconstruir el esfuerzo real, estimar la eficiencia de las mejoras y distinguir avances algorítmicos de mera fuerza bruta.

El efecto acumulado de esta dinámica de aumento de cómputo es que el desarrollo de modelos frontera ha pasado a estar concentrado en muy pocos actores. Mientras que antes bastaba con un laboratorio universitario para experimentar con algoritmos de IA, hoy solo unas pocas empresas (OpenAI, Anthropic, Google DeepMind en Estados Unidos; Baidu, ZhipuAI o Baichuan en China) y unos cuantos consorcios públicos con apoyo gubernamental pueden afrontar el coste de entrenar modelos de decenas de billones de parámetros (Amey et al., 2024).

Europa, por su parte, se ha situado en una posición ambivalente. Por un lado, carece de las inversiones privadas colosales que caracterizan al ecosistema estadounidense. Por otro, ha demostrado capacidad de liderazgo en proyectos abiertos y colaborativos como BLOOM y Mistral que han replicado o rozado el estado del arte con una fracción de los recursos. A ello se suma una infraestructura pública y consorciada que no es marginal: redes EuroHPC, supercomputadores nacionales, y ecosistemas de datos sectoriales impulsados por fondos públicos. La fragmentación regulatoria y de mercado sigue siendo un lastre, pero coexiste con una ventaja diferencial: disponibilidad de datos multilingües y marcos de privacidad que, bien aprovechados, elevan la calidad de los conjuntos de entrenamiento en ámbitos sensibles.

La conclusión es inequívoca: la historia de la inteligencia artificial en las últimas dos décadas puede leerse como la historia del escalado computacional. Cada avance significativo no ha sido únicamente fruto de una mejora de los algoritmos, sino sobre todo de la capacidad de movilizar más cómputo, más datos y más recursos energéticos (de Vries, 2023). El resultado es que, hablar hoy de inteligencia artificial equivale, en buena medida, a hablar de economía política del cómputo.

### 3. La ventana de difusión: casos descriptivos

Uno de los rasgos más llamativos del desarrollo de la inteligencia artificial contemporánea es la creciente reducción del tiempo que transcurre entre el lanzamiento de un modelo propietario de frontera y la aparición de una alternativa abierta de características comparables. Ese intervalo, que hemos denominado aquí ventana de difusión, constituye una medida indirecta de la rapidez con la que el conocimiento tecnológico se propaga más allá de los laboratorios cerrados y llega a comunidades académicas, investigadores independientes y empresas emergentes.

En esta sección repasamos los casos más representativos de los últimos años. No se trata de un catálogo exhaustivo, sino de ejemplos paradigmáticos que ilustran cómo la ventana de difusión ha pasado de años a meses, con implicaciones profundas para la innovación, la gobernanza y la autonomía tecnológica de Europa.

Cuando OpenAI anunció GPT-3 en mayo de 2020, el ecosistema de la IA entró en una nueva era. Con 175.000 millones de parámetros y un coste de entrenamiento estimado en decenas de millones de dólares, el modelo superaba en varios órdenes de magnitud todo lo que existía hasta entonces. Durante más de dos años, GPT-3 permaneció sin equivalente abierto.

La respuesta llegó en julio de 2022, cuando el consorcio internacional BigScience, con fuerte liderazgo europeo y el apoyo de la infraestructura pública del supercomputador Jean Zay en Francia, presentó BLOOM: un modelo de 176.000 millones de parámetros entrenado en 46 lenguas y 13 lenguas de programación. Más allá de la potencia técnica, lo relevante fue el cambio cultural: por primera vez, un modelo de escala comparable a GPT-3 se ponía a disposición de la comunidad global bajo una licencia abierta. La ventana temporal en este caso fue de aproximadamente 26 meses. Representa el último gran ejemplo de un intervalo prolongado, propio de una etapa en la que solo unos pocos actores privados podían financiar entrenamientos a gran escala.

En marzo de 2023, OpenAI presentó GPT-4, un modelo cuyo tamaño exacto no se ha hecho público, pero que estableció un nuevo estándar en capacidades de razonamiento y comprensión del lenguaje. Apenas cuatro meses después, en julio de 2023, Meta anunció LLaMA-2, un modelo abierto de hasta 70.000 millones de parámetros que, en numerosos puntos de referencia, mostraba un rendimiento comparable al de los modelos propietarios más avanzados.

Aquí la ventana de difusión se redujo a meses, no años. El salto refleja no solo la maduración de la comunidad investigadora y el acceso más amplio a infraestructuras de cómputo, sino también un cambio estratégico: grandes empresas tecnológicas, en este caso Meta, comenzaron a liberar modelos de alto nivel como parte de su estrategia competitiva.

Además, en febrero de 2023, Meta distribuyó el modelo LLaMA-1 bajo una licencia restringida a investigadores. Pocas semanas después, los pesos del modelo fueron filtrados en internet, lo que permitió que cualquier desarrollador pudiera experimentar con él. Aunque se trató de una filtración no autorizada, sus efectos fueron comparables a los de una liberación oficial: desencadenó una explosión de proyectos derivados, y adaptaciones que impulsaron una ola de innovación en todo el ecosistema abierto.

