

Ética e IA: Agitado, no revuelto

Aspectos Éticos, Legislación y Profesión

Índice

1. Introducción y motivación	2
1.1. Introducción y motivación I	2
1.2. Introducción y motivación II	2
1.3. ¿Has sido capaz de detectarlo?	3
2. ¿Qué es y qué no es la IA?	4
2.1. Definiciones de Inteligencia Artificial	4
2.2. Breve historia de la Inteligencia Artificial	4
2.3. Qué es la Inteligencia Artificial	4
2.4. Tipos/Niveles de IA	5
3. Legislación y acuerdos supranacionales sobre IA	6
3.1. España	6
3.2. Unión Europea	6
3.3. Naciones Unidas	7
4. Ética e IA: entrenamiento de modelos	8
5. Ética e IA: supervisión y corrección	9
6. Ética e IA: Brecha de acceso	10
7. Ética e IA: Impacto climático	11
7.1. La IA es positiva	11
7.2. La IA es negativa	11
8. Ética e IA: ¿Revolución o conflicto educativo?	13
9. Conclusión	14
10. Bibliografía	15

1. Introducción y motivación

Antes de continuar, hazte la siguiente pregunta: ¿serías capaz de detectar un texto creado por una Inteligencia Artificial?

1.1. Introducción y motivación I

La inteligencia artificial (IA) es una de las áreas más innovadoras y de mayor crecimiento en la industria tecnológica. La IA es una rama de la informática que se enfoca en la creación de sistemas inteligentes que pueden realizar tareas que normalmente requieren inteligencia humana, como el aprendizaje, el razonamiento y la resolución de problemas. Algunas de las aplicaciones más populares de la IA incluyen los chatbots, los sistemas de recomendación, los asistentes virtuales, la automatización de procesos y los vehículos autónomos.

Sin embargo, la IA también plantea importantes implicaciones éticas que deben ser consideradas en su desarrollo y uso. En particular, la privacidad y la seguridad son preocupaciones clave debido al gran volumen de datos que se utilizan en los sistemas de IA y la amenaza que esto representa para la información personal y confidencial. Además, la discriminación es otro problema importante relacionado con la IA, ya que los sistemas de IA pueden perpetuar prejuicios y sesgos de género, raza o clase social.

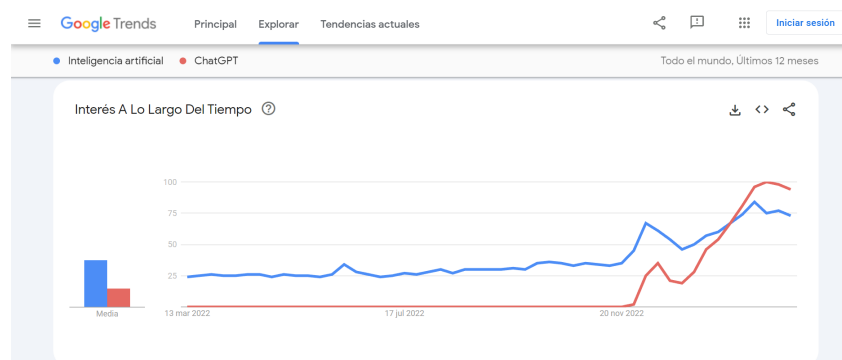
Otro tema importante en la ética de la IA es el poder que se concentra en las empresas que desarrollan y utilizan estas tecnologías. Las grandes empresas de tecnología tienen acceso a una gran cantidad de datos y recursos, lo que les permite crear sistemas de IA más avanzados y sofisticados. Esto les da una ventaja competitiva significativa y les permite consolidar su posición en el mercado, lo que puede ser perjudicial para la competencia y la innovación.

En resumen, la IA es una tecnología con un gran potencial para cambiar el mundo, pero también con importantes implicaciones éticas que deben ser abordadas. Es importante que los desarrolladores y usuarios de la IA consideren cuidadosamente estas implicaciones y trabajen juntos para garantizar que la IA se utilice de manera responsable y ética.

1.2. Introducción y motivación II

Inteligencia Artificial, IA, ChatGPT... Son conceptos que en los últimos meses todos hemos oído. En algunas ocasiones con estupor, en otras con curiosidad, e incluso algunas con miedo.

Este interés queda demostrado en las estadísticas de infinidad de organizaciones. Por poner un ejemplo, en la imagen a continuación podemos ver el interés por los términos *Inteligencia Artificial* y *ChatGPT* en la Búsqueda de Google. (Google, 2023):

