



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

ADE

Facultad de Administración
y Dirección de Empresas /UPV

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Facultad de Administración y Dirección de Empresas

Crowdsourcing y su implementación en empresas para
entrenar modelos de machine learning y algoritmos de
inteligencia artificial.

Trabajo Fin de Máster

Máster Universitario en Gestión de Empresas, Productos y
Servicios

AUTOR/A: Muñoz Gomez, Sebastian

Tutor/a: González Ladrón de Guevara, Fernando Raimundo

CURSO ACADÉMICO: 2023/2024

Índice contenido

1.	Abstract	5
2.	Palabras clave.....	5
3.	Capítulo 1: Introducción.....	6
	3.1. Introducción.....	7
	3.2. Objetivos.....	8
	3.3. Hipótesis	9
4.	Capítulo 2: Enfoque Metodológico	10
	4.1. Metodología.....	11
	4.2. Justificación.....	12
	4.3. Resultados esperados	13
5.	Capítulo 3: Marco teórico.....	14
	5.1. Inteligencia Artificial	15
	5.2. Aprendizaje automático	17
	5.3. Crowdsourcing.....	18
	5.4. Calidad de datos.....	21
	5.5. Colaboradores	23
	5.6. Política y gobernanza.....	25
6.	Capítulo 4: Análisis de plataformas.....	29

6.1. Amazon Mechanical Turk (MTurk).....	30
Experiencias exitosas con MTurk.....	31
6.2. Figure Eight.....	34
Experiencias exitosas con Figure Eight.....	35
6.3. Clickworker.....	37
Experiencias exitosas con Clickworker.....	38
6.4. Resumen análisis plataformas.....	39
6.5. Aplicación del crowdsourcing.....	40
7. Capítulo 5: Resultados.....	44
8. Capítulo 6: Conclusiones, limitaciones y futuras líneas de investigación.....	47
9. Referencias bibliográficas.....	50
10. Anexos.....	58
11. ANEXO: Objetivos de Desarrollo Sostenible (ODS).....	65

Índice tablas

- Tabla #1: Motivadores intrínsecos (pag. 24)
- Tabla #2: Motivadores extrínsecos (pag. 24)
- Tabla #3: Comparativa entre MTurk, Figure Eight y Clickworker (pag. 38)
- Tabla #4: Tabla ventajas y desventajas (pag. 40)

Índice anexos

- #1. Porcentaje de trabajadores en MTurk (2018) (pag. 58)
- #2. Satisfacción salarial entre MTurk y Población general USA (2023) (pag. 58)
- #3. Implementación de IA en diferentes países y áreas (pag. 59)
- #4. Comparativa salarial, opinión (pag. 59)
- #5. Definiciones de crowdsourcing (pag. 60)
- #6. Evolución modelos ML (pag. 60)
- #7. Modelos ML por país (pag. 61)
- #8. Patentes de IA (pag. 61)
- #9. IA y sus preocupaciones (pag. 62)
- #10. Regulaciones de IA en USA (pag. 62)
- #11. Regulaciones de IA en UE (pag. 63)
- #12. Beneficios del ML (pag. 63)
- #13. DataRobot y gobernanza (pag. 64)

Abstract

Este estudio examina la relevancia de las plataformas de crowdsourcing Amazon Mechanical Turk, Figure Eight y Clickworker, en la calidad y diversidad de los datos utilizados para entrenar modelos de Inteligencia Artificial (IA de ahora en adelante) y Aprendizaje Automático (del inglés Machine Learning, de ahora en adelante ML). Analiza cómo el crowdsourcing puede satisfacer estas necesidades, y tiene como principales objetivos evaluar estas plataformas en términos de sus beneficios, desventajas y aplicaciones para la IA y el ML, e identificar los desafíos presentes relacionados con los incentivos de los colaboradores de las plataformas, la calidad de los datos, y la gobernanza sobre estos. Los resultados de la investigación indicaron que las plataformas de crowdsourcing son esenciales para brindar soluciones de recopilación y anotación de datos escalables y rentables, permiten automatizar y optimizar tareas complejas en varias micro tareas, esto se evidencia con empresas como Autodesk que querían obtener datos completos de sus clientes, o la plataforma médica C-SATS de Johnson & Johnson que publica videos de cirugías para recibir comentarios y recomendaciones de varios médicos y expertos. Además, para obtener una estrategia exitosa depende de un equilibrio entre la eficiencia operativa y el cumplimiento de estándares éticos y de calidad. Este estudio contribuye al creciente desarrollo de tecnologías de IA y ML gracias al crowdsourcing, ofreciendo perspectivas para los que buscan aprovechar la innovación basada en datos.

Palabras clave

Crowdsourcing, inteligencia artificial, aprendizaje automático, calidad de datos, lenguajes de gran escala, algoritmos de aprendizaje.

Capítulo 1: Introducción

Introducción

La inteligencia artificial (IA), que es la ciencia e ingeniería de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes (McCarthy, 2007), abarca diversos enfoques, desde sistemas basados en reglas hasta aprendizaje profundo, cada uno con sus propias fortalezas y aplicaciones (Russell & Norvig, 2021). El ML, un conjunto de métodos que puede detectar automáticamente patrones en los datos y luego usar los patrones descubiertos para predecir datos futuros o para realizar otros tipos de toma de decisiones bajo incertidumbre (Murphy, 2012), empieza a ser relevante para que las empresas se mantengan competitivas.

El crowdsourcing, que se puede definir como la obtención de servicios, ideas o contenido mediante la solicitud de contribuciones de un grupo amplio de personas (Estelles & González-Ladrón-De-Guevara, 2012), se perfila como una estrategia efectiva para recolectar datos diversos y de alta calidad, esenciales para el entrenamiento de estos modelos. Su importancia en IA ha crecido exponencialmente, con un aumento del 300% en su uso para proyectos de ML entre 2018 y 2023 (LeCun et al., 2023). Esta estrategia en la IA y ML se utiliza para tareas como etiquetado de datos, validación de resultados de modelos y generación de conjuntos de datos diversos. Por ejemplo, en visión por computadora, plataformas de crowdsourcing se usan para anotar imágenes, mejorando la precisión de los modelos de reconocimiento de objetos (Kovashka et al., 2016).

Este trabajo analiza las plataformas de crowdsourcing Amazon Mechanical Turk, Figure Eight y Clickworker porque son opciones ideales para investigar la recopilación y depuración de datos para modelos de IA debido a su versatilidad y eficacia comprobada, un análisis de 1,000 artículos publicados en conferencias de IA entre 2015 y 2020 reveló que MTurk fue utilizado en el 45% de los estudios que emplearon crowdsourcing, seguido por Figure Eight (20%) y Clickworker (15%), justificando su selección para este estudio (Paolacci & Chandler, 2014).

Su popularidad en la comunidad científica facilita el acceso a una amplia base de conocimientos y mejores prácticas (Peer et al., 2017). La similitud en sus funcionalidades básicas permite una comparación efectiva, mientras que sus diferencias sutiles ofrecen perspectivas únicas sobre diversos enfoques de crowdsourcing (Difallah et al., 2018). La facilidad de uso de estas plataformas, combinada con sus robustas capacidades de control de calidad, las hace particularmente adecuadas para proyectos de investigación que requieren datos precisos y diversos (Vaughan, 2017). Además, su integración con herramientas de análisis de datos y aprendizaje automático simplifica el flujo de trabajo de la investigación, desde la recopilación de datos hasta el entrenamiento de modelos (Lease, 2011). La abundancia de literatura académica sobre estas plataformas proporciona un sólido fundamento teórico para futuras investigaciones, permitiendo a los investigadores construir sobre un cuerpo de conocimiento establecido y contribuir significativamente al campo del crowdsourcing para IA.

Objetivos

- Analizar las características de las plataformas de crowdsourcing Amazon Mechanical Turk, Figure Eight y Clickworker, enfocándose en los beneficios, funciones, desventajas, y aplicaciones que poseen en las industrias, a través de revisión de literatura, investigaciones pasadas, experiencias exitosas y benchmarking.
- Documentar los desafíos del crowdsourcing en el contexto de la inteligencia artificial y el aprendizaje automático, centrándose en la calidad de datos, la relevancia de tener políticas y gobernanzas, y la importancia de incentivos dignos por realizar las tareas.

Hipótesis

Este estudio plantea las siguientes hipótesis interrelacionadas sobre el impacto del crowdsourcing en la IA y el ML:

- El uso de plataformas de crowdsourcing como MTurk, Figure Eight y Clickworker aumenta la precisión de los modelos de IA y ML, esto se podrá observar por su amplio campo de aplicación en diferentes industrias soportado por experiencias exitosas de empresas que las utilizaron. Se espera un incremento en la optimización del tiempo, más satisfacción por la precisión en los outputs de los sistemas, reducción de costes en los procesos.
- La variabilidad en las remuneraciones, siendo un motivador principal, y las condiciones laborales de los colaboradores que participan en las campañas de crowdsourcing serán causas principales que pueden afectar la calidad de los datos que se obtienen, aumentando la probabilidad de crear sesgos en los datos de las campañas. Comparaciones entre aspectos de remuneración entre plataformas nos darán una idea de las diferencias y cuál beneficia más.
- Sistemas de IA y ML que se analicen y que no implementen una gobernanza y políticas adecuadas sobre el manejo de datos, revelarán brechas significativas en la protección de datos personales y privacidad, subrayando la necesidad de mejorar la transparencia y seguridad en el manejo de datos para las plataformas de crowdsourcing, AI y ML.

Capítulo 2: Enfoque Metodológico

Metodología

Se opta por un enfoque exploratorio de metodología cualitativa con el objetivo de identificar las funciones, ventajas y desafíos que presentan las plataformas de crowdsourcing dentro del ámbito de la inteligencia artificial y el aprendizaje automático. Se realizó una revisión de literatura académica de investigaciones sobre crowdsourcing y las plataformas seleccionadas, su desarrollo, sus aplicaciones en diferentes áreas y su evolución. Estas investigaciones se buscaron en bases de datos como Scopus e IEEE Xplore, utilizando palabras clave como: MTurk, Clickworker, Figure Eight, AI, Machine Learning, crowdsourcing, publicadas entre 2005 y 2024. Se aplicaron criterios de inclusión y exclusión, obteniendo alrededor de 500 artículos. También, para poder identificar y explicar los desafíos en el contexto de la IA y ML se examinaron informes anuales reconocidos en el mundo de instituciones como Stanford University Human-Centered Artificial Intelligence y McKinsey & Company, para evidenciar la evolución en la creación modelos de IA y ML, como la incertidumbre por la seguridad y privacidad en el manejo de datos personales crece, y diferentes acciones que se recomiendan para regular el origen y manejo de los datos.

Además, una revisión técnica de las diferentes plataformas para comparar sus funciones, alcances, y aplicaciones; se analizaron criterios como interfaces de usuarios, casos exitosos de uso publicados y sus resultados, encuestas de personas que fueron colaboradores. El objetivo de esta metodología no solo es explorar, sino además generar nuevos conocimientos teóricos y prácticos acerca de cómo implementar el crowdsourcing en el ámbito empresarial.

Justificación

La llegada de la inteligencia artificial a diversos sectores y sus múltiples aplicaciones están impulsando la competencia en los mercados y teniendo un efecto significativo en la economía global (Manyika & Sneider, 2018). Diferentes gobiernos han creado marcos oficiales relacionados con la IA con el fin de fomentar el desarrollo económico y tecnológico (para más detalles, véase el **Anexo 3**). Deloitte expresaba que cada vez más se tiene en cuenta las ventajas competitivas que genera la inteligencia artificial. Las organizaciones deben definir una estrategia, seleccionar las aplicaciones adecuadas, y sentar las bases para el uso de los datos adecuados. (Deloitte Spain, 2020).