Desde el punto de vista de la ventana de difusión, este caso muestra cómo los procesos formales de transferencia de conocimiento pueden verse acelerados por fugas no planificadas, con consecuencias difíciles de controlar tanto para las empresas propietarias como para los reguladores.

La dinámica de aceleración se refleja con claridad en la secuencia Claude de Anthropic, a Mistral, ambos de 2023. Mientras Anthropic avanzaba en el terreno del alineamiento y la seguridad, perfeccionando mecanismos de control para reducir sesgos y respuestas dañinas en sus modelos cerrados, la joven startup francesa Mistral irrumpía con un golpe de efecto: la publicación de un modelo abierto, compacto y altamente eficiente que alcanzaba niveles de rendimiento comparables con una fracción de los recursos invertidos por los gigantes del sector (Jiang et al., 2023). Este contraste ilustra la velocidad con la que la frontera tecnológica puede desplazarse en direcciones divergentes, combinando el esfuerzo intensivo en sofisticación algorítmica con estrategias radicales de eficiencia y apertura que democratizan el acceso y multiplican la capacidad de réplica en el ecosistema global.

En 2025 se ha dado el caso inverso, en donde el modelo abierto se adelantó al cerrado en días. El 5 de abril de 2025 Meta publica LLaMA-4 con pesos descargables y que permite uso y modificación con restricciones para casos de uso comerciales de gran escala. Nueve días después, el 14 de abril de 2025, OpenAI



anuncia GPT-4.1 como su nueva serie de modelos. La ventana de difusión aquí se invierte: el abierto precede al propietario y de hecho precipita la publicación de este modelo. Este patrón sugiere que cuando existen pesos abiertos y licencias de uso amplio, las comunidades y empresas pueden trabajar en su propio hardware o nube sin esperar APIs propietarias, reduciendo la fricción de réplica y acelerando la difusión.

Estos casos, sintetizados en la Tabla 1, refuerzan la idea de que la ventana de difusión ya no es uniforme ni previsible: depende de decisiones estratégicas, de la disponibilidad de hardware y, en ocasiones, de eventos inesperados como filtraciones. Asimismo, el repaso a estos hitos revela una tendencia clara: la ventana de difusión se está acortando de manera sostenida. Lo que en 2020 suponía más de dos años se ha reducido a unos pocos meses. En paralelo, el ecosistema abierto ha adquirido una capacidad organizativa y técnica sin precedentes, capaz de movilizar recursos distribuidos y coordinar proyectos de gran escala.

Para Europa, esta aceleración supone tanto una oportunidad como un desafío. Oportunidad, porque iniciativas como BigScience o Mistral muestran que es posible situarse en la frontera mediante estrategias colaborativas y abiertas. Desafío, porque la rapidez de la difusión obliga a repensar las políticas de apoyo, las infraestructuras de cómputo y la regulación para evitar quedar atrapados en una posición secundaria.

Tabla 1. Ejemplos de ventanas de difusión en IA reciente

Modelo propietario (fecha)	Modelo abierto (fecha)	Intervalo aproximado
GPT-3 (OpenAI, mayo 2020)	BLOOM (BigScience, jul. 2022)	26 meses
GPT-4 (OpenAI, mar. 2023)	LLaMA-2 (Meta, jul. 2023)	4 meses
PaLM (Google, 2022)	Falcon (TII, 2023)	12 meses
Claude (Anthropic, 2023)	Mistral (Francia, 2023)	Menos de 12 meses
LLaMA-1 (Meta, feb. 2023)	Filtración (mar. 2023)	Semanas
GPT-4.1 (OpenAI, abr. 2025)	LLaMA-4 (Meta, abr. 2025)	Modelo abierto antes

Fuente: Elaboración propia

La evidencia descriptiva es inequívoca: la ventana de difusión se ha estrechado radicalmente. Lo que falta por explorar, y que será el objeto de la sección siguiente, son los motores de esa aceleración, desde la caída de los costes de cómputo hasta la competencia geopolítica por el liderazgo en inteligencia artificial.

4. Tendencias recientes en costes y hardware

El acelerado ritmo de avance en inteligencia artificial se explica por un factor material ineludible: el coste del cómputo. Durante los últimos quince años, entrenar modelos de frontera ha requerido cantidades crecientes de operaciones de cálculo, medibles en FLOP, que se traducen en costes económicos y energéticos cada vez más elevados. Sin embargo, en paralelo a este aumento absoluto, se ha producido una caída sostenida del coste relativo por unidad de cómputo, lo que explica que iniciativas abiertas y consorcios académicos hayan podido competir con proyectos corporativos de miles de millones de dólares.

Desde los primeros modelos de *deep learning* en la década de 2010, el coste de una operación de punto flotante ha seguido una trayectoria decreciente casi exponencial, heredera del espíritu de la ley de Moore. Estudios recientes (Sevilla et al., 2022; Epoch, 2025) muestran que, aunque la velocidad de mejora del hardware se ha ralentizado en algunos segmentos, el coste por FLOP sigue reduciéndose a un ritmo que hace que sea posible duplicar los FLOPs por cada dólar de gasto en alrededor de dos años como se observa en la Figura 2.