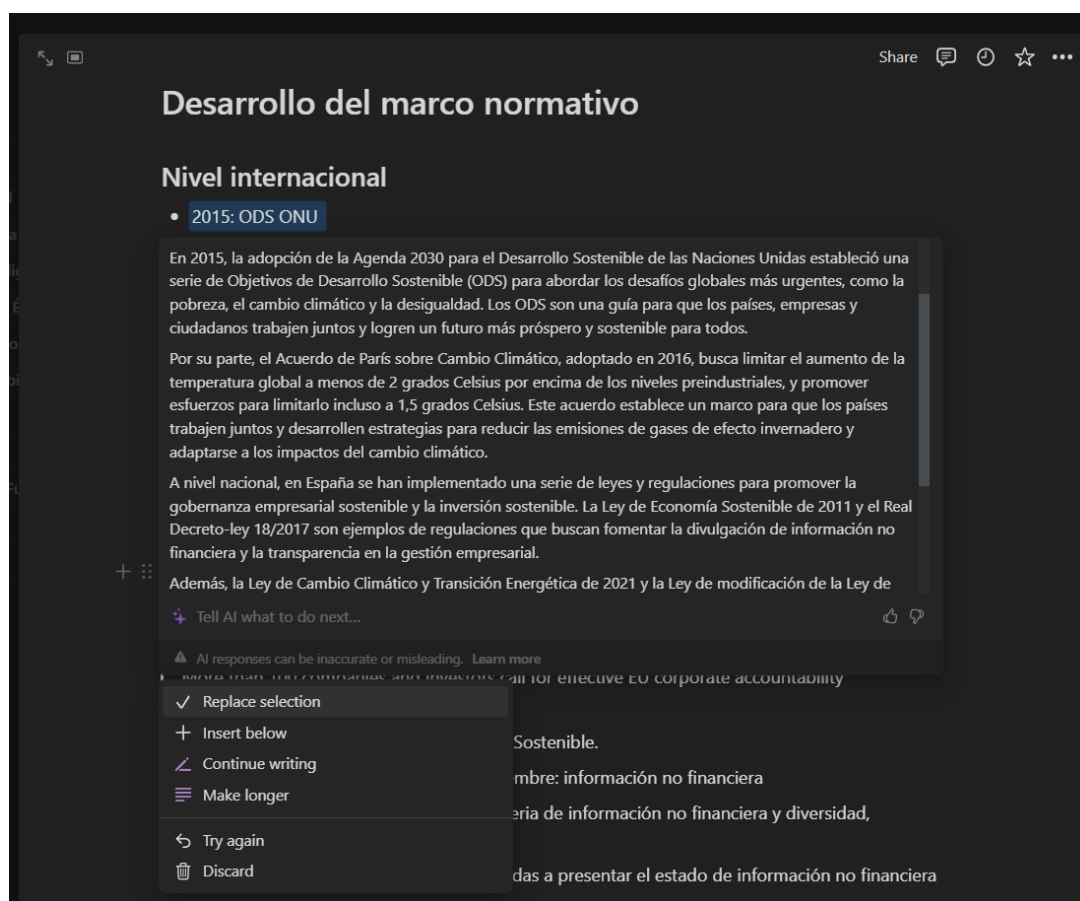


La motivación para realizar este trabajo es, principalmente, la curiosidad que me genera el tema. Personalmente, aunque al principio me consideraba bastante escéptico respecto a su fiabilidad, he de confesar que tras probarlo estoy gratamente sorprendido.

He usado ChatGPT y Notion AI de forma puntual durante las últimas semanas para poder tener una imagen más completa de la realidad. *ChatGPT* lo he utilizando principalmente para dudas tecnológicas y de desarrollo. No es perfecto, de hecho está muy lejos de serlo, pero con una explicación detallada del problema es capaz de *entenderlo* y orientarte correctamente. Por supuesto, el usuario/a debe ser capaz de interpretar la respuesta y discernir entre lo que es correcto y lo que no es.

En tanto a *Notion AI*, mi uso ha sido limitado pues es posterior a la aparición de ChatGPT. Lo he utilizado para pruebas de concepto y para conocer la solución, más que para trabajo *productivo* en sí. En cualquier caso, con muy buenas sensaciones.

Aunque ChatGPT me ha ayudado mucho, mi mayor asombro fue probando (casi sin quererlo) *Notion AI*. A partir de un documento en esta plataforma, y seleccionando unas pocas palabras, *Notion AI* pudo generar un texto que era aparentemente correcto y coherente:



1.3. ¿Has sido capaz de detectarlo?

La sección *Introducción y motivación I* de este documento ha sido generada utilizando Notion AI. Sí, lo ha hecho un ordenador. Esta herramienta únicamente ha necesitado la siguiente oración como *input* para generar los cuatro párrafos: *Qué es la IA y sus implicaciones éticas*.

Como se ha podido comprobar con este pequeño ejemplo, las implicaciones de esta nueva tecnología son realmente importantes.

2. ¿Qué es y qué no es la IA?

2.1. Definiciones de Inteligencia Artificial

El significado de *Inteligencia Artificial* para la Real Academia Española (RAE) es *Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico*. (RAE, Real Academia Española, 2023)

Para el Parlamento Europeo, la IA es *el presente y el futuro de la tecnología* y lo define como *la habilidad de una máquina de presentar las mismas capacidades que los seres humanos, como el razonamiento, el aprendizaje, la creatividad y la capacidad de planear*. (Parlamento Europeo, 2020)

Para John McCarthy, tal y como recoge la empresa IBM, *es la ciencia y la ingeniería de crear máquinas inteligentes, especialmente programas informáticos inteligentes. Está relacionada con la tarea similar de utilizar ordenadores para comprender la inteligencia humana, pero la IA no se limita a métodos que sean observables biológicamente*. (Learn Hub IBM, 2023)

Ahora que ya sabemos qué es la IA, veamos algo incluso más importante: qué **no** es: la IA no es una forma de inteligencia humana completa. Aunque los sistemas de IA pueden ser muy sofisticados y capaces de realizar tareas complejas, todavía no pueden igualar la inteligencia humana en todas las áreas. La IA también es diferente de la automatización, que se refiere a la capacidad de las máquinas para realizar tareas repetitivas de manera eficiente sin la necesidad de intervención humana.

2.2. Breve historia de la Inteligencia Artificial

Tal y como arroja la obra de de Crosas y Mora, la Inteligencia Artificial se remota a 1956. En este año John McCarthy ofreció su visión sobre la IA en la Universidad de Darmouth. (Wikipedia, 2023a) (Crosas Batista & Mora Ayala, 2022)

Para esa época la Informática no era siquiera parecida a lo que hoy conocemos. No es hasta el año 1997 cuando una máquina de IBM, conocida como *Deep Blue* consiguió ganar al ajedrez al entonces vigente campeón (Wikipedia, 2023b), Gary Kasparov. La prensa española se hizo eco de este hecho. (Archivo ABC, 1996)

En 2011, IBM creó Watson. Este sistema fue capaz de ganar a todos los humanos en el concurso televisivo Jeopardy como si fuera un participante más. La clave y la auténtica importancia de este hecho no fue la capacidad de búsqueda de los datos, sino la capacidad de entender las preguntas en lenguaje natural. (Crosas Batista & Mora Ayala, 2022)

2.3. Qué es la Inteligencia Artificial

- La IA no es un robot, un ordenador o algo similar. Es un área de las ciencias de la computación. Es decir, un campo de investigación y aplicación.
- El punto clave para distinguir un agente de un programa normal (como por ejemplo una hoja de cálculo, o la aplicación para acceder a una red social) es que intenta simular procesos asociados a la inteligencia humana. Cosas tan sencillas como distinguir objetos de una foto, leer una carta, reconocer a una persona, identificar el sarcasmo, etc. Estos procesos típicamente humanos son muy complicados para los ordenadores. Cuando un agente es capaz de simular alguno de ellos, se dice que está asociado con la IA.