El crowdsourcing emerge como una herramienta crucial para superar uno de los mayores desafíos en el desarrollo de la IA: la obtención de datos de alta calidad, entendiendo esta como la precisión, completitud y consistencia de los datos utilizados para entrenar modelos de IA (Batini & Scannapieco, 2016) y la diversidad de datos. Según un estudio de Accenture (2023), el 78% de las empresas que implementan IA exitosamente utilizan alguna forma de crowdsourcing para la recolección y etiquetado de datos. Un informe de Grand View Research (2024) proyecta que el mercado global de crowdsourcing para IA alcanzará los \$25 mil millones para 2028, con una tasa de crecimiento anual compuesta del 22% entre 2024 y 2028, subrayando la creciente importancia de esta intersección tecnológica. Sin embargo, el uso de crowdsourcing en IA también plantea desafíos éticos significativos, incluyendo preocupaciones sobre la privacidad de los datos, la justa compensación de los trabajadores y los sesgos potenciales en los datos recopilados (Whittaker et al., 2023)

Resultados esperados

- Se espera desarrollar un análisis detallado de Amazon Mechanical Turk, Figure Eight y Clickworker, evaluando sus características, funcionalidades, ventajas y desventajas en el contexto de proyectos de IA y ML. Este análisis incluirá métricas cualitativas sobre la eficacia de cada plataforma en tareas como etiquetado de datos, clasificación de texto y generación de conjuntos de datos para entrenamiento de modelos.
- Identificar los principales desafíos asociados con el uso del crowdsourcing en proyectos de IA y ML. Estos desafíos abarcarán aspectos como la calidad de los datos, la remuneración de los colaboradores, la transparencia en el manejo de datos y las consideraciones éticas. Se proporcionará un análisis de cómo estos desafíos impactan en el desarrollo y la precisión de los modelos de IA.
- Desarrollar un marco teórico que guíe la implementación efectiva de estrategias de crowdsourcing en proyectos de IA y ML. Este marco integrará los hallazgos del análisis de plataformas y los desafíos identificados, proporcionando pautas prácticas para empresas que busquen adoptar estas tecnologías.

Capítulo 3: Marco teórico.

Este capítulo presenta el marco teórico que sustenta la investigación sobre el uso de plataformas de crowdsourcing en Inteligencia Artificial y Aprendizaje Automático. Comenzaremos explorando los conceptos fundamentales de IA y ML, seguido de una discusión sobre el crowdsourcing y su relevancia en este contexto. Luego, abordaremos la importancia de la calidad de los datos, el papel de los colaboradores, y finalizaremos con una revisión de las políticas y gobernanza relacionadas con estas tecnologías.

Inteligencia Artificial

La aparición de nuevas tecnologías digitales soportadas por IA promete una revolución industrial en la que empresas optarán por estas tecnologías para transformar y optimizar sus procesos aumentando la rentabilidad y manteniendo la competitividad mediante digitalización y automatización de procesos. Stanford Institute HAI (Human-Centered Artificial Intelligence) (cap.4, 2024) indica lo siguiente:

“En 2023, las inversiones en IA en Estados Unidos alcanzaron los 67.200 millones de dólares, casi 8.7 veces más que China, el siguiente mayor inversor. Si bien la inversión privada en IA en China y la Unión Europea, incluido el Reino Unido, disminuyó un 44,2% y un 14,1%, respectivamente, desde 2022, Estados Unidos experimentó un notable aumento del 22,1% en el mismo período.”

Adicional a esto, una investigación realizada por McKinsey & Company (2024) indica que las aplicaciones de IA de última generación pueden sumar hasta 4,4 billones de dólares a la economía global, anualmente. Esto también se puede evidenciar por el número de patentes de IA que ha incrementado un 62.7% desde el 2010 (para más detalles, véase el **Anexo 8**). Estas tecnologías se

han convertido en pilares fundamentales para la toma de decisiones estratégicas, y la innovación. Como lo expresa Philippe Rambach, Chief AI Officer de Schneider Electric (2024):

“La IA representa una oportunidad y una perturbación masiva. Con su adopción sin precedentes, aumenta la productividad de las personas, máquinas y procesos con un inmenso potencial de ampliación. Su capacidad para extraer grandes cantidades de datos es clave para la transición energética: ninguna otra tecnología puede apoyarnos en un uso más inteligente de los escasos recursos mediante optimización incesante y continua.”

Pero eso sí, para que las transformaciones digitales y de inteligencia artificial tengan éxito, las empresas necesitan comprender los problemas que quieren resolver y reconfigurar sus organizaciones para la innovación continua (Lamarre, 2023). Lamarre también afirma que las empresas tradicionales que se toman en serio la transformación digital consiguen el talento adecuado para desarrollarlas, pero tienen que comprometerse con un entorno tecnológico moderno que facilite a los empleados hacer su trabajo, manteniendo la calidad de datos y proporcionando el software y las herramientas necesarias para gestionarlos.

Tavakoli (et al., 2024) comenta que la IA generativa, definida como algoritmos que se pueden utilizar para crear contenido nuevo, incluidos audio, código, imágenes, texto, simulaciones y videos (McKinsey & Company, 2024), es uno de los avances tecnológicos en la IA más prometedores, pero la gestión de datos sigue siendo uno de los principales obstáculos para crear valor con la IA generativa. Tavakoli (et al., 2024) en su artículo comenta que, en un estudio realizado por McKinsey “el 70% de las empresas con mejores resultados dijeron que han experimentado dificultades para integrar datos en modelos de IA, que van desde problemas con la calidad de los datos hasta la definición de procesos para la gobernanza de los datos y la

disponibilidad de datos de capacitación suficientes.” Es un proceso continuo para comprender cómo mejorar las capacidades de datos para que respalden la IA generativa.

Aprendizaje automático

Los modelos de ML son una disciplina de la IA que pueden aprender sin programación explícita por un humano, siendo capaces de detectar patrones y aprender a hacer predicciones y recomendaciones procesando datos y experiencias (McKinsey & Company, 2024). Las funcionalidades y sus aplicaciones en algoritmos de IA han llevado a que se sigan desarrollando nuevos modelos y más avanzados. Stanford Institute HAI (para más detalles, véase el **Anexo 6**) muestra cómo el número de modelos de ML creados ha ido incrementando por sector en el mercado desde el 2003; (para más detalles, véase el **Anexo 7**) como Estados Unidos (con 61 modelos) lidera en la creación por área geográfica, superando a China (con 15 modelos) que está en segundo lugar, y luego Francia (con 8 modelos).

Iberdrola (s.f.) divide los modelos de ML en tres categorías:

1. **Aprendizaje supervisado:** Están desarrollados con un aprendizaje previo basado en un sistema de etiquetas asociadas a datos que les permiten tomar decisiones o realizar predicciones. Un ejemplo es un detector de spam que etiqueta un e-mail como spam o no dependiendo de los patrones que ha aprendido del histórico de correos.
2. **Aprendizaje no supervisado:** Se enfrentan al caos de datos con el objetivo de encontrar patrones que permitan organizarlos de alguna manera. Por ejemplo, en una estrategia de marketing se utilizan para extraer patrones de datos masivos provenientes de las redes sociales y crear campañas de publicidad altamente segmentadas.

3. **Aprendizaje por refuerzo:** Su objetivo es que un algoritmo aprenda a partir de la propia experiencia. Esto es, que sea capaz de tomar la mejor decisión ante diferentes situaciones de acuerdo con un proceso de prueba y error en el que se recompensan las decisiones correctas. En la actualidad se está utilizando para posibilitar el reconocimiento facial, hacer diagnósticos médicos o clasificar secuencias de ADN.

Cada una tiene un objetivo diferente, pero lleva a beneficios similares como optimizar procesos, predicción de tendencias, mayor innovación (para más detalle, véase **Anexo 12**). Por ejemplo, su aplicación en el Procesamiento de Lenguaje Natural (PLN), que es una subcategoría de la inteligencia artificial que se centra en la interacción entre las computadoras y el lenguaje humano, e involucra la capacidad de una máquina para entender, interpretar y generar lenguaje humano de manera útil (Jurafsky & Martin, 2024), colaboró a que asistentes virtuales como Alexa de Amazon o Siri de Apple puedan traducir instantáneamente de un idioma a otro (Iberdrola, s.f.). Al igual que la IA, estos modelos también requieren de datos de calidad, estructurados, para el rendimiento del modelo.

Crowdsourcing

En este contexto, el crowdsourcing ha surgido como una estrategia eficaz para abordar los desafíos de la obtención de datos diversos y de calidad, ya que democratiza el conocimiento, apoya la globalización y empodera a los profesionales en todos los niveles de la organización (Fealy, 2021). Estelles & González-Ladrón-De-Guevara (2012) cuenta que el término de crowdsourcing fue establecido por Jeff Howe en 2006, definiéndolo como el acto de externalizar una tarea tradicional realizada por empleados a una gran masa de personas. Con el tiempo ha ido evolucionando, incluyendo más actividades colaborativas en línea que sólo externalización de tareas, causando

una confusión semántica para establecer una definición unificada. Por ejemplo, Estelles & González muestran algunas de las definiciones más citadas en diferentes estudios, como se muestra a continuación (para más detalles, véase el **Anexo 5**):

- “manera para que una empresa acceda a conocimientos externos.”
- “un caso especial de tan gran inteligencia colectiva.”
- “personas a las que se les paga por realizar tareas en internet por solicitantes.”

Y concluyeron que el 50% de las definiciones coinciden cuando la multitud (crowd) se perfila como un gran grupo de individuos. El número óptimo de personas dependerá de la iniciativa de crowdsourcing, debido a que la información debe ser filtrada y evaluada, y la heterogeneidad de la multitud dependerá del tipo de iniciativa considerada. Algunas campañas de crowdsourcing requerirán la sabiduría de una multitud heterogénea en la que cada persona aporta su conocimiento personal. En otros casos, la heterogeneidad no será tan importante, como en las tareas de traducción propuestas por Amazon Mechanical Turk que serán explicadas posteriormente.

Al ser una estrategia que puede tener diferentes objetivos, existen diferentes tipos de crowdsourcing, cada uno con un enfoque, objetivo y método diferente. Brabham (2013) en su libro *Crowdsourcing* explica los diferentes tipos:

- **Creación colectiva** (del inglés crowdcreation), donde una multitud colabora para crear contenido ideas, o productos. Es el más asociado con la creatividad y la innovación. Un caso de uso de esta estrategia fue cuando la empresa LEGO con su oferta *LEGO Ideas* utilizó esta estrategia de crowdsourcing en la cual los fanáticos creaban y compartían sus propias creaciones y estas se escogían para usarlos como nuevos diseños para nuevos sets de LEGO (Reffell, s.f.).

- **Votación colectiva** (del inglés crowdvoting), donde las multitudes participan en la evaluación y selección de ideas o productos. Se utiliza para medir preferencias o legitimar decisiones.
- **Financiamiento colectivo** (del inglés crowdfunding), para recaudar dinero para proyectos, startups, causas especiales.
- **Solución colectiva** (del inglés crowdsolving), que es la resolución de problemas mediante una convocatoria a expertos para obtener su contribución.

Adicional a estos tipos, existen también las **Micro tareas** (del inglés microtasks), que es la externalización de tareas pequeñas, repetitivas y específicas que se pueden realizar rápidamente, como el análisis de datos y etiquetado (E.A.E. Madrid, 2024).