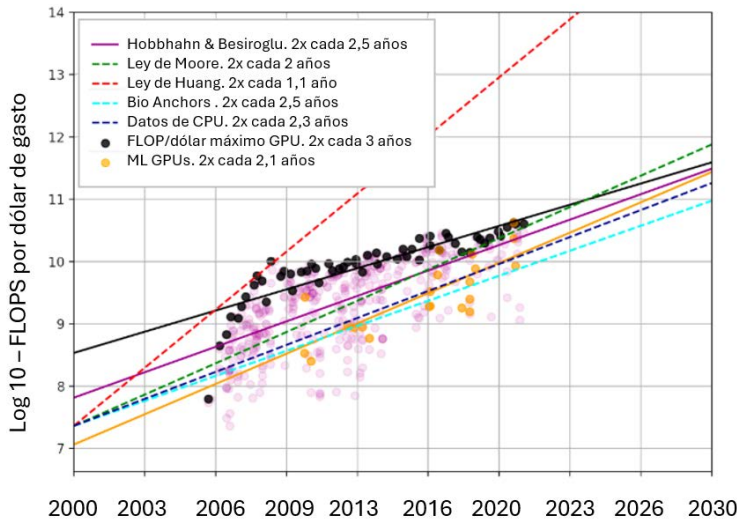


Figura 2. Evolución del coste por FLOP en entrenamiento de IA, 2010-2025

Fuente: Hobbhahn & Besiroglu (2022)

La magnitud del cambio es difícil de exagerar: entrenar un modelo de gran escala en 2012 podía costar decenas de millones de dólares en infraestructura que solo estaba al alcance de grandes empresas tecnológicas. Una década más tarde, proyectos colaborativos como BLOOM pudieron movilizar un cómputo comparable por una fracción del precio gracias a la combinación de hardware más eficiente, mejores algoritmos de optimización y, sobre todo, economías de escala en infraestructuras compartidas.

El motor de esta reducción ha sido la innovación en hardware. Las tarjetas gráficas (GPUs), inicialmente concebidas para videojuegos, fueron el catalizador de la revolución del *deep learning*. Posteriormente, Google introdujo las TPUs (*Tensor Processing Units*), diseñadas específicamente para operaciones de aprendizaje automático. Hoy asistimos a una diversificación del ecosistema con GPUs de nueva generación que multiplican por varios órdenes la eficiencia respecto a generaciones anteriores, chips especializados desarrollados por empresas emergentes y proyectos de código abierto y supercomputadores europeos como Jean Zay en Francia o Leonardo en Italia, que representan intentos de crear una base propia de cómputo para la investigación en IA. El resultado es un entorno competitivo en rápida evolución, donde el acceso a hardware de última generación se ha convertido en un factor geopolítico.

El cómputo no es solo una cuestión de eficiencia técnica; es también un recurso estratégico sujeto a tensiones internacionales. Las restricciones a la exportación de GPUs avanzadas impuestas por Estados Unidos a China desde 2022 han alterado profundamente el mercado, generando incertidumbre sobre el acceso a componentes críticos. Aunque Europa no es el objetivo directo de esas medidas, sus consecuencias indirectas son claras: encarecimiento de la oferta global, competencia por recursos limitados y mayor dependencia de unos pocos proveedores (Schrepel & Pentland, 2024).

Esta dimensión geopolítica subraya la vulnerabilidad de Europa. Sin una base industrial propia en semiconductores de última generación, la región depende de importaciones para sostener cualquier proyecto de IA de frontera. La European Chips Act, aprobado en 2023, es un intento de reducir esta dependencia, pero sus efectos tardarán años en materializarse (Parlamento Europeo y Consejo, 2023).

Paradójicamente, la caída del coste relativo por FLOP no significa que entrenar modelos sea más barato en términos absolutos. Al contrario: la lógica de las *scaling laws* empuja a usar siempre más cómputo, de modo que el gasto total crece incluso mientras el coste unitario disminuye. Es el fenómeno que algunos analistas describen como una “espiral de escalado”: cada mejora en eficiencia se traduce en un aumento en la ambición de los proyectos.

Así, mientras el entrenamiento de AlexNet en 2012 podía realizarse con unas pocas GPUs, GPT-3 en 2020 necesitó miles de ellas funcionando durante semanas, y los modelos más recientes consumen ya millones de horas de GPU equivalentes. La accesibilidad relativa ha mejorado, pero la barrera de entrada absoluta sigue siendo altísima.

La relación entre costes y difusión es directa: Cuando los costes caen, los proyectos académicos y consorcios abiertos pueden acercarse antes a la frontera. En cambio, cuando los costes se disparan, la ventana de difusión se amplía, pues solo unos pocos actores pueden sostener el esfuerzo inicial.

El caso de BLOOM frente a GPT-3 se entiende mejor a la luz de esta dinámica: en 2020, el coste del entrenamiento estaba todavía fuera del alcance de la comunidad abierta; en 2022, la mejora del hardware y la disponibilidad de infraestructuras compartidas permitió acortar la distancia. Desde entonces, cada nueva generación de hardware ha reducido aún más el intervalo.