Es decir, tenemos programas de ordenador normales y programas de ordenador super (o agentes) que son capaces de imitar a un ser humano (o casi). (Crosas Batista & Mora Ayala, 2022)

Extremadamente útil resulta la *Prueba de Turing*. Una IA (o generalizando, una máquina) pasará esta prueba cuando una persona no pueda diferenciar si está hablando (siempre de forma escrita) con una persona o con una máquina. (Wikipedia, 2023c)

2.4. Tipos/Niveles de IA

Aunque existen multitud de clasificaciones de IA. Existe cierta convención en la que aquí se refleja:

Nivel 1: inteligencia artificial débil. En este nivel el agente es capaz de realizar una única tarea sencilla (y muy concreta). Es importante reseñar que este es el nivel que actualmente tenemos de IA. (Crosas Batista & Mora Ayala, 2022).

También es denominada como *IA estrecha*. (Microsoft, 2023)

Nivel 2: inteligencia artificial general. Un agente en este nivel sería capaz de pasar la prueba de Turing. (Crosas Batista & Mora Ayala, 2022)

En teoría, un sistema informático que ha conseguido inteligencia artificial general podría resolver problemas sumamente complejos, emitir juicios en situaciones inciertas e incorporar conocimientos previos a su razonamiento actual. Podría tener creatividad e imaginación a la par que las personas y podría realizar muchísimas más tareas que la inteligencia artificial estrecha. (Microsoft, 2023)

Nivel 3: inteligencia artificial fuerte. Un agente en este nivel sería capaz de tener consciencia de sí mismo. (Crosas Batista & Mora Ayala, 2022)

Un sistema informático que haya logrado una superinteligencia artificial podría superar a las personas en casi todos los campos, incluidos el conocimiento en general, la creatividad científica y las habilidades sociales. (Microsoft, 2023)

3. Legislación y acuerdos supranacionales sobre IA

3.1. España

La mayor actuación de nuestro país de cara a la ordenación y promoción de la Inteligencia Artificial es la llamada **ENIA**, Estrategia Nacional de Inteligencia Artificial. Las actuaciones de esta estrategia se centran en seis ejes: la innovación en Inteligencia Artificial, el desarrollo de infraestructuras, el impulso del talento nacional, la integración de la IA en la cadena de valor, su uso de en la Administración y el impulso de un marco ético. Está dotada con 600 millones para el periodo 2021-2023. (Gobierno de España - PAE CTT, 2020)

En dicha estrategia se asuma la IA como *una de las tecnologías con mayor potencial de transformación e impacto en todas las áreas de actividad productiva, además de suponer un motor de innovación y un importante vector de generación de empleo de calidad*. (Gobierno de España - PAE CTT, 2020)

Pone sobre la mesa el dato de que la IA aportó 1.760 millones de euros al PIB mundial en 2018 y se estima que su contribución superará los 14 billones de euros para el año 2030. (Ministerio de Asuntos Económicos y Transformación Digital, 2020) (Ministerio de Asuntos Económicos y Transformación Digital, 2021)

Como se comenta en la subsección siguiente, relativa a la ordenación de la IA por parte de los organismos de la Unión Europea, los organismos de la Unión están trabajando para regular las implicaciones de esta nueva tecnología. Una de estas acciones consiste en la creación de un *sandbox regulatorio* para *acercar a las autoridades competentes a las empresas que desarrollan IA para definir las mejores prácticas que guiarán la implementación del futuro Reglamento de IA de la Comisión Europea (Artificial Intelligence Act)*. (Comisión Europea, 2022)

La creación de este entorno está liderada por el Gobierno de España, como una forma de materializar su compromiso e interés en la Inteligencia Artificial. (Gobierno de España - PAE CTT, 2021)

De forma adicional a las iniciativas europeas como el *Sandbox* regulatorio, y la ENIA, el Gobierno de España constituyó el Consejo Asesor de Inteligencia Artificial. (Gobierno de España, 2020) Este consejo asesor ha sufrido *bajas* precisamente por la ética: algunos expertos han abandonado este Consejo como protesta por el convenio de colaboración científica firmado entre España y Emiratos Árabes Unidos. (Pascual, 2023)

3.2. Unión Europea

El control y la creación de legislación en relación a la Inteligencia artificial es un tema clave para los legisladores. Tanto es así que la Unión Europea creó un comité especializado ya en 2020. (Prensa Parlamento Europeo, 2020). Este comité recibe el nombre de *AIDA*; del inglés, *Artificial Intelligence in a Digital Age*.

Este comité ha **clasificado** las Inteligencias Artificiales según su **riesgo**, dividiéndolas en cuatro grupos: (Comisión Europea, 2020)

- **Riesgo inadmisibles:** Se prohibirá un conjunto muy limitado de usos especialmente nocivos de la IA que contravienen los valores de la Unión al violar los derechos fundamentales (por ejemplo, puntuación social por parte de los Gobiernos [...]).
- **Alto riesgo:** tienen un impacto negativo en la seguridad de las personas o en sus derechos fundamentales.
- **Riesgo limitado:** En el caso de determinados sistemas de IA se imponen obligaciones específicas de transparencia, por ejemplo, cuando exista un riesgo claro de manipulación (por ejemplo, mediante

el uso de robots conversacionales). Los usuarios deben ser conscientes de que están interactuando con una máquina.

- **Riesgo mínimo:** Todos los demás sistemas de IA pueden desarrollarse y utilizarse con arreglo a la legislación vigente sin obligaciones jurídicas adicionales. La inmensa mayoría de los sistemas de IA utilizados actualmente en la UE pertenece a esta categoría. De forma voluntaria, los proveedores de estos sistemas pueden optar por aplicar los requisitos de una IA digna de confianza y adherirse a códigos de conducta voluntarios.