Brabham (2013) destaca varios beneficios del crowdsourcing, como la diversidad de ideas y soluciones, la eficiencia en la resolución de problemas por la variedad de soluciones en corto plazo, la reducción de costos en la contratación de empleados adicionales o consultores especializados, acceso a talento global con diferentes conocimientos y habilidades, lealtad del consumidor al ser participe en el proceso de creación o selección de productos, la flexibilidad y escalabilidad que ofrece, puedes ajustarlo según tus necesidades específicas.

“El crowdsourcing puede ser una forma útil de entrenar modelos de IA y aprendizaje profundo aprovechando la inteligencia colectiva y las capacidades de un gran grupo de personas. La idea general es externalizar tareas que, de otro modo, realizaría un pequeño grupo de expertos a un gran grupo de personas no expertas. Esto puede incluir tareas como anotar datos, recopilar datos y etiquetar imágenes. Los beneficios del crowdsourcing para el entrenamiento de IA incluyen una mayor eficiencia, menores costos y una mayor precisión y diversidad de los datos de entrenamiento. Sin embargo, también existen

desafíos asociados con el crowdsourcing, como el control de calidad, la privacidad de los datos y la seguridad.” (Reffell, s.f.).

Reffell (s.f.) explica como ChatGPT, desarrollado por la empresa OpenAI y siendo uno de los modelos de lenguaje de gran escala (del inglés Large Language Model, LLM) utiliza crowdsourcing para mejorar la fiabilidad de sus datos y respuestas. Le es posible recolectar y anotar grandes volúmenes de datos textuales etiquetados, lo que simplifica la tarea de entrenar y mejorar el modelo. Cuando OpenAI lanzó el modelo GPT-3 en 2020, este estaba soportado por 175 mil millones de parámetros de datos (Li, 2020). Como indica IBM (2023):

“Los modelos de lenguaje grande (LLM) son una categoría de modelos básicos entrenados en inmensas cantidades de datos, lo que los hace capaces de comprender y generar lenguaje natural y otros tipos de contenido para realizar una amplia gama de tareas dentro de las organizaciones. Entre esas podemos encontrar generación de texto, resumen de contenido, asistencias de IA, generación de código, análisis de sentimientos, traducción de idiomas.”

Pero se ha encontrado que modelos de LLM siguen siendo susceptibles a reflejar sesgos en los datos ocasionando la creación de estereotipos no reales, interpretación errónea de datos, y sin las medidas protectoras adecuadas pueden memorizar y reproducir datos textuales exactos de su entrenamiento, lo que puede comprometer la privacidad al revelar información sensible o protegida (Stanford Institute HAI, cap 3, 2024).

Calidad de datos

La eficacia de estos modelos depende en gran parte de la calidad de los datos, y el reto reside en mantener ese nivel. La creciente disponibilidad de datos y los avances tecnológicos subrayan la

necesidad de nuevos planteamientos y modelos para afrontar estos desafíos. Angeles & García-Ugalde (2009) definen la calidad de datos de la siguiente manera:

La calidad de los datos es un ámbito de área multidisciplinar, en la que intervienen la gestión, la estadística y la informática. Consideramos la calidad de los datos no como el fin, sino como el medio para tomar decisiones informadas. Las propiedades pertinentes de la calidad de los datos, sus prioridades y el nivel de calidad de datos esperado dependen no sólo de la experiencia del consumidor de datos, sino también del tipo de sistema de información.

Otra definición más actual es la de Batini & Scannapieco (2016) que definen la calidad de datos como el grado en que un conjunto de datos es apto para su uso previsto, basado en características medibles y verificables.

Kariluoto (et al., 2021) investigaron si la calidad de los datos en el aprendizaje automático dependía de la cantidad de datos y el número de características de entrada, y sus resultados empíricos se asemejan a lo explicado anteriormente: ambos factores influyen en la calidad del modelo, la calidad sobre la cantidad resalta, pero esto seguirá dependiendo del contexto de aplicación del modelo y de la naturaleza de los datos. Yandrapalli (2024) destaca que los datos organizados y limpios permiten un mejor “feature engineering”, que se refiere a la selección y modificación de características relevantes para identificar nuevos patrones significativos en los modelos ML, y no causar un sobreajuste en el modelo causando incoherencias.

Hubo casos donde el modelo presentó sesgos e incoherencias causando problemas significativos: En 2015, Jacky Alciné, un ingeniero de software señaló que los algoritmos de reconocimiento de imágenes de Google Photos estaban clasificando a sus amigos, personas de color, como "gorilas",

un término particularmente ofensivo por su contexto relacionado con la época de la esclavitud. (Matsakis, 2023). En 2018, la empresa Amazon decidió no continuar usando un algoritmo de contratación de empleados soportado por IA, luego de descubrir que para puestos como desarrollador de software y otros puestos técnicos no se calificaba de manera neutral cada curriculum por su “género” (Dastin, 2018). Otro caso es el del chatbot *Tay*, creado por Microsoft para Twitter para aprender a interactuar con usuarios a través del ML, en menos de 24 horas después de su lanzamiento, el chatbot empezó a responder con mensajes racistas, sexistas y ofensivos, lo que llevó a su desactivación (Wakefield, 2016). Por este motivo cuando un sistema de IA empieza a tener comportamientos inapropiados por datos sesgados se le conoce como “efecto Tay” (Binns, 2018). Estos casos subrayan como el mal etiquetado, manejo y calidad de datos afecta de manera negativa y significativa el resultado del modelo.

Colaboradores

Los colaboradores, individuos que, a través de plataformas digitales contribuyen con pequeñas tareas o actividades que, al combinarse, ayudan a resolver problemas más complejos o a generar contenido a gran escala, contribuyen de manera significativa a la variabilidad en la calidad de los datos, dependiendo tanto de su motivación como de sus habilidades (Estelles & González-Ladrón-De-Guevara, 2012). Brabham (2013) comenta que es crucial comprender los motivadores para diseñar tu proyecto para que sea exitoso y atraiga los colaboradores adecuados. Estos motivadores los divide en dos, intrínsecos y extrínsecos:

Tabla #1: Motivadores intrínsecos.

Motivadores intrínsecos	
Interés personal y pasión	Colaboradores motivados por un interés genuino en el tema.
Altruismo y sentido de comunidad	Participación impulsada por el deseo de ayudar a otros o contribuir a una causa significativa.
Aprendizaje y desarrollo personal	Motivación por la oportunidad de aprender nuevas habilidades o aplicarlas.
Reconocimiento y estatus	El deseo de obtener reconocimiento y estatus dentro de una comunidad, especialmente donde se valora la experiencia.

Tabla #2: Motivadores extrínsecos.

Motivadores extrínsecos	
Compensación financiera	Participación motivada por recompensas financieras o compensaciones monetarias.
Oportunidades profesionales	Oportunidades de construir un portafolio, ganar experiencia o atraer la atención de posibles empleadores.
Premios y competencias	Participación en concursos o competencias con la motivación de ganar premios desafíos.
Acceso a productos o servicios	Motivación por la promesa de acceso anticipado o exclusivo a productos o servicios, especialmente en tecnología.

Pero también subraya que las motivaciones no son estáticas, pueden ir cambiando en función que avanza el proyecto; es importante identificar estos cambios a tiempo para adaptar las estrategias y obtener mejores resultados, tener comunicaciones claras y continuas para crear una comunidad alrededor del proyecto, y diseñar proyectos de crowdsourcing que sean transparentes y justos, en términos de compensación y reconocimiento, para mejorar la eficacia (Brabham, 2013).

Política y gobernanza

Para aprovechar el potencial de los modelos de IA y transformar la forma en que hacen negocios, las organizaciones requieren una base sólida en prácticas de gobernanza para establecer políticas y procedimientos para mantener la calidad y seguridad de los datos, pero también optando por un enfoque de uso ético. Esto implica tener acceso a herramientas y tecnología de IA que sean seguras, confiables, transparentes y responsables (IBM, 2023). Así como lo expresa Christina Montgomery, vicepresidente de privacidad y confianza en IBM "Si tú no tienes una gobernanza para IA, no podrás adoptar soluciones de IA a escala" (IBM, 2023).

La creación de esta gobernanza les permitirá tener más trazabilidad del origen de los datos, explicar el cómo y el porqué de los resultados, monitorizar y controlar el rendimiento para evitar desviaciones, y para fomentar una cultura inteligente de la IA (IBM, s.f.). Una encuesta global fue realizada por Stanford Institute HAI (para más detalles, véase el **Anexo 9**) sobre IA responsable y destaca que las principales preocupaciones de las empresas en relación con la IA incluyen la privacidad, la gobernanza, la seguridad, y la transparencia. Las regulaciones aplicadas a la IA en los Estados Unidos han incrementado de menos de 5 regulaciones en el 2016 a tener 25 en el 2023 (para más detalles, véase el **Anexo 10**), y en la Unión Europea de menos de 5 regulaciones en el 2017 ascendió a 32 en el 2023 (para más detalles, véase el **Anexo 11**). Stanford Institute HAI (cap.7, 2024) muestra como el número de agencias reguladoras estadounidenses que emiten regulaciones sobre IA aumentó de 17 en 2022 a 21 en 2023, lo que indica una creciente preocupación por la regulación de la IA entre una gama más amplia de organismos reguladores estadounidenses.

En Estados Unidos en 2023, el gobierno presidencial de Joe Biden obtuvo compromisos voluntarios de siete grandes empresas de IA: Google, Microsoft, Meta, Amazon, OpenAI,

Anthropic e Inflection, para promover el desarrollo de una IA segura y fiable. Estos compromisos implican la realización de una evaluación de seguridad interna y externa de los sistemas de IA antes de su lanzamiento, compartiendo información sobre los riesgos identificados, reporte de problemas, e identificación cuándo el contenido está generado por IA (Stanford Institute, cap.7, 2024). Adicional, el Reino Unido anuncia el Instituto de Seguridad de IA (AI Safety Institute) para promover la seguridad entre el interés público, buscando proteger a la humanidad de avances de la IA creando un marco sociotécnico necesario para comprender y gobernar los riesgos asociados (Stanford Institute, cap.7, 2024).

Solo en 2023, la industria de la IA experimentó un aumento del 40 % en los incidentes denunciados relacionados con violaciones de datos y sesgo de modelos, lo que pone en relieve la necesidad de contar con marcos de gobernanza sólidos. Según una encuesta reciente de PwC, el 85 % de los líderes de la IA citan la gobernanza como su principal preocupación, lo que enfatiza la importancia de la confianza y la seguridad de la propiedad intelectual valiosa (DataRobot, 2023). DataRobot, la plataforma de IA catalogada como la #1 en casos de uso de gobernanza (para más detalle, véase **Anexo 13**) explica las características clave para ofrecer un marco de gobernanza efectivo, y las divide en cuatro fases que evalúan de principio a fin un modelo de IA:

1. **Fase de construcción:** los científicos de datos y los profesionales de IA sientan las bases para crear soluciones de IA sólidas.
2. **Fase de prueba:** Los modelos se someten a pruebas rigurosas para garantizar que cumplan con nuestros estándares y funcionen de manera confiable en diversas condiciones.
3. **Fase de producción:** los modelos se implementan y gestionan en un entorno en vivo.

4. **Fase de monitoreo y gestión:** Las herramientas de supervisión y gobernanza ayudan a los equipos a mantener el cumplimiento, la integridad y la precisión de las soluciones de IA para lograr la excelencia operativa.