En síntesis, la evolución reciente muestra una doble cara. Por un lado, la caída del coste relativo por FLOP ha abierto la puerta a que iniciativas abiertas y actores europeos puedan competir en el terreno de la IA de gran escala. Por otro, la dependencia de cadenas de suministro externas y el aumento del gasto absoluto consolidan un escenario en el que el cómputo se convierte en recurso estratégico de primer orden. La ventana de difusión no puede entenderse sin atender a esta paradoja: el hardware abarata la réplica, pero también eleva las apuestas.

## 5. Europa en la carrera global

Hablar de inteligencia artificial en la actualidad es hacerlo de una carrera geopolítica en la que Estados Unidos y China concentran los mayores recursos, las empresas más avanzadas y la mayor parte de los modelos de frontera. En ese contexto, Europa se encuentra en una posición intermedia: rezagada en la competición de escala bruta, pero con fortalezas notables en el terreno de los modelos abiertos, la regulación y la cooperación internacional.

Estados Unidos sigue siendo el epicentro de la inteligencia artificial de frontera. Empresas como OpenAI, Anthropic, Google DeepMind y Microsoft han establecido un oligopolio de facto en la producción de modelos con miles de millones de parámetros, respaldados por presupuestos de miles de millones de dólares y acceso privilegiado a infraestructuras de cómputo masivo.

La estrategia estadounidense combina tres elementos: Una escala de cómputo sin precedentes, apoyada en el dominio de NVIDIA en GPUs, un capital privado abundante, con rondas de financiación que superan los miles de millones, y alianzas estratégicas entre startups y gigantes tecnológicos con el caso paradigmático es la inversión de Microsoft en OpenAI. El resultado es un liderazgo tecnológico indiscutible, pero también una creciente concentración del poder en manos de pocas empresas.

China, aunque más tardía en la carrera, ha logrado reducir distancias a gran velocidad. Institutos como ZhipuAI, Baichuan o SenseTime han desarrollado modelos de escala similar a los occidentales, con un fuerte apoyo estatal y una estrategia explícita de autonomía tecnológica.

Las restricciones a la exportación de hardware impuestas por Estados Unidos desde 2022 han obligado a China a buscar soluciones alternativas, como el desarrollo de chips nacionales y la construcción de supercomputadores optimizados para IA. Aunque estas limitaciones ralentizan el acceso al hardware más avanzado, también estimulan la creación de una base industrial propia.

China, en suma, representa una estrategia estatal de largo plazo, orientada a asegurar que la ventana de difusión no dependa de decisiones externas, sino de la capacidad de movilizar recursos internos.

Frente a estas dos potencias, Europa aparece en ocasiones como un actor secundario. Carece de empresas capaces de invertir decenas de miles de millones en un único modelo y depende en gran medida de hardware importado. Sin embargo, en un aspecto crucial Europa ha conseguido proyectar liderazgo global: el de los modelos abiertos y colaborativos.

El caso más emblemático es el de BLOOM (2022), fruto de la iniciativa *BigScience* impulsada desde París por HuggingFace y financiada en parte con recursos públicos franceses. BLOOM no solo fue un logro técnico, sino también un símbolo político: demostró que Europa podía coordinar un consorcio internacional, movilizar supercomputación pública y entregar a la comunidad científica un modelo de escala GPT-3 con licencia abierta. En la Figura 3 se observa como BLOOM fue el modelo abierto más preciso y en la actualidad los modelos LLaMA, de Meta, han convergido en cuanto a precisión respecto a los modelos cerrados de frontera.



Figura 3: Mapa comparativo de iniciativas en IA cerrados y abiertos

Fuente: Cottier et al. (2024)

A este ejemplo se suman iniciativas recientes como Mistral, una startup francesa fundada en 2023 que ha conseguido situar modelos abiertos altamente competitivos en el centro del debate internacional, o proyectos apoyados por universidades y supercomputadores nacionales (Jean Zay en Francia, Leonardo en Italia, MareNostrum en España).

Europa cuenta además con un activo estratégico poco reconocido: su red de supercomputadores públicos, muchos de ellos integrados en la red EuroHPC. Aunque originalmente concebidos para simulaciones científicas y cálculos de física o biología, estos centros han comenzado a destinar parte de su capacidad a proyectos de inteligencia artificial.

No obstante, el acceso sigue siendo limitado y fragmentado. La falta de un programa europeo coordinado de cómputo abierto para IA dificulta que universidades y startups puedan competir en igualdad de condiciones con sus homólogos estadounidenses o chinos. Así, hay tres modelos distintos de aproximación a la frontera de la IA. Estados Unidos cuenta con dominio corporativo, impulsado por grandes empresas privadas con acceso a capital y hardware global. China ha optado por una estrategia estatal, con inversión masiva y apuesta por la autonomía en hardware. Por su parte, Europa cuenta con liderazgo en el espacio abierto y colaborativo, pero dependiente en hardware y financiación de actores externos.

El panorama no es necesariamente desfavorable para Europa. Precisamente porque la ventana de difusión se está acortando, el espacio de ventaja no reside en reproducir la escala bruta de los modelos estadounidenses o chinos, sino en liderar la esfera abierta. El capital social, la tradición académica y la infraestructura pública de Europa le permiten desempeñar un papel crucial como garante de transparencia, accesibilidad y responsabilidad en la IA.