La **infracción** por parte de los sistemas de IA, en caso de no respetar los requisitos del reglamento propuesto, los Estados miembros pueden sancionar con hasta treinta millones de euros o el 6 % del volumen de negocios anual total a escala mundial del ejercicio financiero anterior, optándose por la de mayor cuantía. *La UE se lo toma en serio.*

Todo lo expuesto está recogido en la *Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO POR EL QUE SE ESTABLECEN NORMAS ARMONIZADAS EN MATERIA DE INTELIGENCIA ARTIFICIAL (LEY DE INTELIGENCIA ARTIFICIAL) Y SE MODIFICAN DETERMINADOS ACTOS LEGISLATIVOS DE LA UNIÓN*, disponible de forma íntegra en la web EUR-Lex. (Parlamento Europeo y Consejo, 2021)

3.3. Naciones Unidas

En noviembre de 2021, los 193 estados miembros de la UNESCO (organismo dependiente de la ONU), aprobaron el primer marco ético sobre la inteligencia artificial. En palabras de la propia ONU: *Este marco ético define valores y principios comunes que guiarán y garantizarán un desarrollo saludable de esta tecnología. Afirma que todos los individuos deberían poder acceder a sus registros de datos personales o incluso borrarlos y prohíbe explícitamente el uso de sistemas de inteligencia artificial para la calificación social y la vigilancia masiva.* (Naciones Unidas, 2021)

La ONU reconoce la relevancia (*[...] es omnipresente*) y la utilidad (*[...] hace posibles muchas de nuestras rutinas[...] dando resultados notables[...]*) de la inteligencia artificial; así como las implicaciones y retos que esta nueva tecnología está suponiendo (*[...]también está trayendo consigo nuevos retos sin precedentes*). (Naciones Unidas, 2021)

Las recomendaciones tiene por objetivo ayudar a que la IA no suponga riesgo para la sociedad y contribuya a la consecución de los Objetivos de Desarrollo Sostenible.

En tanto a la **protección de datos**, esta recomendación aboga por asegurar la transparencia, la capacidad de actuar y el control de los datos personales por parte de los interesados. Afirma que todos los individuos deberían poder acceder a sus registros de datos personales o incluso borrarlos.

Respecto a la **prohibición de los marcadores sociales** y la vigilancia masiva, prohíbe explícitamente el uso de sistemas de inteligencia artificial para la calificación social y la vigilancia masiva. En este mismo sentido apunta la recomendación de la Comisión Europea comentada anteriormente en lo que califica como *riesgo inadmisibile*.

Por último, se reconoce que las implicaciones de la IA, así como su funcionamiento, deben ser fáciles de **aplicar y evaluar** en tanto al impacto en las personas, la sociedad y el medio ambiente. (UNESCO, 2022)

4. Ética e IA: entrenamiento de modelos

En el ya comentado reglamento propuesto por la Comisión Europea para la regulación de la Inteligencia Artificial se incluyen las siguientes directrices para evitar sesgos raciales y de género en la IA, con especial importancia en su entrenamiento: *los sistemas de alto riesgo tendrán que entrenarse y ensayarse con conjuntos de datos lo suficientemente representativos como para reducir al mínimo el riesgo de sesgos injustos incorporados al modelo y velar por que puedan solventarse mediante las medidas apropiadas de detección y corrección de sesgos y otras medidas paliativas.*

Uno de estos casos en el que el entrenamiento supuso un posterior sesgo negativo de género hacia las mujeres lo recoge la publicación de Harvard Business Review, en el que se expone que el equipo de ingeniería de Amazon estuvo trabajando durante años en un programa que usaba Inteligencia Artificial para los procesos de contratación de nuevo personal. Finalmente, tuvieron que descartar su uso al no conseguir que esta no discriminase de forma sistemática a las mujeres. (Reuters, 2018) (Blackman, 2020)

Esta problemática está recogida en numerosas publicaciones científicas, algunos profesionales apuestan por dar visibilidad, de principio a fin, sobre los datos que son usados para entrenar los modelos de IA; así como facilitar que la información de entrenamiento sea lo más variada posible y no contenga sesgos previos. (Borenstein & Howard, 2021)

La empresa OpenAI, conocida por su producto *ChatGPT* también ha trabajado en IAs de generación de imágenes. Productos conocidos como *texto-to-image*. En un intento previo de evitar sesgos se encontró con el problema contrario: filtrar demasiado o de forma incorrecta puede introducir sesgos no contemplados inicialmente. *Los modelos generativos intentan ajustarse a la distribución de sus datos de entrenamiento, incluidos los sesgos que puedan contener. En consecuencia, filtrar los datos de entrenamiento puede crear o amplificar sesgos en los modelos posteriores. En general, corregir los sesgos del conjunto de datos original es una difícil tarea sociotécnica que seguimos estudiando y que queda fuera del alcance de este artículo. El problema que abordamos aquí es la amplificación de los sesgos causados específicamente por el propio filtrado de datos. Con nuestro enfoque, pretendemos evitar que el modelo filtrado esté más sesgado que el modelo sin filtrar, reduciendo esencialmente el desplazamiento de la distribución causado por el filtrado de datos.* (OpenAI & Perrogo, 2022)

5. Ética e IA: supervisión y corrección

En esta sección me centraré, no tanto en la supervisión del uso que los usuarios hacen de la IA (*user-side*), sino en qué ocurre cuando los modelos devuelven contenido violento, sexista y/o racista y cómo evitar que ocurra. En tanto a la moderación del contenido en el lado del usuario, las empresas disponen de sistemas para este fin y existen numerosos trabajos científicos publicados. Se debe tener en cuenta que la moderación de contenido, en lo relativo a texto, es algo que lleva siendo necesario desde el comienzo de Internet. (Todor Markov et al., 2022) (Röttger et al., 2021) (Markov et al., 2023)

El mayor problema que se presenta es el aumento de la complejidad en la detección de contenido peligroso. Es relativamente sencillo cuando se trata de texto, sobre simplificando: *matar* es malo, *amor* es bueno. Pero, ¿qué ocurre cuando este contenido peligroso no está etiquetado y hay que detectarlo y evitarlo? Hay que etiquetarlo.