Además, debido a que DataRobot (2023) prioriza la privacidad, la seguridad y la eficiencia de los datos, esta posee seis funciones únicas que potencia cada fase para que destaque aún más:

1. **Visibilidad y trazabilidad:** la trazabilidad total de los datos, el linaje del modelo y el control de versiones garantiza que cada cambio se rastree y documente, lo que hace que las aplicaciones sean seguras y útiles.
2. **Auditorías y documentación de cumplimiento:** generación automatizada de informes de cumplimiento y registros de auditoría para cumplir con los requisitos reglamentarios y garantizar la transparencia.
3. **Evaluación y pruebas LLM únicas:** detecte riesgos potenciales utilizando conjuntos de datos sintéticos y reales para evaluar sus modelos de IA predictivos y generativos y comparar el rendimiento.
4. **Pruebas CI/CD:** la capacidad de ejecutar pruebas de prototipos y evaluar ML o soluciones generativas con métricas de calidad para clasificar los experimentos RAG.
5. **Intervención, moderación y alertas en tiempo real:** monitoreo continuo con notificaciones instantáneas y capacidades de intervención habilitadas por modelos de guardia y métricas para abordar los problemas a medida que surgen.

6. **Catálogo de IA:** registre, realice un seguimiento y cree versiones de todos los activos de IA fácilmente, ya sea que se hayan creado dentro o fuera de la plataforma de IA DataRobot, todo a través de un centro centralizado seguro.

Capítulo 4: Análisis de plataformas.

Amazon Mechanical Turk (MTurk)

Lanzada en noviembre de 2005, es un mercado de crowdsourcing que facilita que individuos y empresas subcontraten sus procesos y trabajos a una fuerza laboral distribuida que pueda realizar estas tareas virtualmente. Esto puede incluir varias cosas, desde realizar investigaciones y validaciones de datos simples hasta tareas más subjetivas como participación en encuestas, moderación de contenido y más. MTurk permite a las empresas aprovechar la inteligencia, las habilidades y los conocimientos colectivos de una fuerza laboral global (para más detalles, véase el **Anexo 1**) para optimizar los procesos comerciales, aumentar la recopilación y el análisis de datos y acelerar el desarrollo del aprendizaje automático (MTurk, s.f.). Es ideal para realizar tareas simples y repetitivas en sus flujos de trabajo que deben manejarse manualmente, como validación y entrada de datos, investigación de mercado, moderación de contenido, y anotación de datos. Al aprovechar las habilidades de los trabajadores distribuidos en un modelo de pago por tarea, puede reducir significativamente los costos y, al mismo tiempo, lograr resultados que tal vez no hubieran sido posibles con solo un equipo dedicado (MTurk, s.f.). Hay dos modos de uso, uno como solicitante y otro como colaborador:

Como solicitante debes:

1. Crear una cuenta en MTurk Requester.
2. Diseñar tareas (HITs) especificando lo que necesitas y estableciendo recompensas atractivas para los trabajadores.
3. Realizar pruebas en el *sandbox*: entorno de prueba dentro de MTurk que permite a los solicitantes simular y probar sus Human Intelligence Tasks (HITs) antes de publicarlas en el Marketplace real.
4. Publicar y gestionar HITs, revisar el trabajo realizado, y pagar a los colaboradores.

Como colaborador debes:

1. Crear una cuenta en MTurk Worker.
2. Completar el perfil y las calificaciones: Crear el perfil y buscar oportunidades para obtener calificaciones adicionales que te permitirán acceder a HITs mejor remunerados.
3. Buscar HITs: Navegar por las tareas disponibles (HITs) en la plataforma y seleccionar aquellas que te interesen.
4. Realizar las tareas: Completa las HITs según las instrucciones proporcionadas.
5. Monitorear la tasa de aprobación: Mantener una alta tasa de aprobación para acceder a mejores HITs en el futuro.
6. Recibir pagos: Después de que las HITs sean aprobadas, recibirás el pago directamente en tu cuenta de Amazon Payments.

Experiencias exitosas con MTurk

Allen Institute for Artificial Intelligence (AI2) es uno de los casos de uso como solicitante. Es un instituto de investigación sin fines de lucro fundado en 2014 con la misión de realizar investigaciones e ingeniería de IA de alto impacto al servicio del bien común. AI2 ha emprendido varios proyectos ambiciosos para impulsar avances fundamentales en ciencia, medicina y conservación a través de la IA (AI2, 2024). Uno de esos proyectos fue el de la plataforma de búsqueda *Semantic Scholar*, el cual fue diseñado para facilitar a investigadores su búsqueda de literaturas científicas relevantes y más citadas en su campo, potenciado mediante IA (Allen Institute for AI, 2015). Hoy en día posee más de 214 millones papers, ha generado más de 2,49 mil millones de citas, de más de 79 millones de autores, esto gracias a su gran modelo soportado por etiquetado de datos de calidad.

Michael Schmitz, director de Ingeniería del Instituto Allen de Inteligencia Artificial, comparte su éxito y experiencia con MTurk:

“En AI2, estamos impulsando el estado del arte de la Inteligencia Artificial, que a menudo requiere datos anotados por humanos para entrenar nuevos sistemas y medir nuestro progreso. En particular, utilizamos plataformas de crowdsourcing como Amazon Mechanical Turk para crear conjuntos de datos que ayuden a nuestros modelos a aprender conocimientos de sentido común, que a menudo son necesarios para responder preguntas básicas que son fáciles para los humanos, pero aún bastante difíciles para las máquinas. Amazon Mechanical Turk proporciona una plataforma flexible que nos permite aprovechar el conocimiento humano para avanzar en la investigación del aprendizaje automático.”

(Michael Schmitz, director of engineering AI2, s.f.).

Otra experiencia exitosa es el del programa C-SATS de la empresa Johnson & Johnson (C-SATS, s.f.), que ha utilizado MTurk para que cirujanos puedan cargar sus videos de cirugías y recibir evaluaciones por parte de expertos. En este proceso los cirujanos pueden recibir comentarios anónimos, o sugerencias de habilidades técnicas, lo que puede contribuir a la calidad de las cirugías y beneficia tanto a pacientes como a profesionales de la salud (Svetlana, s.f.). Por esas retroalimentaciones cualitativas y cuantitativas sobre técnicas o procedimientos cirujanos que usan esta plataforma han logrado reducir el tiempo de una cirugía en un promedio de 27 minutos, que sus pacientes tengan un 34% menos de probabilidad de prolongar su estadía en un hospital, también cada procedimiento realizado por un médico que utiliza C-SATS, logra un ahorro en costes de \$586 dólares en promedio (C-SATS, s.f.).

Baidu Research, una división de la empresa Baidu Inc., reúne investigadores globalmente para trabajar en tecnologías como reconocimiento de imágenes, comprensión de videos,

reconocimiento de voz, procesamiento del lenguaje natural e inteligencia semántica (Diamos, s.f.). Greg Diamos, investigador principal en Baidu, también asegura que los investigadores mejoraron la tecnología de clonación de voz utilizando las evaluaciones críticas proporcionadas por los participantes de MTurk, quienes escuchan muestras de audio generadas por estos sistemas y las comparan con grabaciones humanas. Resalta la rapidez con la que MTurk ha permitido obtener evaluaciones de calidad de una gran cantidad de personas, acelerando el desarrollo de interfaces personalizadas para voz. “Gracias a la utilización de Amazon Mechanical Turk, conseguimos evaluar rápidamente la calidad de las muestras de audio al contar con un amplio grupo de oyentes que nos permitieron comparar el audio sintético con grabaciones realizadas por humanos.” (Diamos, s.f.).

La empresa KITT.AI, que también hace parte de Baidu, desarrolló *Snowboy*, un motor de detección de palabras clave personalizable. Utilizando MTurk, KITT.AI pudo recolectar una amplia variedad de grabaciones de palabras clave personalizadas por los usuarios. Esta diversidad de datos fue crucial para entrenar modelos de IA, que sean precisos y eficaces en la detección de comandos de voz específicos (Yao, s.f.). Xuchen Yao, CEO de KITT.AI, señala que MTurk proporcionó una solución económica y escalable para obtener los datos necesarios para el entrenamiento de sus algoritmos. “Amazon Mechanical Turk nos proporcionó una fuerza laboral global y escalable a bajo costo, lo que nos permitió generar conjuntos de datos diversos necesarios para entrenar nuestros modelos de IA.” (Yao, s.f.). Desafortunadamente el producto Snowboy fue desactivado a finales del 2020 por decisión interna.

Ahora, analizando desde el punto de vista de un colaborador, investigaciones pasadas indican que numerosos colaboradores de MTurk y otras plataformas percibían remuneraciones mínimas por el tiempo y esfuerzo invertidos, lo cual genera inquietudes éticas acerca de la explotación laboral

(Difallah et al., 2018). Además, aunque aproximadamente el 60% de los participantes están satisfechos con la cantidad de dinero que ganan en MTurk, esta cifra es menor que la satisfacción general con los salarios en la población de EE. UU (para más detalles, véase el **Anexo 2**).

Pero estudios recientes muestran cómo ha mejorado en aspectos relacionados con la remuneración; la plataforma en línea Swift Salary (2023), que se centra en proporcionar información detallada a personas sobre toma de decisiones financiera, cómo ganar dinero en diferentes plataformas, realizó una evaluación de diferentes plataformas de trabajo online, centrándose en aspectos relacionados con el pago a los trabajadores (para más detalles, véase el **Anexo 4**). En este estudio, MTurk se posicionó en primer lugar, destacando en dos aspectos cruciales: En primer lugar, su velocidad de pago, MTurk tarda solo 3 días en pagar a sus colaboradores, mientras que otras plataformas pueden tardar entre 15 días y un mes. En segundo lugar, el pago mínimo por tarea, MTurk obtuvo una calificación de 4,5 sobre 5 en este aspecto, esto demuestra como la plataforma ha ido invirtiendo en mejorar la remuneración de los colaboradores. Es importante señalar que estos datos provienen de encuestas realizadas a empleados de las plataformas evaluadas.

Figure Eight

Fundada en diciembre de 2007 con el nombre CrowdFlower, es una plataforma de enriquecimiento de datos con acceso a una fuerza laboral en línea y bajo demanda de personas que completan tareas. Algunas de esas tareas son categorización de datos, creación de metadatos, análisis de los sentimientos, anotación de imagen, entre otras (CrowdFlower, s.f.). CrowdFlower cambió su nombre a Figure Eight en 2018. Posteriormente, en 2019, Figure Eight fue adquirida por Appen, una empresa especializada en soluciones de datos de alta calidad para el aprendizaje automático y

la inteligencia artificial (Appen, 2019). Mark Brayan, director ejecutivo de Appen, comentó lo siguiente sobre la adquisición:

“Ahora tenemos lo mejor de ambos mundos: nuestra plataforma de gestión de la nube altamente eficiente y un público multilingüe, hábil y escalable, combinado con la innovadora plataforma SaaS orientada al cliente de Figure Eight con anotaciones asistidas por ML. Combinados, cumpliremos y superaremos los requisitos de escala, velocidad y calidad de nuestros clientes.” (Appen, 2019).

Figure Eight combina lo mejor de la inteligencia humana y automática para proporcionar datos de capacitación anotados de alta calidad que impulsan las soluciones comerciales y de aprendizaje automático más innovadoras del mundo (Appen, 2019). Como MTurk, Figure Eight tiene procesos de uso diferentes si eres solicitante o colaborador, y se asemejan en ambos casos.