En un mundo donde la innovación se mide no solo en FLOP, sino también en legitimidad social y capacidad de regulación, Europa puede convertir su aparente debilidad en fortaleza. La apuesta por lo abierto, lejos de ser un recurso secundario, puede ser la clave para marcar la diferencia en la próxima década.

## 6. Implicaciones para la política europea

Si algo demuestra el análisis de la ventana de difusión en inteligencia artificial es que la cuestión no es solo técnica, sino profundamente política y estratégica. La rapidez con que los modelos abiertos alcanzan a los

propietarios redefine el equilibrio global, y Europa debe decidir cómo posicionarse en un terreno donde Estados Unidos y China marcan el paso.

Las implicaciones para la política europea pueden agruparse en cuatro dimensiones principales: la dependencia tecnológica, las oportunidades del espacio abierto, la integración con las agendas comunitarias y la necesidad de infraestructuras compartidas.

Europa carece hoy de un actor corporativo capaz de competir con OpenAI, Anthropic o Google DeepMind en escala de inversión y cómputo. Esto genera una dependencia tecnológica que afecta a múltiples niveles. En el ámbito de la investigación, universidades y centros europeos se ven obligados a trabajar con APIs propietarias de empresas extranjeras, lo que limita su autonomía y condiciona el tipo de experimentos avanzados que pueden realizar. En el plano industrial, startups y pymes europeas desarrollan sus productos y servicios sujetos a licencias, precios y restricciones impuestas desde fuera, lo que frena la capacidad de innovación endógena y aumenta la vulnerabilidad competitiva. En materia de gobernanza, la definición de estándares y prácticas de seguridad queda en manos de corporaciones que operan fuera de la jurisdicción europea, lo que reduce el margen de acción normativa y limita la soberanía tecnológica de la región. Reducir esta dependencia tecnológica no significa replicar el modelo estadounidense de grandes empresas privadas, sino construir un ecosistema alternativo que aproveche las fortalezas de Europa en apertura y regulación.

Además, la ventana de difusión se ha convertido en una palanca de oportunidad para Europa. Iniciativas como BLOOM o Mistral demuestran que, con coordinación y acceso a cómputo público, es posible situarse en la frontera en cuestión de meses. Esta estrategia ofrece ventajas claras, tales como transparencia y reproducibilidad, esenciales para la investigación científica, democratización del acceso, que favorece la innovación distribuida en startups, universidades y pymes, así como legitimidad social, en un momento de desconfianza creciente hacia la concentración de poder tecnológico en unas pocas corporaciones. Europa puede liderar no en escala bruta, sino en la definición de estándares abiertos y éticos que guíen el desarrollo global de la IA.

La inteligencia artificial no se desarrolla en el vacío: debe articularse con las grandes agendas europeas, que marcan el rumbo estratégico de la Unión y condicionan tanto el marco regulatorio como las capacidades industriales y tecnológicas disponibles.

- La AI Act, aprobada en 2024, constituye el primer marco regulador integral sobre inteligencia artificial en el mundo. Su enfoque en clasificación de riesgos, transparencia y derechos fundamentales convierte a Europa en pionera en gobernanza tecnológica (Parlamento Europeo y Consejo, 2024). En este contexto, la disponibilidad de modelos abiertos y auditables resulta decisiva: al exponer su funcionamiento interno, facilitan la verificación de sesgos, el control de usos indebidos y la adaptación de la normativa en función de la evidencia empírica (Balcioglu et al., 2025)
- La Chips Act responde a una vulnerabilidad crítica: la dependencia europea de semiconductores importados para el cómputo avanzado. La IA de frontera exige hardware especializado, cuya producción está hoy concentrada en pocas regiones del mundo. El desafío europeo es que la construcción de una base industrial propia en semiconductores requiere plazos largos de inversión, desarrollo y escalado, mientras que la ventana de difusión de la IA avanza con gran rapidez, creando un desajuste entre necesidad inmediata y capacidad instalada.
- La Agenda Digital Europea persigue una transición digital inclusiva y competitiva, donde la IA ocupa un papel central en productividad, servicios públicos y transformación sectorial. La disponibilidad de modelos abiertos de alta calidad permite que pymes, administraciones y actores sin grandes presupuestos accedan a herramientas avanzadas sin quedar relegados a depender de licencias costosas o soluciones cerradas, democratizando así los beneficios de la digitalización.
- El Pacto Verde Europeo conecta la estrategia climática con la innovación tecnológica. La IA de gran escala implica consumos energéticos y de recursos significativos, desde el entrenamiento hasta la inferencia. Europa tiene la posibilidad de marcar una diferencia global impulsando modelos más eficientes en cómputo y energéticamente sostenibles, alineando la investigación en IA con sus compromisos de neutralidad climática y convirtiendo la sostenibilidad en un vector de competitividad tecnológica.

De entre todos los obstáculos, el más inmediato para que Europa aproveche la ventana de difusión es el acceso al cómputo. Hoy, incluso proyectos abiertos dependen de acuerdos bilaterales con proveedores de hardware o de un acceso limitado a supercomputadores nacionales. Una opción es avanzar hacia un "European Open Compute Cloud", una infraestructura compartida que permita a universidades, startups y consorcios acceder a recursos de cómputo de última generación con reglas claras y financiación pública. Este modelo sería comparable al CERN en física de partículas: una gran infraestructura paneuropea al servicio de la ciencia y la innovación.