Y sí, lo ha de etiquetar una persona. Esta información la recoge la revista TIME en un artículo reciente. Para obtener esos marcadores, OpenAI envió miles de fragmentos de texto a una empresa de externalización de servicios en Kenya. *Gran parte de ese texto parecía sacado de los recovecos más oscuros de Internet. Algunos describían situaciones con detalles gráficos como abusos sexuales a menores, zoofilia, asesinatos, suicidios, torturas, autolesiones e incesto..* (Times, 2023)

Esta empresa, Sama, pagaba entre 1.32\$ y 2\$, dependiendo del nivel y rendimiento de los empleados. Como referencia, el salario mínimo por hora en España es de 8,45€. Teniendo en cuenta el cambio de USD a EUR, estaríamos hablando de sueldos hasta 8 veces más bajos que en España. (Aradas, 2023)

El artículo de TIME cuenta que un trabajador de Sama encargado de leer y etiquetar texto para OpenAI contó a TIME que sufrió visiones recurrentes tras leer una descripción gráfica de un hombre manteniendo relaciones sexuales con un perro en presencia de un niño pequeño. *Fue una tortura*, dijo. *A lo largo de la semana se leen varias declaraciones como esa. Para cuando llega el viernes, estás perturbado de pensar en esa imagen.* (Times, 2023)

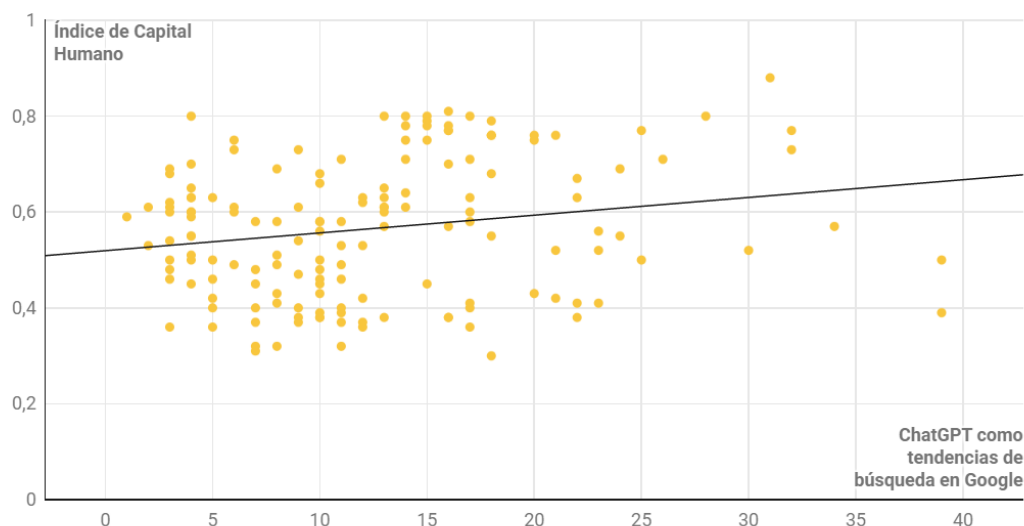
Menos ético se torna el caso cuando los contratos establecían que OpenAI pagaría una tarifa horaria de 12,50 dólares a Sama por el trabajo, lo que suponía entre seis y nueve veces la cantidad que los empleados de Sama cobraban.

6. Ética e IA: Brecha de acceso

Se define la brecha de acceso como *la imposibilidad que tienen algunos grupos de acceder a las TIC. Si bien las causas son diversas, suelen estar relacionadas por las diferencias socioeconómicas que existen entre grupos de población o entre países.* (Becas Santander, 2021)

Cada vez son más las voces que apuntan que el aumento en la implantación de la IA creará una nueva brecha digital. (McSherry, 2018) Este aumento en el uso supone una gran ventaja para las empresas, personas y estados que se pueden permitir usarlos, pero supone un aumento en la desigualdad con respecto a los que no se pueden permitir. (Lombana, 2018) (Internet Society, 2017) (Dans, 2016)

Utilizando Google Trends, relacionándolos con los datos del Índice de Capital Humano que ofrece el Banco Mundial, se puede observar como los países con menor índice tienen, de forma general, menores búsquedas para *ChatGPT*. (Stefan Pahl, 2023)



Por tanto, se vuelve a ver como la tecnología sí distingue por clases sociales y localización de los usuarios: las personas que no puedan acceder a estas tecnologías estarán en una posición inferior respecto a los que sí puedan acceder.

7. Ética e IA: Impacto climático

Como en este sentido no hay un acuerdo suficiente para generalizar, se ofrecerán los dos puntos de vista.

7.1. La IA es positiva

Por un lado, diversos expertos y consultoras apuntan a que la Inteligencia Artificial puede ser *esencial* para solucionar la crisis climática. (Maher et al., 2022) España es, de hecho, uno de los países mejor posicionados en este sentido según el informe de *Digital Future Society* gracias a su sector público, su experiencia y talento, la posibilidad de financiación y las características del sector privado.

La UIT, Unión Internacional de Telecomunicaciones, organismo dependiente de las Naciones Unidas ofrece 7 ejemplos de como la IA puede ayudar a combatir el cambio climático: (Kalejs, 2022)

- Tecnologías ecológicas de refrigeración y reducción de residuos electrónicos.
- Prevención de la contaminación por plásticos.
- Agricultura digital ecológica.
- Prevención de incendios forestales.
- Ciudades inteligentes y sostenibles.
- Imágenes por satélite para predecir la deforestación.
- Mejora de la eficiencia energética.

7.2. La IA es negativa

MIT Technology Review informó de que el entrenamiento de un solo modelo de IA puede emitir más de 626.000 libras equivalentes de dióxido de carbono. (Marr, 2023) También hay desarrolladores trabajando en herramientas para poder calcular cuánto dióxido de carbono emite el entrenamiento de un modelo de *Machine Learning*. (Schmidt et al., 2019) (Karen Hao, 2019)

Common carbon footprint benchmarks

in lbs of CO2 equivalent

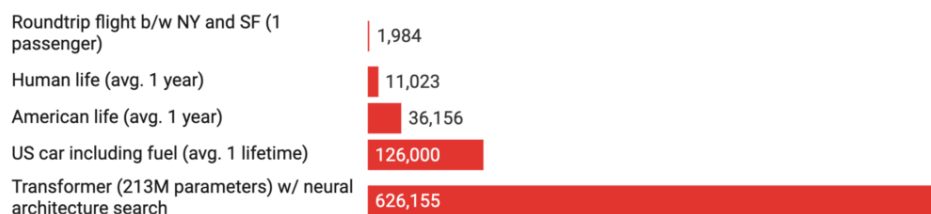


Chart: MIT Technology Review • Source: Strubell et al. • Created with Datawrapper