Experiencias exitosas con Figure Eight

Autodesk, una compañía que ofrece software de diseño, ingeniería y entretenimiento a clientes de los sectores de arquitectura, fabricación, construcción, medios y entretenimiento notó que sus clientes necesitaban más soluciones integrales, en vez de productos individuales; era necesario entender mejor a sus clientes y entrenar a su equipo de ventas para ofrecer soluciones en lugar de imponer productos. Para lograr identificar las oportunidades de venta e incentivar a su equipo, era necesario datos completos, como sector de la empresa, tamaño, empresa matriz. (Appen, s.f.).

“Al igual que muchas grandes empresas, los datos de CRM de Autodesk provienen de múltiples fuentes con una calidad de datos variable. Los datos ingresados por los clientes suelen ser inexactos y difíciles de normalizar. Las diferentes fuentes de datos tienen

distintos niveles de integridad y todo debe estar integrado para producir un único almacén de datos confiable.” (Appen, s.f.).

Como solución realizaron lo siguiente: Los registros de clientes a los que les falten datos clave se envían automáticamente a Appen a través de una Interfaz de Programación de Aplicaciones (por sus siglas en inglés, API) que es un conjunto de reglas y protocolos que permiten a diferentes aplicaciones interactuar entre sí (Fielding & Taylor, 2022). El equipo primero encuentra la URL oficial de la empresa de destino y la utiliza (así como otras fuentes de datos) para identificar la industria de la empresa, el tamaño y las empresas matrices. Si es necesario, se puede agregar información adicional como la dirección de la empresa, el número de teléfono y la descripción, o información ejecutiva de la empresa como nombres de los equipos, cargos y biografías.” (Appen, s.f.), el porcentaje de registros con información completo aumentó del 70% al 85%. Sólo en octubre del 2012, Appen procesó más de 450,000 registros para Autodesk. Patrick Booher, director de Gestión de datos empresariales de Autodesk manifestó lo siguiente: “Appen fue parte de la solución general que nos permitió pasar de ser direccionales a ser transaccionales: no se puede realizar transacciones con el 70 %.” (Appen, s.f.).

Aleix Canet, Product Owner y Team Lead en Airbus cuenta su experiencia exitosa implementando la plataforma: “Estamos satisfechos con la alta calidad y la rapidez de los resultados entregados y estamos felices de haber iniciado esta colaboración con Appen.” (Appen, s.f.). Esto fue después de aumentar la precisión de sus tiempos de vuelo estimados y el tiempo de llegada en su herramienta de análisis avanzado AirSense. Para esto, se necesitaban anotar miles de rutas de vuelo basadas en datos históricos y secuencias de coordenadas de GPS (Global Positioning System). La solución propuesta por Appen (s.f.) fue la siguiente:

“Airbus capacitó a nuestro equipo de anotadores en el lugar mediante videos y comentarios constantes para garantizar una alta precisión en las anotaciones de imágenes. Las tareas requerían que los anotadores mapearan el comienzo, el medio y el final de la trayectoria de vuelo. También necesitaban coordinar una posición en la imagen con una coordenada GPS. Las herramientas de control de calidad de la plataforma permitieron al equipo de Airbus revisar los datos de manera rápida y precisa.”

Como resultado, el proyecto piloto se lanzó con éxito con un 98% de precisión, pudieron anotar datos de manera segura y precisar miles de imágenes.

Clickworker

Fundada en 2005, es un proveedor de servicios que cuenta con una red global de millones de clickworkers (colaboradores) en más de 130 países, los cuales ofrecen servicios de gestión de datos y entrenamiento en inteligencia artificial. Entre los servicios ofrecidos están la generación de datos para visión artificial, reconocimiento de voz y texto, investigación en línea, optimización del contenido para motores de búsqueda y gestión eficiente de puntos de venta. Diferentes industrias usan estos servicios para mejorar sus modelos de IA, optimizar la disposición de productos en tiendas y realizar investigaciones competitivas en línea. Estos proyectos se dividen automáticamente en micro trabajos y son procesados por clickworkers calificados. Luego, los resultados se vuelven a ensamblar y se transmiten al cliente con garantía de calidad (Clickworker, 2009). Para cada proyecto individual, la plataforma puede proporcionarle conjuntos de datos de IA únicos y recién creados, como fotos, audio, grabaciones de video y texto para ayudarlo a desarrollar su algoritmo basado en el aprendizaje. Los datos de entrenamiento de IA son la base

para construir y mejorar modelos de IA. Sus algoritmos necesitan interacción humana si desea que proporcionen resultados similares a los humanos (Clickworker, 2016).

Experiencias exitosas con Clickworker

En un caso de estudio que desarrollaron sobre datos de entrenamiento de reconocimiento facial para ayudar a entrenar un software, clickworkers de todo el mundo se tomaron fotografías de sus rostros con diferentes expresiones faciales y desde varias perspectivas. De esta manera, proporcionan al desarrollador de software importantes datos de entrenamiento en reconocimiento facial que se utilizan para entrenar el sistema (Clickworker, 2017). Con estas fotografías, el fabricante puede entrenar el software para el reconocimiento facial, optimizarlo y especializar el sistema para diferentes tareas. George Eastman House, el museo de fotografía más antiguo del mundo fundado en 1947 en la propiedad del fundador de Kodak, George Eastman, implementó una estrategia similar. El proyecto implicaba en el etiquetado fotográfico y la catalogación de más de 400.000 imágenes del Museo Internacional de Fotografía y Cine George Eastman House (Clickworker, 2011).

La importancia de Clickworker en la generación de distintos conjuntos de datos para el aprendizaje automático es indiscutible. La creación de preguntas en formato de texto para chatbots con fines de soporte técnico, así como la recolección de grabaciones vocales utilizadas en software de reconocimiento para voz y la recopilación conjuntos fotográficos destinados a sistemas que realizan reconocimiento facial. Estos proyectos garantizan que los modelos de IA se entrenen con datos diversos y de excelente calidad, lo que es fundamental para desarrollar aplicaciones precisas y confiables basadas en IA (Clickworker, 2018).

Cabe añadir que la auditoría de certificación confirma que Clickworker cumple con los requisitos de la norma ISO/IEC 27001:2017. Esto incluye la implementación exitosa de un sistema de gestión de seguridad de la información (SGSI), la disponibilidad asegurada de los sistemas y procesos de TI, la prevención de riesgos de TI y la confidencialidad de los datos dentro de clickworker.com Inc. y clickworker GmbH (Maione, 2022).

Resumen análisis plataformas

Tabla #3: Comparativa entre MTurk, Figure Eight y Clickworker.

	Amazon Mechanical Turk (MTurk)	Figure Eight (Appen)	Clickworker
Descripción	Mercado de crowdsourcing global.	Plataforma de enriquecimiento de datos.	Proveedor de servicios completos para IA.
Tareas comunes	Investigaciones, validación de datos, encuestas, moderación de contenido.	Categorización de datos, creación de metadatos, anotación de imágenes.	Creación de conjuntos de datos de IA (fotos, audio, video, texto), etiquetado fotográfico, catalogación de imágenes.
Beneficios	Amplia fuerza laboral, flexibilidad.	Alta precisión en tareas complejas.	Alta calidad de datos.

Calidad y Supervisión	Sistema de calificación y cualificación de trabajadores para garantizar la calidad de los resultados.	Ofrece anotaciones asistidas por ML y validaciones de calidad integradas.	Certificación ISO/IEC 27001:2017 que asegura la seguridad y confidencialidad de los datos.
Desafíos y Limitaciones	Remuneración y motivación de trabajadores.	Equilibrar la inteligencia humana y automática usada para mantener la calidad de los datos, garantizar la satisfacción de los empleados y pagar adecuadamente.	Calidad y seguridad de los datos a gran escala, asegurar una compensación justa y adecuada para los Clickworkers.

Tabla #4: Tabla ventajas y desventajas.

	Amazon Mechanical Turk (MTurk)	Figure Eight	Clickworker
Ventajas	<ul style="list-style-type: none"> - Red global de trabajadores. - Flexibilidad y variedad de tareas. - Fácil integración con herramientas de 	<ul style="list-style-type: none"> - Herramientas avanzadas para la anotación y etiquetado de datos. - Capacidad para manejar tareas complejas. 	<ul style="list-style-type: none"> - Red global de trabajadores. - Herramienta para verificación de datos y creación de contenido. - Pago basado en rendimiento

	<p>Amazon Web Services (AWS).</p> <ul style="list-style-type: none"> - Pago rápido y fiable para los colaboradores. 	<ul style="list-style-type: none"> - Amplio equipo de trabajadores calificados. 	<p>para aumentar la calidad de las tareas.</p>
Desventajas	<ul style="list-style-type: none"> - Control de calidad es un desafío por la gran cantidad de información - Alta competitividad para realizar las tareas mejor recompensadas (competencia interna). 	<ul style="list-style-type: none"> - Inconstancia con la calidad de los datos. 	<ul style="list-style-type: none"> - Los salarios pueden ser bajos, lo que puede afectar la calidad del trabajo.

Aplicación del crowdsourcing

Task Us (2023) sugiere con los siguientes pasos cómo crear una estrategia bien planificada de crowdsourcing para aprovechar el crecimiento de valor del mercado, que se estima que alcanzará un valor de más de 1,5 billones de dólares en 2030:

1. Definir el objetivo de la campaña, qué datos se van a necesitar para el algoritmo de IA.
2. Identificar la audiencia objetivo, quiénes son los participantes más aptos.

3. Comunicar instrucciones claras para el proyecto; al crear directrices garantiza que comprendan el objetivo y mejore la calidad de datos.
4. Acordar pagos, que generalmente se manejan con un sistema de pago por tarea y generan rentabilidad.
5. Garantizar la seguridad y calidad de los datos.

De igual manera también aconseja cómo seleccionar la plataforma de crowdsourcing indicada, porque esto afectará en la calidad y resultados de tu campaña. Task Us (2023) comenta qué aspectos hay que tener en cuenta en esta decisión:

- **Características de la plataforma:** Como funcionalidades, practicidad, eficiencia, monitoreo de datos y su desempeño.
- **El grupo de trabajadores colaborativos:** Si es diverso, diferentes niveles de experiencia, habilidades, idiomas y comunicación.
- **Historial de éxito / estudios de caso:** Revelación de fortalezas, objetivos cumplidos, anticipación de desafíos.
- **Seguridad y privacidad de los datos.**
- **Precios y estructuras de costos.**

Al seleccionar una plataforma de crowdsourcing, es crucial considerar las necesidades específicas del proyecto. MTurk destaca por su versatilidad, permitiendo la realización de una amplia gama de tareas con gran flexibilidad. Figure Eight se especializa en tareas complejas que demandan un alto grado de precisión en la anotación y etiquetado de datos. Clickworker se posiciona como una opción ideal para la verificación de datos y la creación de contenido. Es importante resaltar que el control de calidad de los datos representa un desafío común para todas estas plataformas. La capacidad de mantener la transparencia y seguridad de la información recopilada es un factor

crítico que debe evaluarse cuidadosamente, como se explicó anteriormente en el punto sobre Calidad de Datos.

Capítulo 5: Resultados.