La combinación de estos elementos apunta hacia una estrategia que podríamos denominar autonomía estratégica abierta. Europa no puede ni debe aspirar a replicar la escala de Estados Unidos o China en solitario, pero sí puede consolidar un modelo propio basado en tres pilares:

- I. Infraestructuras de cómputo públicas y accesibles.
- II. Iniciativas abiertas y colaborativas como BLOOM o Mistral.
- III. Un marco regulatorio robusto que garantice seguridad, derechos y sostenibilidad (Bolgouras et al., 2025).



De esta manera, la ventana de difusión deja de ser un factor de vulnerabilidad para convertirse en un recurso estratégico: un intervalo que Europa puede aprovechar para posicionarse como líder en transparencia, legitimidad y cooperación internacional.

## 7. Perspectivas de futuro

Si algo enseña la trayectoria reciente de la inteligencia artificial es que las tendencias que hoy parecen sólidas pueden transformarse de forma abrupta. En 2020 nadie habría predicho que un consorcio abierto como BigScience conseguiría entrenar un modelo comparable a GPT-3. Tampoco que una filtración como la de LLaMA-1 en 2023 desencadenaría una ola de innovación distribuida que alteraría los calendarios de empresas y reguladores. En 2025 se produjo una inversión puntual de la ventana: LLaMA-4 (5 de abril) precedió a GPT-4.1 (14 de abril), mostrando que el abierto puede adelantarse al propietario en cuestión de días. El futuro de la ventana de difusión en IA es, por tanto, incierto, pero se pueden identificar tres grandes escenarios plausibles que orientan la reflexión estratégica de Europa.

En el primer escenario, la dinámica observada en 2023 y de forma más reciente se consolida: la distancia entre un modelo propietario y su equivalente abierto se vuelve consistentemente reducida. La explicación sería doble: Por un lado, la maduración de comunidades abiertas capaces de movilizar rápidamente infraestructuras distribuidas y coordinar equipos internacionales. Por otro, la decisión de grandes empresas de liberar modelos avanzados como parte de su estrategia competitiva, acelerando la disponibilidad de capacidades de frontera.

Las consecuencias son ambivalentes. Una ventana muy corta permitiría a universidades, startups y gobiernos acceder casi en tiempo real a tecnologías de vanguardia. Pero también aumentaría los riesgos de proliferación incontrolada, dificultando los esfuerzos de seguridad y gobernanza.

El segundo escenario es que la ventana vuelva a ampliarse, no por razones técnicas, sino por restricciones políticas y comerciales. Si los costes absolutos de entrenamiento siguen creciendo y las empresas propietarias optan por cerrar aún más el acceso a sus modelos, la difusión podría ralentizarse.

En este caso, los modelos abiertos quedarían limitados a escalas intermedias, y los actores europeos dependerían de APIs privadas gestionadas por empresas estadounidenses o chinas. La ventana de difusión se convertiría así en un mecanismo de poder, prolongado deliberadamente para mantener ventajas competitivas. El riesgo para Europa sería alto: una posición de dependencia estructural, sin capacidad real de influir en la agenda tecnológica global.

El tercer escenario sería el más favorable para Europa y daría si se combina la aceleración de la ventana con una estrategia deliberada de autonomía. Aquí, la región aprovecharía su capacidad de coordinación institucional y su tradición de ciencia abierta para situarse como referente en modelos accesibles, transparentes y auditables.

En la práctica, esto supone consolidar un ecosistema europeo de supercomputación para IA, articulado sobre EuroHPC y nodos nacionales, financiado con fondos comunitarios, con acceso federado y tarifas subvencionadas para universidades, centros públicos y startups y colas de prioridad para proyectos estratégicos; respaldar proyectos abiertos de gran escala (BLOOM, Mistral y sucesores) con financiación plurianual, repositorios de datos auditados, licencias claras y gobernanza técnica independiente que garantice reproducibilidad y transferencia a la industria; impulsar espacios comunes de datos sectoriales (salud, energía, movilidad, manufactura) con estándares de interoperabilidad, anonimización sólida y acuerdos de cooperación que permitan entrenamiento y evaluación transfronteriza; desplegar un marco regulatorio con alcance internacional que alinee el IA Act con guías técnicas de evaluación ex ante y ex post, informe de huella de cómputo y trazabilidad de bases de datos; establecer bancos de pruebas y puntos de referencia públicos europeos con métricas de seguridad, robustez, eficiencia energética y desempeño multilingüe, vinculando resultados a la elegibilidad de ayudas; y crear instrumentos financieros mixtos como IPCEI (*Important Projects of Common European Interest*), compra pública precomercial y créditos blandos que reduzcan el coste de capital para escalar del laboratorio al producto, de modo que transparencia y responsabilidad dejen de ser carga regulatoria y pasen a operar como ventajas competitivas exportables.