En un blog técnico de Microsoft se indicaba que la Inteligencia Artificial aún no está cerca de ser inteligente en ningún sentido significativo de la palabra. Los sistemas de *Machine Learning* aprenden a realizar una tarea específica observando montones de ejemplos y comparando patrones de datos mediante lo que puede considerarse un enfoque computacional de fuerza bruta. (Buchanan, 2020)

Esto requiere cantidades asombrosas de energía (datos y recursos informáticos) para realizar la comparación de patrones y un análisis estadístico sobrehumano. Como resultado, los enfoques más avanzados de la Inteligencia Artificial están inmersos en una carrera de armamento computacional para alcanzar el siguiente rendimiento de referencia. Un estudio de OpenAI (cuando todavía no era conocida) reveló que la computación utilizada en varios modelos de entrenamiento de IA de gran tamaño se había duplicado cada 3,4 meses desde 2012 -una desviación salvaje de la Ley de Moore, que la sitúa en 18 meses-, lo que supone un aumento de 300.000 veces. Este estudio es del año 2018, por lo que las implicaciones de estos datos 5 años después pueden ser aún mayores. (Amodei & Hernandez, 2018)

Ya existen expertos apostando por una *IA Verde* en todos los ámbitos posibles: desde el tamaño de los modelos y el entrenamiento, hasta donde están físicamente los servidores que realizan las tareas. (Andrews, 2020) Para ayudar a valorar el impacto que un proyecto de *Machine Learning* tendrá, Stanford ha desarrollado una herramienta sencilla bajo la premisa de que «no puedes resolver un problema si no lo puedes medir». (Henderson et al., 2022)

8. Ética e IA: ¿Revolución o conflicto educativo?

Que la inteligencia artificial se está haciendo un hueco en el mundo educativo es un hecho indiscutible. La pregunta, por tanto, no es si se pueden evitar sus *efectos* en el sector educativo, sino cómo usarlo de forma productiva y segura. Según un informe de *Impact Research*, *un tercio de los estudiantes estadounidenses de entre 12 y 17 años, y el 51 % de los docentes, utilizan ya esta herramienta, ya sea para planificar sus clases, encontrar actividades creativas o elaborar un conocimiento de base con el que empezar sus lecciones. Sus conclusiones no pueden ser mejores: el 89 % de los enseñantes y el 79 % de los alumnos que han usado ChatGPT piensan que el impacto ha sido francamente positivo.* (Research, 2023) (Meneses, 2023).

Para Guillem García Brustenga, experto del eLearning Innovation Center de la UOC, es necesario un cambio de metodología: contextualización, reflexión, reinención del rol del profesorado.

Si bien la abundancia de conocimiento ya estaba disponible y fácilmente accesible con la llegada de Internet, con ChatGPT (y, en general, las soluciones de IA *generativas*) es más sencillo porque el estudiante no tiene siquiera que modificar el contenido encontrado o saber hacer las búsquedas correctas. Basta con copiar las instrucciones facilitadas por el/la docente y ¿tarea hecha?

No. La tarea puede no estar hecha, puede no estar bien hecha. Joaquín Rodríguez, director de Diseño, Innovación y Tecnologías para el Aprendizaje de la Institución Educativa SEK, advierte que “uno de los errores de ChatGPT es que tiene lo que se conoce como “alucinaciones”: tiene siempre pretensiones de verdad, por lo que, si no sabe algo, se lo inventa”. (Meneses, 2023)

Para Joel Manuel Prieto Andreu, de UNIR, *La inteligencia artificial en educación puede analizar los datos del alumnado y adaptar los contenidos educativos para satisfacer las necesidades individuales de cada estudiante, evaluar el aprendizaje de los estudiantes en tiempo real y proporcionar retroalimentación inmediata, lo que permite al profesorado intervenir y corregir los errores de los estudiantes de manera oportuna.*

Sin embargo, la implementación efectiva de la IA en la educación requiere una planificación cuidadosa y un enfoque ético. Es importante que los educadores y los responsables políticos trabajen juntos para garantizar que esta herramienta se utilice de manera responsable y para mejorar la calidad de la educación. (Prieto Andreu, 2023)

9. Conclusión

En vista de lo expuesto en el presente trabajo, podemos concluir que sí, que la Inteligencia Artificial puede ser realmente positiva en muchos aspectos. También cabe asumir que la extensión de esta tecnología trae cambios estructurales que habrá que asumir y a los que la sociedad tendrá que adaptarse. Sí, entre ellos la pérdida de muchos puestos de trabajo. (Eloundou et al., 2023).

La clave está en cómo, de forma conjunta como sociedad y desde todos los poderes, afrontamos estos cambios. Estos cambios deben ser abordados de forma ética y cuando sean generadores de desigualdad es necesaria la aplicación de políticas para su corrección. Legislaciones valientes y tempranas son imprescindibles para prevenir problemas en un futuro próximo. A su vez, es necesario que expertos y expertas en ética supervisen, de principio a fin, el ciclo de vida de las aplicaciones y modelos de Inteligencia Artificial. Y cuando haya trabajos especialmente duros, se limite la exposición de las personas y se paguen de forma digna.

Respecto a la brecha de acceso, es necesario limitar la brecha de acceso ya existente respecto a Internet en general, y en concreto para las IA, respecto a los modelos (Workshop et al., 2023)

Como conclusión, como para todos los cambios trascendentales, habrá efectos negativos. La clave está en cómo se afrontan y cómo se evita que estos efectos lastren, como siempre, a los sectores ya vulnerables de la población.