Los hallazgos de este estudio muestran que las plataformas de crowdsourcing analizadas son útiles en términos de su uso para la recolección y anotación de datos para entrenar algoritmos de inteligencia artificial (IA) y aprendizaje automático (ML), y esto logra suceder por los siguientes motivos:

- Las plataformas como Amazon Mechanical Turk (MTurk), Figure Eight y Clickworker son cruciales para mejorar la diversidad y la calidad de los datos utilizados en el entrenamiento de modelos de IA y mejorar los outputs obtenidos, casos como el del programa C-SATS utilizado por médicos, que ayudó a reducir tiempos de cirugía, costos por procedimiento, acortar los tiempos de estadía en el hospital de los pacientes; también el de la empresa de software Autodesk que quería entrenar y mejorar su equipo de ventas obteniendo datos más completos de los clientes, aumentando de un 70% al 85% los datos de los clientes; o el etiquetado y catalogación de 400,000 fotos para el Museo Internacional de Fotografía y Cine George Eastman House por parte de colaboradores de Clickworker.
- Las ventajas evidentes de estas plataformas incluyen su flexibilidad operativa y su capacidad de optimizar procesos, lo que ha permitido a las organizaciones expandir sus estrategias de recolección de datos de manera eficiente.
- La capacidad de dividir tareas complejas en micro tareas y distribuir las a una amplia base de colaboradores permite optimizar tanto el tiempo como los recursos necesarios para obtener grandes cantidades de datos etiquetados y anotados.

Sin embargo, desafíos como la calidad y variedad de datos se vieron reflejados en modelos de IA, como en el caso de Microsoft con el chatbot que presentaba “efecto Tay” o Google Photos con su etiquetado de fotos con sesgos raciales. La falta de gobernanza y su importancia fue identificada por

el número de denuncias relacionados con manejo de datos personales y sesgos que mostró DataRobot.

La remuneración, siendo uno de los principales motivadores extrínsecos de los colaboradores, si tiende a ser muy variable o bajo, va a desincentivar a los colaboradores calificados resultando en datos con sesgos o erróneos, como lo explica Brabham en su libro Crowdsourcing, afectando negativamente el desarrollo de modelos efectivos. También, aunque las plataformas como MTurk y Clickworker tengan acceso a una fuerza global, la heterogeneidad cultural entre los colaboradores no está garantizada porque, aunque geográficamente estén dispersos, pueden compartir aspectos culturales o sociodemográficos, limitando la diversidad de datos y creando sesgos en los modelos entrenados.

Establecer políticas de gobernanza responsables es necesario para no poner en riesgo la integridad de los datos y dar fiabilidad a los modelos de IA desarrollados, los desafíos relacionados con la coherencia y calidad de los datos, así como la necesidad de garantizar la privacidad y seguridad de los datos, son aspectos sensibles que deben ser gestionados con máxima rigurosidad para lograr una implementación exitosa de gobernanza. Es por eso que países en todo el mundo están invirtiendo en crear nuevas regulaciones y políticas sobre los sistemas de IA y ML.

En resumen, el éxito de una estrategia de crowdsourcing depende entre el equilibrio de la eficiencia operativa (escoger la plataforma indicada, crear una estrategia, definir tus objetivos), y la implementación de estándares éticos y de calidad (remuneración justa, etiquetado de datos por expertos, gobernanzas sobre los datos). El continuo éxito dependerá de la capacidad de las organizaciones de gestionar estos desafíos al mismo tiempo que las tecnologías siguen avanzando.

Capítulo 6: Conclusiones, limitaciones y futuras líneas de investigación.

La investigación realizada a creado una comprensión del papel crucial que juegan las plataformas de crowdsourcing en el desarrollo y entrenamiento de modelos de aprendizaje automático (ML) e inteligencia artificial (IA). Las plataformas analizadas han demostrado ser cruciales para la recolección de datos a gran escala y la integración de diversas perspectivas, lo que enriquece los conjuntos de datos utilizados en IA y ML. Las plataformas de crowdsourcing se están posicionando como un puente entre la tecnología avanzada y la inteligencia colectiva, además de su funcionalidad técnica. El modelo de cooperación de una fuerza global ha brindado nuevas posibilidades para democratizar la participación en proyectos tecnológicos, lo que permite a personas de todo el mundo contribuir a los avances en IA que de otro modo estarían limitados a entornos más centralizados. Sin embargo, la tensión entre los principios éticos que deben guiar el uso de estas plataformas y la eficiencia operativa es un tema importante que se ha identificado. Es innegable que el crowdsourcing ha permitido a las organizaciones expandir sus procesos, pero este éxito debe equilibrarse con la necesidad de garantizar condiciones laborales y remuneración justa para los empleados y proteger la privacidad y seguridad de los datos que se manejan.

De igual modo, esta investigación presenta limitaciones por dos motivos: Primero, se centró en analizar un conjunto específico de plataformas, probablemente excluyendo otras plataformas emergentes o con características y ventajas más eficientes. Segundo, no de todas las experiencias exitosas que se expusieron tenemos todos los resultados y beneficios que obtuvieron, hizo falta obtener más información y detallada para tener un análisis más holístico de los resultados. La combinación de ambos tipos de datos ayudaría a reducir los sesgos en los resultados obtenidos ofreciendo una visión más completa de su implementación y resultado.

El desarrollo y la prueba de sistemas de control de calidad automatizados que utilicen algoritmos de inteligencia artificial para supervisar y mejorar continuamente la calidad de los datos

recolectados a través de plataformas de crowdsourcing podrían ser el foco de futuras investigaciones. Este método ayudaría a reducir la dependencia de la intervención humana y la variabilidad de la calidad de los datos. También, con el crecimiento exponencial de estas tecnologías, otra futura investigación podría enfocarse en evaluar el impacto ambiental de las plataformas de crowdsourcing, incluyendo el consumo energético de las infraestructuras tecnológicas necesarias y las implicaciones de trabajar de manera distribuida a nivel global.

Referencias bibliográficas

- Accenture. (2023). *AI: Built to Scale. From experimental to exponential*. Accenture Applied Intelligence.
- Allen Institute for AI. (s.f.). *About*. <https://allenai.org/about>
- Allen Institute for AI. (2015). *Semantic Scholar*. <https://www.semanticscholar.org>
- Amazon Web Services. (s.f.). *¿Qué es la inteligencia artificial?* Amazon. <https://aws.amazon.com/es/what-is/artificial-intelligence/>
- Angeles, M. del P., & García-Ugalde, F. (2009). *A data quality practical approach*. International Journal on Advances in Software, 2(2&3). <http://www.iariajournals.org/software/>
- Appen. (s.f.). *Case study: Autodesk - Data enrichment*. Appen Success Center. Retrieved from <https://success.appen.com/hc/en-us/articles/202768719-Case-Study-Autodesk-Data-Enrichment>
- Appen. (s.f.). *Airbus: Computer vision and image annotation case study*. Appen. Retrieved from <https://www.appen.com/case-studies/airbus-computer-vision-image-annotation-case-study>
- Appen. (2019). *Appen to acquire Figure Eight to create industry-leading solution for high-quality machine learning training data*. <https://www.appen.com/press-release/appen-to-acquire-figure-eight>
- Batini, C., & Scannapieco, M. (2016). *Data and information quality dimensions*. In Data quality (pp. 1-22). Springer, Cham. https://doi.org/10.1007/978-3-319-24106-7_1
- Bertolazzi, P., Felici, G., Festa, P., Fiscon, G., & Weitschek, E. (2016). *Integer programming models for feature selection: New extensions and a randomized solution algorithm*. European Journal of Operational Research, 250(2), 389–399.
- Binns, R. (2018). *Fairness in Machine Learning: Lessons from Political Philosophy*. Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency, 149-159. <https://doi.org/10.1145/3287560.3287600>
- Brabham, D. C. (2013). *Crowdsourcing*. MIT Press.

- Chandler, J., Mueller, P., & Paolacci, G. (2014). *Nonnaivety among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers*. *Behavior Research Methods*, 46(1), 112-130.
- Chui, M., Manyika, J., Bughin, J., et al. (2012). *The social economy: Unlocking value and productivity through social technologies*. McKinsey Global Institute.
- Cios, K. J., Pedrycz, W., & Swiniarski, R. W. (2012). *Data mining methods for knowledge discovery* (Vol. 458). Springer Science & Business Media.
- Clickworker. (2009, June 24). *About clickworker - Your data and content provider*. <https://www.clickworker.com/about-us/>
- Clickworker. (2011, August 25). *George Eastman House taps Clickworker for iconic global crowdsourcing project*. <https://www.clickworker.com/press-releases/george-eastman-house-taps-clickworker-for-iconic-global-crowdsourcing-project/>
- Clickworker. (2016, December 13). *AI training data: Get original data for your algorithm*. <https://www.clickworker.com/ai-datasets-for-machine-learning/>
- Clickworker. (2017, March 29). *Training data for a face recognition software - case study*. <https://www.clickworker.com/case-studies/training-data-for-a-face-recognition-software/>
- Clickworker. (2018, March 8). *Training data for Chatbots*. <https://www.clickworker.com/case-studies/training-data-for-chatbots/>
- CrowdFlower. (s.f.). *People-Powered Data Enrichment*. https://visit.figure-eight.com/People-Powered-Data-Enrichment_T
- C-SATS. (s.f.). *How it works*. <https://www.csats.com/how-it-works>
- Dastin, J. (2018). *Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
- DataRobot. (2023, September 1). *DataRobot ranked #1 for governance use case by Gartner*. DataRobot. <https://www.datarobot.com/blog/datarobot-ranked-1-for-governance-use-case-by-gartner/>
- Davenport, T. H., & Ronanki, R. (2018). *Artificial intelligence for the real world*. *Harvard Business Review*, 96(1), 108–116.

- Deloitte Spain. (2020, July 22). *Impacto de la IA en las empresas*.
<https://www2.deloitte.com/es/es/pages/strategy-operations/articles/impacto-ia-empresas-nivel-mundial.html>
- Dhirani, L. L., Mukhtiar, N., Chowdhry, B. S., & Newe, T. (2023). *Ethical dilemmas and privacy issues in emerging technologies: A review*. *Sensors* (Basel, Switzerland), 23(3), 1151.
- Diamos, G. (s.f.). *Baidu Research*. <https://www.mturk.com/customers>
- Difallah, D., Filatova, E., & Ipeirotis, P. (2018). *Demographics and dynamics of Mechanical Turk workers*. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 135-143.
- E.A.E Madrid. (2024, January 26). *Explorando el crowdsourcing: Desentrañando sus ventajas*. EAE Business School Madrid. <https://www.eaemadrid.com/es/blog/que-es-crowdsourcing>
- Estelle, A., Gardet, E., Guittet, L., & Larivière, D. (2019). *Comparing Amazon Mechanical Turk and SurveyMonkey for assessing topical expertise of Twitter users*. *Quality & Quantity*, 53(5), 2645-2662.
- Estellés-Arolas, E., & González-Ladrón-De-Guevara, F. (2012). *Towards an integrated crowdsourcing definition*. *Journal of Information Science*, 32(2), 189-200.
- Fealy, L. (2021, February 8). *How crowdsourcing can help reimagine work*. Ernst & Young Global Limited. https://www.ey.com/en_dk/workforce/how-crowdsourcing-can-help-reimagine-work
- Fielding, R. T., & Taylor, R. N. (2002). *Principled design of the modern web architecture*. *ACM Transactions on Internet Technology (TOIT)*, 2(2), 115-150.
<https://doi.org/10.1145/514183.514185>
- Garnett, A. G. (s.f.). *5 ethical questions about artificial intelligence*. Britannica Money. <https://www.britannica.com/money/ai-ethical-issues>
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2018). *Datasheets for Datasets*. arXiv. <https://arxiv.org/abs/1803.09010>
- Google Cloud. (s.f.). *¿Qué es el aprendizaje automático?* Google.
<https://cloud.google.com/learn/what-is-machine-learning?hl=es-419>