En este escenario, Europa no lideraría en escala bruta, pero sí en legitimidad y gobernanza, actuando como árbitro de facto en la fijación de estándares y buenas prácticas. La palanca sería regulatoria y de compra: requisitos de interoperabilidad y documentación técnica, auditorías de seguridad para despliegues de alto riesgo, y acuerdos tipo para compartición de datos en espacios de confianza con anonimización fuerte. La gobernanza se consolidaría mediante foros europeos abiertos, bancos de pruebas neutrales operados por consorcios académicos-industriales, y un régimen de conformidad graduado que favorezca modelos abiertos y eficientes sin eximirlos de pruebas. El resultado sería un “sello UE” de calidad técnica y ética que terceros adoptan por portabilidad regulatoria y por coste de cumplimiento marginal, con Europa marcando el listón de evaluabilidad, seguridad y transparencia incluso cuando el mayor cómputo esté fuera (Cantero Gamito, 2024).

El rumbo de la ventana de difusión dependerá de variables aún indeterminadas: la evolución del hardware, es decir, si aparecen cuellos de botella sostenidos en GPUs avanzadas y si la capacidad de réplica y escala se comprimirá; las dinámicas corporativas de las grandes empresas del sector; los eventos inesperados como filtraciones, fugas de modelos, rupturas algorítmicas o mejoras de eficiencia de entrenamiento capaces de alterar la frontera en semanas; y la política internacional, por ejemplo con unas crecientes tensiones EE.UU.-China que reconfiguren las cadenas de suministro o conduzcan a controles a la exportación condicionando el margen de maniobra europeo.

En cualquier escenario, Europa afronta un dilema: ser mero receptor de tecnologías definidas en otros lugares, o convertirse en un actor normativo y científico central. Los riesgos son claros: dependencia tecnológica, pérdida de competitividad, incapacidad de controlar los efectos sociales y económicos de la IA. Las oportunidades también: liderar el espacio abierto, consolidar un modelo alternativo basado en la transparencia y la cooperación, y reforzar su autonomía estratégica en el terreno digital.

En síntesis, el futuro de la ventana de difusión no está predeterminado. Europa se encuentra en un punto de inflexión: la velocidad creciente de la difusión puede ser una amenaza si se aborda pasivamente, pero puede convertirse en una ventaja si se acompaña de políticas activas, infraestructuras compartidas y una visión estratégica de largo plazo.

## 8. Conclusión

La historia reciente de la inteligencia artificial puede leerse como una sucesión de avances en algoritmos y datos, pero, sobre todo, como la historia del cómputo y su difusión. Cada salto en la frontera ha estado acompañado de un aumento exponencial en los recursos necesarios y de un intervalo, la ventana de difusión, que determina cuándo y cómo esas innovaciones llegan a la comunidad abierta.

En 2020, la diferencia entre el anuncio de GPT-3 y la aparición de BLOOM era de más de dos años. Tres años más tarde, con GPT-4 y LLaMA-2, la ventana se había reducido a apenas cuatro meses. En 2025, la ventana llegó a invertirse puntualmente: LLaMA-4 (5 de abril) se publicó con pesos abiertos bajo licencia comunitaria y precedió a GPT-4.1 (14 de abril).

La evidencia es clara: la ventana de difusión se está acortando. Lo que antes era un intervalo prolongado que consolidaba la hegemonía de unos pocos actores hoy se convierte en un margen estrecho en el que comunidades abiertas, startups y consorcios internacionales pueden disputar parte de la frontera. Este cambio altera profundamente el mapa de la innovación y plantea dilemas urgentes de gobernanza.

Para Europa, la lección es doble. Por un lado, existe un riesgo real de dependencia tecnológica, agravado por la concentración de hardware avanzado en manos de proveedores externos y por la ausencia de empresas capaces de competir en escala con los gigantes estadounidenses o chinos. Por otro, se abre una ventana de oportunidad inédita: el liderazgo en el terreno de los modelos abiertos, transparentes y colaborativos.

La apuesta europea por proyectos como BLOOM, HuggingFace o Mistral demuestra que la región puede situarse en el centro del debate internacional sin necesidad de replicar el modelo corporativo de Estados Unidos ni la estrategia estatal de China. Lo que Europa puede ofrecer es un modelo alternativo, basado en tres pilares:

- I. Infraestructuras compartidas de cómputo, que garanticen acceso justo y sostenible a investigadores y empresas emergentes.
- II. Modelos abiertos y auditables, que refuercen la legitimidad social y científica del desarrollo en IA.
- III. Un marco regulatorio pionero, como la AI Act, que convierta la seguridad, la transparencia y la sostenibilidad en ventajas competitivas.

El reto es que la ventana de difusión no espere. Su aceleración significa que las decisiones deben tomarse con rapidez, antes de que la concentración corporativa o las tensiones geopolíticas cierren el margen de maniobra. Europa se encuentra en una encrucijada: puede resignarse a ser un consumidor pasivo de tecnologías desarrolladas fuera, o puede convertirse en un referente global en gobernanza y apertura.

La inteligencia artificial es, en última instancia, una tecnología de propósito general, llamada a transformar la economía, la política y la vida cotidiana en las próximas décadas. El modo en que Europa gestione la ventana de difusión determinará no solo su competitividad económica, sino también su capacidad de preservar valores fundamentales como la democracia, la equidad y la sostenibilidad.

El mensaje es claro: en la carrera global por la inteligencia artificial, la apertura es la ventaja estratégica de Europa. Convertir esa intuición en una política coherente y ambiciosa es quizás uno de los mayores desafíos y oportunidades del proyecto europeo en el siglo XXI.