10. Bibliografía

- Amodei, D., & Hernandez, D. (2018, 16 de mayo). *AI and compute*. <https://openai.com/research/ai-and-compute> (Consultado 2023-04-05)
- Andrews, E. L. (2020, 2 de julio). *AI's carbon footprint problem* [Stanford HAI]. <https://hai.stanford.edu/news/ais-carbon-footprint-problem> (Consultado 2023-04-05)
- Aradas, A. (2023, 29 de marzo). *Cuánta del SMI por hora, día, mes y anual / 2023* [Section: Condiciones laborales]. <https://www.cuestioneslaborales.es/cuanta-del-smi/> (Consultado 2023-04-05)
- Archivo ABC. (1996). *ABC MADRID 11-02-1996 página 94*. <https://www.abc.es/archivo/periodicos/abc-madrid-19960211-94.html> (Consultado: 07/03/2023)
- Becas Santander. (2021, 29 de noviembre). *La brecha digital: qué es y cómo podemos reducirla*. <https://www.becas-santander.com/es/blog/brecha-digital-que-es.html> (Consultado 2023-04-05)
- Blackman, R. (2020). A Practical Guide to Building Ethical AI. *Harvard Business Review*. <https://hbr.org/2020/10/a-practical-guide-to-building-ethical-ai>. (Consultado: 05/04/2023)
- Borenstein, J., & Howard, A. (2021). Emerging challenges in AI and the need for AI ethics education. *AI and Ethics*, 1(1), 61-65. <https://doi.org/10.1007/s43681-020-00002-7>. (Consultado: 05/04/2023)
- Buchanan, W. (2020, 26 de octubre). *The carbon footprint of AI* [Sustainable software]. <https://devblogs.microsoft.com/sustainable-software/the-carbon-footprint-of-ai/> (Consultado 2023-04-05)
- Comisión Europea. (2020). *FAQ Nuevas normas sobre la IA*. https://ec.europa.eu/commission/presscorner/api/files/document/print/es/qanda_21_1683/QANDA_21_1683_ES.pdf (Consultado: 05/04/2023)
- Comisión Europea. (2022). *First regulatory sandbox on Artificial Intelligence presented*. <https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented> (Consultado: 05/04/2023)
- Crosas Batista, M., & Mora Ayala, E. (2022). Qué es la inteligencia artificial? Entendamos primero lo que es (pero sobre todo lo que no es). En *La Era de Los Asistentes Conversacionales*. Editorial UOC.
- Dans, E. (2016, 22 de noviembre). *Artificial intelligence is the new digital divide* [Enrique dans]. <https://medium.com/enrique-dans/artificial-intelligence-is-the-new-digital-divide-736cd52fe906> (Consultado 2023-04-05)
- Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023, 23 de marzo). GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. <https://doi.org/10.48550/arXiv.2303.10130> (Consultado 2023-04-09)
- Gobierno de España. (2020, 20 de julio). *El Gobierno constituye el Consejo Asesor de Inteligencia Artificial*. <https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/asuntos-economicos/Paginas/2020/200720-inteligencia.aspx> (Consultado 2023-04-05)
- Gobierno de España - PAE CTT. (2020). *Presentación de la Estrategia Nacional de Inteligencia Artificial (ENIA) para el periodo 2021-2023*. https://administracionelectronica.gob.es/pae/Home/pae_Actualidad/pae_Noticias/Anio2020/Diciembre/Noticia-2020-12-04-presenta-Estrategia-Nacional-Inteligencia-Artificial-2021-2023.html (Consultado: 05/04/2023)
- Gobierno de España - PAE CTT. (2021). *Presentado el primer marco normativo de Inteligencia Artificial*. https://administracionelectronica.gob.es/pae/Home/pae_Actualidad/pae_Noticias/Anio2022/Junio/Noticia-2022-06-28-Presentado-primer-marco-normativo-Inteligencia-Artificial.html (Consultado: 05/04/2023)
- Google. (2023). *Resultados de Google Trends*. <https://trends.google.com/trends/explore?q=%5C%2Fm%5C%2F0mkz,ChatGPT&hl=es> (consultado: 07/03/2023)

- Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., & Pineau, J. (2022, 29 de noviembre). Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning. <https://doi.org/10.48550/arXiv.2002.05651> (Consultado 2023-04-05)
- Internet Society. (2017). *Global internet report 2017: Paths to our digital future* [Internet society]. <https://www.internetsociety.org/resources/doc/2017/global-internet-report-2017/> (Consultado 2023-04-05)
- Kalejs, E. (2022, 10 de noviembre). *7 AI innovations helping to combat climate change* [AI for good]. <https://aiforgood.itu.int/7-ai-innovations-helping-to-combat-climate-change/> (Consultado 2023-04-05)
- Karen Hao. (2019, 6 de junio). *Training a single AI model can emit as much carbon as five cars in their lifetimes* [MIT technology review]. <https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/> (Consultado 2023-04-05)
- Learn Hub IBM. (2023). *Inteligencia artificial (IA)*. <https://www.ibm.com/es-es/cloud/learn/what-is-artificial-intelligence> (Consultado: 07/03/2023)
- Lombana, A. (2018, 19 de septiembre). *La evolución de las brechas digitales*. https://scholar.googleusercontent.com/scholar?q=cache:I_F_VyN1G0kJ:scholar.google.com/+Brecha+de+acceso+Inteligencia+Artificial&hl=es&as_sdt=0,5 (Consultado 2023-04-05)
- Maher, H., Meinecke, H., Gromier, D., Garcia-Novelli, M., & Fortmann, R. (2022, 4 de julio). *AI is essential for solving the climate crisis* [BCG global]. <https://www.bcg.com/publications/2022/how-ai-can-help-climate-change> (Consultado 2023-04-05)
- Markov, T., Zhang, C., Agarwal, S., Eloundou, T., Lee, T., Adler, S., Jiang, A., & Weng, L. (2023, 14 de febrero). A Holistic Approach to Undesired Content Detection in the Real World. <https://doi.org/10.48550/arXiv.2208.03274> (Consultado 2023-04-05)
- Marr, B. (2023, 22 de marzo). *Green intelligence: Why data and AI must become more sustainable* [Forbes] [Section: Enterprise Tech]. <https://www.forbes.com/sites/bernardmarr/2023/03/22/green-intelligence-why-data-and-ai-must-become-more-sustainable/> (Consultado 2023-04-05)
- McSherry, M. (2018, 21 de mayo). *Will AI widen or weaken the global digital divide?* [Ada-AI]. <https://medium.com/ada-ai/will-ai-widen-or-weaken-the-global-digital-divide-a0dc6785e8a> (Consultado 2023-04-05)
- Meneses, N. (2023, 30 de marzo). *ChatGPT y educación: ¿un nuevo enemigo o aliado de los profesores?* [El País] [Section: Formación]. <https://elpais.com/economia/formacion/2023-03-30/chatgpt-y-educacion-un-nuevo-enemigo-o-aliado-de-los-profesores.html> (Consultado 2023-04-08)
- Microsoft. (2023). *¿Qué es la Inteligencia Artificial?* <https://azure.microsoft.com/es-es/resources/cloud-computing-dictionary/what-is-artificial-intelligence/#types> (Consultado: 07/03/2023)
- Ministerio de Asuntos Económicos y Transformación Digital. (2020). *Estrategia ENIA - Resumen ejecutivo*. https://administracionelectronica.gob.es/pae/Home/pae_Actualidad/pae_Noticias/Anio2020/Diciembre/Noticia-2020-12-04-presenta-Estrategia-Nacional-Inteligencia-Artificial-2021-2023.html (Consultado: 05/04/2023)
- Ministerio de Asuntos Económicos y Transformación Digital. (2021). *Presentación Estrategia ENIA - Nota de prensa*. https://portal.mineco.gob.es/es-es/comunicacion/Paginas/201202_np_enia.aspx (Consultado: 05/04/2023)
- Naciones Unidas. (2021). *Primer acuerdo mundial sobre la ética de la inteligencia artificial*. <https://news.un.org/es/story/2021/11/1500522> (Consultado: 05/04/2023)
- OpenAI & Perrogo, B. (2022, 28 de junio). *DALL·e 2 pre-training mitigations*. <https://openai.com/research/dall-e-2-pre-training-mitigations> (Consultado 2023-04-05)