- Grand View Research. (2024). *Artificial Intelligence Crowdsourcing Market Size, Share & Trends Analysis Report By Application, By Region, And Segment Forecasts, 2024 - 2028*.
- Hargrave, M. (2010, September 19). *Crowdsourcing: Definition, how it works, types, and examples*. Investopedia. <https://www.investopedia.com/terms/c/crowdsourcing.asp>
- Iberdrola. (s.f.). *¿Qué es el 'Machine Learning' y para qué sirve?* Iberdrola. Retrieved from <https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico>
- IBM. (2023, November 2). *What are large language models (LLMs)*. <https://www.ibm.com/topics/large-language-models>
- IBM. (s.f.). *AI ethics*. <https://www.ibm.com/impact/ai-ethics>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). *Machine learning and deep learning*. *Electronic Markets*, 31(3), 685–695.
- Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing* (3rd ed.). Draft available at https://web.stanford.edu/~jurafsky/slp3/ed3bookaug20_2024.pdf
- Kariluoto, A., Pärnänen, A., Kultanen, J., Soininen, J., & Abrahamsson, P. (2021). *Quality of Data in Machine Learning*. En 2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C) (pp. 216-221). IEEE. <https://doi.org/10.1109/QRS-C55045.2021.00040>
- Kittur, A., Chi, E. H., & Suh, B. (2008). *Crowdsourcing user studies with Mechanical Turk*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Kovashka, A., Russakovsky, O., Fei-Fei, L., & Grauman, K. (2016). *Crowdsourcing in Computer Vision*. *Foundations and Trends in Computer Graphics and Vision*, 10(3), 177-243.
- Lacity, M., & Willcocks, L. (2021). *Becoming strategic with intelligent automation*. *MIS Quarterly Executive*, 20(2), 169–182.
- Lamarre, É. (2023, November 22). *In digital and AI transformations, start with the problem, not the technology*. McKinsey & Company. Retrieved from <https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/in-digital-and-ai-transformations-start-with-the-problem-not-the-technology>
- LeCun, Y., Bengio, Y., & Hinton, G. (2023). *Deep learning for AI*. *Nature*, 521(7553), 436-444.

- Lease, M. (2011). *On quality control and machine learning in crowdsourcing*. Human Computation, 11(11).
- Li, C. (2020, June 3). *OpenAI's GPT-3 Language Model: A Technical Overview*. Lambda. <https://lambdalabs.com/blog/demystifying-gpt-3>
- McCarthy, J. (2007). *What is Artificial Intelligence?* Stanford University. <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html>
- Maione, I. (2022, March 15). *Clickworker receives internationally recognized ISO 27001 certification*. <https://www.clickworker.com/press-releases/clickworker-receives-iso-27001-certification/>
- Manyika, J., & Sneider, K. (2018). *Notes from the AI frontier: Modeling the impact of AI on the world economy*. McKinsey & Company. <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-AI-frontier-modeling-the-impact-of-ai-on-the-world-economy>
- Matsakis, L. (2023, May 22). *A.I. is getting better at generating labels for images. Here's why that's tricky*. The New York Times. <https://www.nytimes.com/2023/05/22/technology/ai-photo-labels-google-apple.html>
- McKinsey & Company. (2024, April 2). *What is generative AI?* <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>
- McKinsey & Company. (2024, April 30). *What is machine learning?* <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-machine-learning>
- Moss, A. J., Rosenzweig, C., Robinson, J., Jaffe, S. N., & Litman, L. (2023). *Is it ethical to use Mechanical Turk for behavioral research? Relevant data from a representative survey of MTurk participants and wages*. Behavior Research Methods, 55(8), 4048–4067. <https://doi.org/10.3758/s13428-022-02005-0>
- Mturk. (s.f.). *Amazon Mechanical Turk*. <https://www.mturk.com/>
- MTurk Guide. (s.f.). *MTurk Guide: Resources and tips for Amazon Mechanical Turk workers*. <https://www.mturkguide.com/>
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.

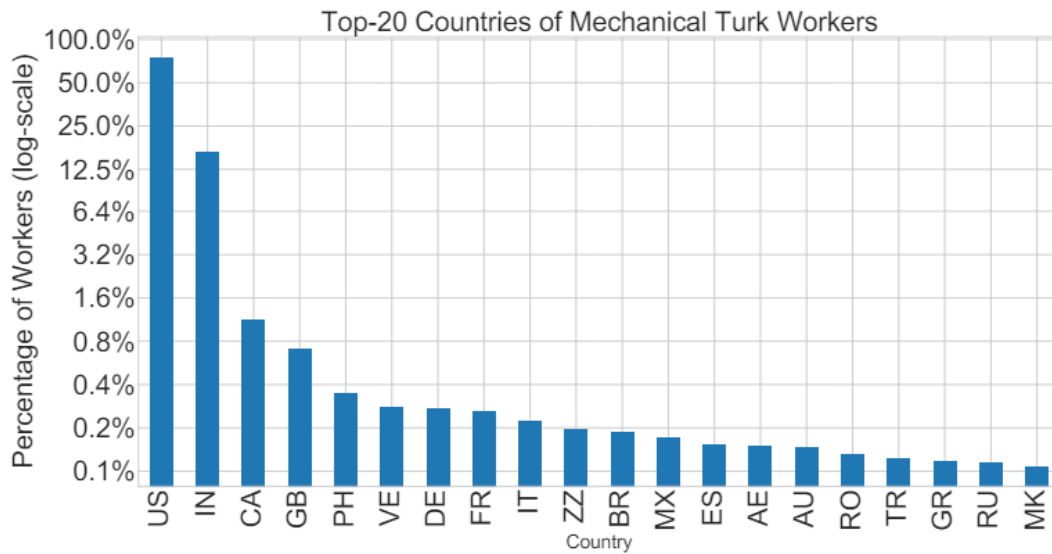
- Nafizah, U. Y., Roper, S., & Mole, K. (2024). *Estimating the innovation benefits of first-mover and second-mover strategies when micro-businesses adopt artificial intelligence and machine learning*. *Small Business Economics*, 62(1), 411–434.
- Olan, F., Spanaki, K., Ahmed, W., & Zhao, G. (2024). *Enabling explainable artificial intelligence capabilities in supply chain decision support making*. *Production Planning & Control*, 1–12.
- Paolacci, G., & Chandler, J. (2014). *Inside the Turk: Understanding Mechanical Turk as a participant pool*. *Current Directions in Psychological Science*, 23(3), 184-188.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). *Beyond the Turk: Alternative platforms for crowdsourcing behavioral research*. *Journal of Experimental Social Psychology*, 70, 153-163.
- Rabiee, M., Mirhashemi, M., Pangburn, M. S., Piri, S., & Delen, D. (2024). *Towards explainable artificial intelligence through expert-augmented supervised feature selection*. *Decision Support Systems*, 181, 114214.
- Reffell, C. (s.f.). *5 formas en las que el crowdsourcing mejora la fiabilidad de ChatGPT*. *Crowdsourcing Week*. <https://crowdsourcingweek.com/blog/chatgpt-reliability/>
- Reffell, C. (s.f.). *Beneficios notables y riesgos simplistas de la anotación de datos mediante crowdsourcing*. *Crowdsourcing Week*. <https://crowdsourcingweek.com/blog/crowdsourcing-data-annotation/>
- Reffell, C. (s.f.). *Cómo LEGO utilizó el crowdsourcing para alcanzar el éxito en el siglo XXI*. *Crowdsourcing Week*. <https://crowdsourcingweek.com/blog/lego-success-through-crowdsourcing/>
- Reffell, C. (s.f.). *Entrenamiento de IA y Deep Learning con crowdsourcing*. *Crowdsourcing Week*. <https://crowdsourcingweek.com/blog/crowdsourcing-ai-training/>
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.) Pearson. p. 1.
- Schmitz, M. (s.f.). *Amazon Mechanical Turk*. <https://www.mturk.com/customers>
- Schneider Electric. (2024). *Annual Reports*. <https://www.se.com/ww/en/about-us/investor-relations/regulatory-information/annual-reports.jsp>
- Stanford University. (2024). *The AI Index 2024 Annual Report*. Stanford Institute for Human-Centered AI. <https://aiindex.stanford.edu/report/>

- Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). *Artificial Intelligence Index Report 2024: Chapter 1 - Research and Development*. Stanford University. <https://aiindex.stanford.edu/>
- Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). *Artificial Intelligence Index Report 2024: Chapter 2 - Technical Performance*. Stanford University. <https://aiindex.stanford.edu/>
- Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). *Artificial Intelligence Index Report 2024: Chapter 3 - Responsible AI*. Stanford University. <https://aiindex.stanford.edu/>
- Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). *Artificial Intelligence Index Report 2024: Chapter 4 – Economy* Stanford University. <https://aiindex.stanford.edu/>
- Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). *Artificial Intelligence Index Report 2024: Chapter 7 – Policy and Governance* Stanford University. <https://aiindex.stanford.edu/>
- Svetlana. (s.f.). *C-SATS*. <https://www.mturk.com/customers>
- Swift Salary. (2023, March 21). *Amazon Mechanical Turk vs Appen*. <https://www.swiftsalary.com/compare-platforms/amazon-mechanical-turk-vs-appen/>
- Swift Salary. (2023, March 21). *Amazon Mechanical Turk vs Clickworker*. <https://www.swiftsalary.com/compare-platforms/amazon-mechanical-turk-vs-clickworker/>
- Task Us. (2023, August 9). *El poder del pueblo: Recopilación de datos colaborativos para un desarrollo exitoso de la IA*. <https://www.taskus.com/insights/crowdsourced-data-collection/>
- Tavakoli, A., Giovine, C., Caserta, J., Machado, J., & Rowshankish, K. (2024, July 8). *Guía técnica de un líder de datos para escalar la IA generativa*. McKinsey & Company. <https://www.mckinsey.com/featured-insights/destacados/guia-tecnica-de-un-lider-de-datos-para-escalar-la-ia-generativa/es>
- Vaughan, J. W. (2017). *Making better use of the crowd: How crowdsourcing can advance machine learning research*. *Journal of Machine Learning Research*, 18(193), 1-46.

- Wakefield, J. (2016). *Microsoft chatbot is taught to swear on Twitter*. BBC News. <https://www.bbc.com/news/technology-35890188>
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., ... & Schwartz, O. (2023). *AI Now 2023 Report*. AI Now Institute, New York University.
- Yandrapalli, V. (2024). *AI-powered data governance: A cutting-edge method for ensuring data quality for machine learning applications*. 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE) (pp. 1-10). IEEE. <https://doi.org/10.1109/IC-ETITE58242.2024.10493601>
- Yao, X. (s.f.). *KITT.AI*. <https://www.mturk.com/customers>

Anexos

#1. Porcentaje de trabajadores en MTurk.



Tomado de: (Difallah et al., 2018).

#2. Satisfacción salarial entre MTurk y Población general USA.

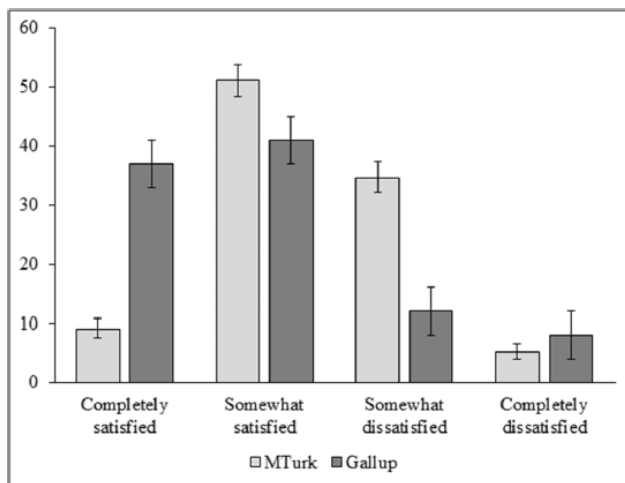


Fig. 9 Pay satisfaction among people on MTurk and in the U.S. population. *Note: Error bars for MTurk represent 95% confidence intervals. Error bars for the Gallup data are based on a maximum $\pm 4\%$ margin of error*

Tomado de: (Moss, Rosenzweig, Robinson, Jaffe, & Litman, 2023)

#3. Implementación de IA en diferentes países y áreas.