## Referencias

- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2018). *The history began from alexnet: A comprehensive survey on deep learning approaches*. arXiv preprint arXiv:1803.01164. <https://doi.org/10.48550/arXiv.1803.01164>
- Ameye, N., Bughin, J., & van Zeebroeck, N. (2024). From experimentation to scaling: what shapes the funnel of AI adoption?. *Economics of Innovation and New Technology*, 1-15. <https://doi.org/10.1080/10438599.2024.2413940>
- Balcioğlu, Y. S., Çelik, A. A., & Altındağ, E. (2025). A Turning Point in AI: Europe's Human-Centric Approach to Technology Regulation. *Journal of Responsible Technology*, 100128. <https://doi.org/10.1016/j.jrt.2025.100128>
- BigScience Workshop (2022). *BLOOM: A 176B Parameter Open-Access Multilingual Language Model*. arXiv preprint arXiv:2211.05100. <https://doi.org/10.48550/arXiv.2211.05100>
- Bolgouras, V., Zarras, A., Leka, C., Stylianou, I., Farao, A., & Xenakis, C. (2025). EU regulatory ecosystem for ethical AI. *AI and Ethics*, 5, 5063–5080 (2025). <https://doi.org/10.1007/s43681-025-00749-x>

- Cantero Gamito, M. (2024). The Role of ETSI in the EU's Regulation and Governance of Artificial Intelligence. *Innovation: The European Journal of Social Science Research*, 37(5), 1425-1440. <https://doi.org/10.1080/13511610.2024.2349627>
- Comin, D., & Hobijn, B. (2010). *An Exploration of Technology Diffusion*. *American Economic Review*, 100(5), 2031-2059. <https://doi.org/10.1257/aer.100.5.2031>
- Cottier, B., You, J., Martemianova, N., & Owen, D. (2024). How far behind are open models. Epoch AI. Disponible en: <https://epoch.ai/blog/open-models-report>
- de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, 7(10), 2191-2194. <https://doi.org/10.1016/j.joule.2023.09.004>
- Epoch (2025). Epoch AI Database: Data on AI Models [conjunto de datos]. Epoch AI. Disponible en: <https://epoch.ai/data/ai-models>
- Hobbbahn, M., & Besiroglu, T. (2022). Trends in GPU price-performance. Epoch AI. Disponible en: <https://epoch.ai/blog/trends-in-gpu-price-performance>
- Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., ... & Irving, G. (2022). *Training Compute-Optimal Large Language Models*. arXiv preprint arXiv:2203.15556. <https://doi.org/10.48550/arXiv.2203.15556>
- Jiang, A. Q., Sablayrolles, A., Mensch, A., Bressand, F., Lample, G., Lacroix, T., & Synnaeve, G. (2023). *Mistral 7B*. arXiv preprint arXiv:2310.06825. <https://doi.org/10.48550/arXiv.2310.06825>
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T., Chess, B., Child, R., ... & Amodei, D. (2020). *Scaling Laws for Neural Language Models*. arXiv preprint arXiv:2001.08361. <https://doi.org/10.48550/arXiv.2001.08361>
- Parlamento Europeo y Consejo (2023). *European Chips Act*. Bruselas: Diario Oficial de la Unión Europea. BOE.es - DOUE-L-2023-81291 Reglamento (UE) 2023/1781 del Parlamento Europeo y del Consejo de 13 de septiembre de 2023 por el que se establece un marco de medidas para reforzar el ecosistema europeo de semiconductores y por el que se modifica el Reglamento (UE) 2021/694 (Reglamento de chips).
- Parlamento Europeo y Consejo (2024). *AI Act: Regulation of Artificial Intelligence*. Bruselas: Diario Oficial de la Unión Europea. BOE.es - DOUE-L-2024-81079 Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial).
- Pyzer-Knapp, E. O., Manica, M., Staar, P., Morin, L., Ruch, P., Laino, T., ... & Curioni, A. (2025). Foundation models for materials discovery-current state and future directions. *Npj Computational Materials*, 11(1), 61. <https://doi.org/10.1038/s41524-025-01538-0>
- Sevilla, J., Villalobos, P., Gray, C., Kumbier, K., Ho, A., Besiroglu, T., ... & Kanan, C. (2022). *Compute Trends Across Three Eras of Machine Learning*. arXiv preprint arXiv:2202.05924. <https://doi.org/10.1109/IJCNN55064.2022.9891914>
- Schrepeel, T., & Pentland, A. S. (2024). Competition between AI foundation models: dynamics and policy recommendations. *Industrial and Corporate Change*, dtae042. <https://doi.org/10.1093/icc/dtae042>
- Shetty, D. K., Arjunan, R. V., Cenitta, D., Makkithaya, K., Hegde, N. V., Salu, S., ... & Pallela, P. K. (2025). Analyzing AI Regulation through Literature and Current Trends. *Journal of Open Innovation: Technology, Market, and Complexity*, 100508. <https://doi.org/10.1016/j.joitmc.2025.100508>
- Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). *LLaMA 2: Open Foundation and Fine-Tuned Chat Models*. arXiv preprint arXiv:2307.09288. <https://doi.org/10.48550/arXiv.2307.09288>