- Parlamento Europeo. (2020). *¿Qué es la inteligencia artificial y cómo se usa?* <https://www.europarl.europa.eu/news/es/headlines/society/20200827STO85804/que-es-la-inteligencia-artificial-y-como-se-usa> (Consultado: 07/03/2023)
- Parlamento Europeo y Consejo. (2021). *Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO POR EL QUE SE ESTABLECEN NORMAS ARMONIZADAS EN MATERIA DE INTELIGENCIA ARTIFICIAL (LEY DE INTELIGENCIA ARTIFICIAL) Y SE MODIFICAN DETERMINADOS ACTOS LEGISLATIVOS DE LA UNIÓN*. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206> (Consultado: 05/04/2023)
- Pascual, M. G. (2023, 28 de marzo). *Un grupo de expertos en inteligencia artificial rompe con el Gobierno por discrepancias éticas* [El País] [Section: Tecnología]. <https://elpais.com/tecnologia/2023-03-28/un-grupo-de-expertos-en-inteligencia-artificial-rompe-con-el-gobierno-por-discrepancias-eticas.html> (Consultado 2023-04-05)
- Prensa Parlamento Europeo. (2020). *New committee on Artificial Intelligence begins its work* / Noticias / Parlamento Europeo. <https://www.europarl.europa.eu/news/es/press-room/20200923IPR87711/new-committee-on-artificial-intelligence-begins-its-work> (Consultado: 05/04/2023)
- Prieto Andreu, J. M. (2023, 16 de marzo). *La Inteligencia Artificial en Educación y la aplicabilidad de ChatGPT* [UNIR]. <https://www.unir.net/educacion/revista/inteligencia-artificial-educacion-chatgpt/> (Consultado 2023-04-08)
- RAE, Real Academia Española. (2023). *Significado de 'Inteligencia' [artificial]*. <https://dle.rae.es/inteligencia> (Consultado: 07/03/2023)
- Research, I. (2023, 1 de marzo). Teachers and Students Embrace ChatGPT for Education. <https://8ce82b94a8c4fdc3ea6d-b1d233e3bc3cb10858bea65ff05e18f2.ssl.cf2.rackcdn.com/ae/84/133976234126a2ad139411c1e770/impact-research-teachers-and-students-tech-poll-summary-memo.pdf> (Consultado 2023-04-08)
- Reuters. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. Consultado el 5 de abril de 2023, desde <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>. (Consultado: 05/04/2023)
- Röttger, P., Vidgen, B., Nguyen, D., Waseem, Z., Margetts, H., & Pierrehumbert, J. (2021). HateCheck: Functional tests for hate speech detection models. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 41-58. <https://doi.org/10.18653/v1/2021.acl-long.4>. (Consultado 2023-04-05)
- Schmidt, V., Luccioni, A., Lacoste, A., & Dandres, T. (2019). *Machine learning CO2 impact calculator*. <https://mlco2.github.io/impact> (Consultado 2023-04-05)
- Stefan Pahl. (2023, 16 de febrero). *Una brecha emergente: ¿Quién se beneficia de la IA?* [Industrial Analytics Platform]. <https://iap.unido.org/es/articles/una-brecha-emergente-quien-se-beneficia-de-la-ia> (Consultado 2023-04-05)
- Times, T. (2023, 18 de enero). *Exclusive: The \$2 per hour workers who made ChatGPT safer* [Time]. <https://time.com/6247678/openai-chatgpt-kenya-workers/> (Consultado 2023-04-05)
- Todor Markov, Chong Zhang, Sandhini Agarwal, Tyna Eloundou, Teddy Lee, Steven Adler, Angela Jiang & Lilian Weng. (2022, 10 de agosto). *New and improved content moderation tooling*. <https://openai.com/blog/new-and-improved-content-moderation-tooling> (Consultado 2023-04-05)
- UNESCO. (2022). *Recomendación sobre la ética de la inteligencia artificial*. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa (Consultado: 05/04/2023)
- Wikipedia. (2023a). *Conferencia de Dartmouth*. https://es.wikipedia.org/wiki/Conferencia_de_Dartmouth (Consultado: 07/03/2023)

- Wikipedia. (2023b). *Deep Blue (computadora)*. [https://es.wikipedia.org/wiki/Deep_Blue_\(computadora\)](https://es.wikipedia.org/wiki/Deep_Blue_(computadora)) (Consultado: 07/03/2023)
- Wikipedia. (2023c). *Prueba de Turing*. https://es.wikipedia.org/wiki/Prueba_de_Turing (Consultado: 07/03/2023)
- Workshop, B., Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., Castagné, R., Luccioni, A. S., Yvon, F., Gallé, M., Tow, J., Rush, A. M., Biderman, S., Webson, A., Ammanamanchi, P. S., Wang, T., Sagot, B., Muennighoff, N., ... Wolf, T. (2023, 13 de marzo). BLOOM: A 176B-Parameter Open-Access Multilingual Language Model. <https://doi.org/10.48550/arXiv.2211.05100> (Consultado 2023-04-09)