Fuente: segunda edición de la encuesta sobre el Estado de la IA en la empresa de Deloitte, 2018.

Deloitte Insights | deloitte.com/insights

Tomado de: (Deloitte Spain, 2020)

#4. Comparativa salarial, opinión.

			Calificación de 1 a 5			
			Valor mínimo para retirar dinero	Velocidad de pago (en días)	Ganancias potenciales	Pago mínimo
Mturk	\$	1,00	3	2,5	4,5	4
Clickworker	\$	10,00	2 a 16	2	3,5	3,5
Figure Eight	\$	5,00	1 a 30	1,5	4	2,5

Tomado de: (Swift Salary, 2023)

#5. Definiciones de crowdsourcing.

Table 3. Collected definitions of crowdsourcing. Source: author

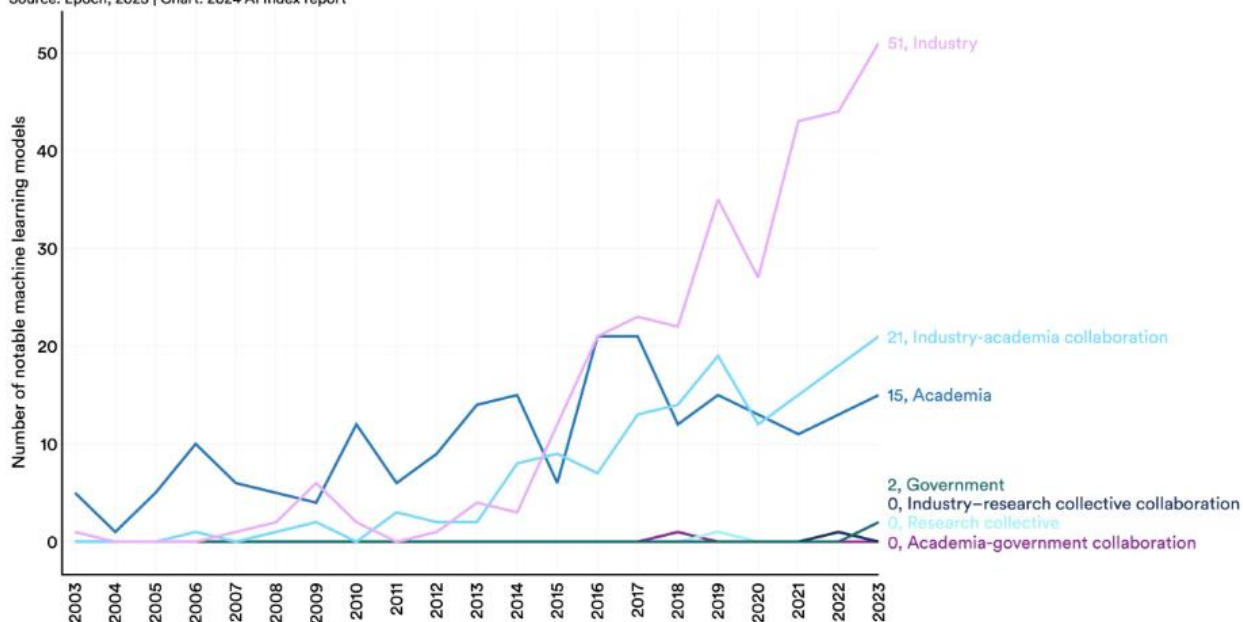
Document	Page	Definition: Crowdsourcing is...
Alonso and Lease [25]	1	... the outsourcing of tasks to a large group of people instead of assigning such tasks to an in-house employee or contractor.
Bederson and Quinn [26]	1	... people being paid to do web-based tasks posted by requestors.
Brabham [9]	75	... an online, distributed problem solving and production model already in use by for profit organizations such as Threadless, iStock...
Brabham [4]	79	... a strategic model to attract an interested, motivated crowd of individuals capable of providing solutions superior in quality and quantity to those that even traditional forms of business can.
Buecheler et al. [11]	1	... a special case of such collective intelligence.
Burger-Helmchen and Penin [10]	2	... one way for a firm to access external knowledge.
Chanal and Caron-	5	... the opening of the innovation process of a firm to integrate numerous and

Tomado de: (Estelles & González-Ladrón-De-Guevara, 2012)

#6. Evolución modelos ML.

Number of notable machine learning models by sector, 2003–23

Source: Epoch, 2023 | Chart: 2024 AI Index report

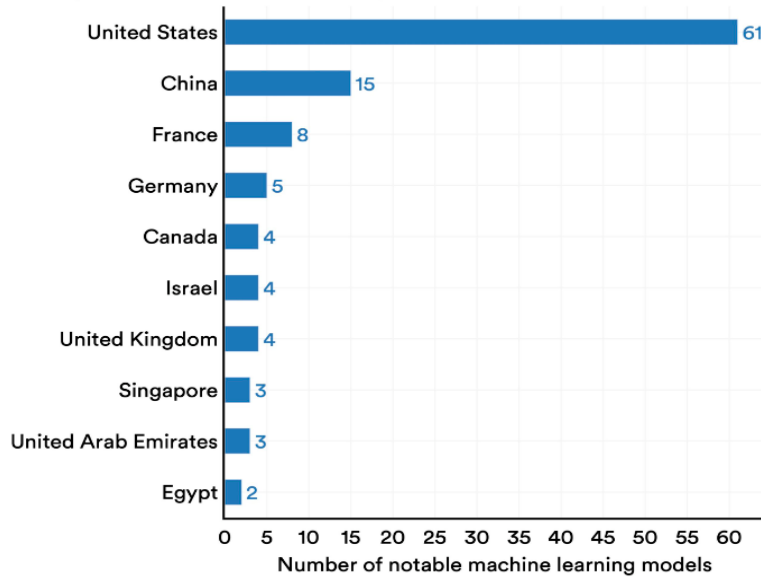


Tomado de: (Stanford University, cap.1, 2024)

#7. Modelos ML por país

Number of notable machine learning models by geographic area, 2023

Source: Epoch, 2023 | Chart: 2024 AI Index report



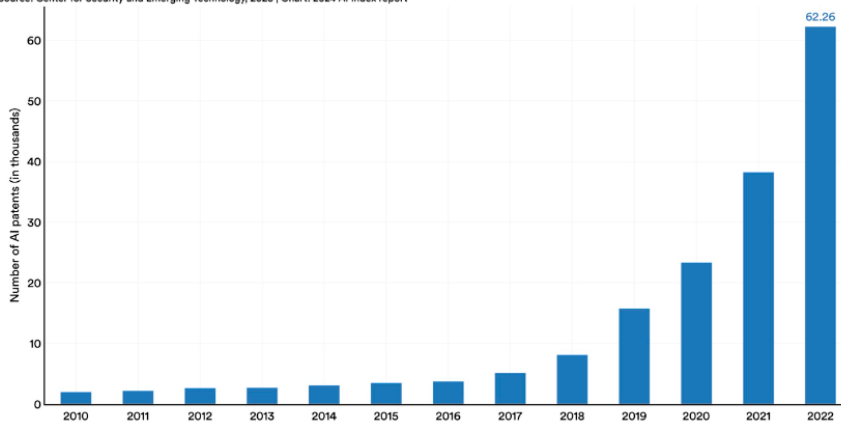
In 2023, 61 notable AI models originated from U.S.-based institutions, far outpacing the European Union's 21 and China's 15.

Tomado de: (Stanford University, cap.1, 2024)

#8. Patentes de IA.

Number of AI patents granted, 2010–22

Source: Center for Security and Emerging Technology, 2023 | Chart: 2024 AI Index report



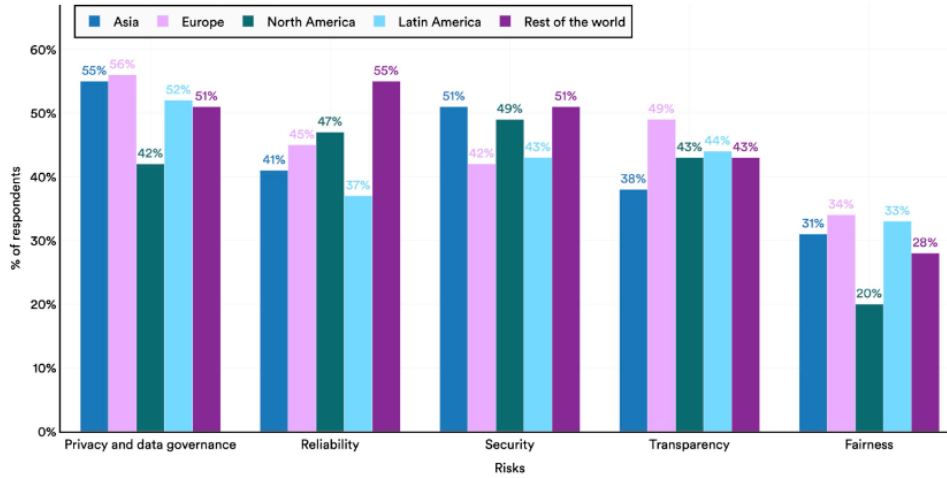
From 2021 to 2022, AI patent grants worldwide increased sharply by 62.7%. Since 2010, the number of granted AI patents has increased more than 31 times.

Tomado de: (Stanford University, cap.1, 2024)

#9. IA y sus preocupaciones.

Relevance of selected responsible AI risks for organizations by region

Source: Global State of Responsible AI report, 2024 | Chart: 2024 AI Index report

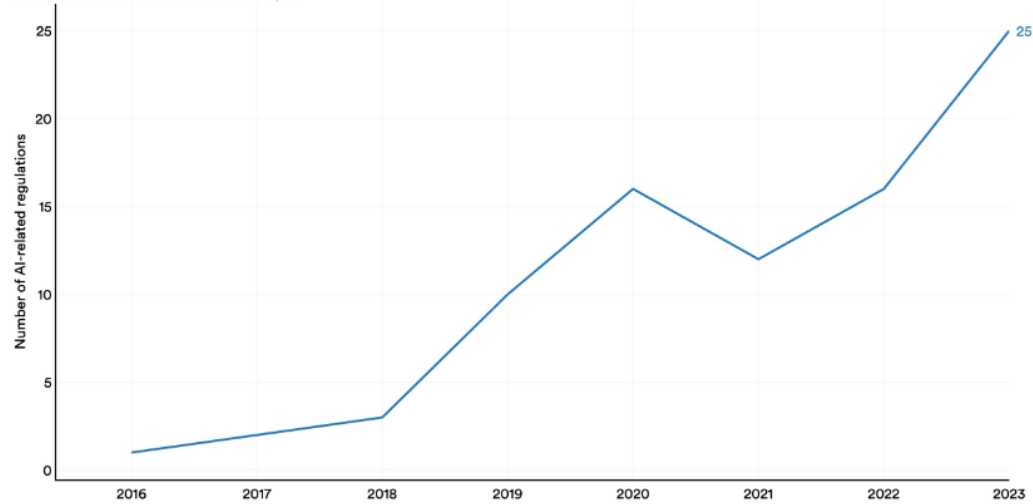


Tomado de: (Stanford University, cap.3, 2024)

#10. Regulaciones de IA en USA.

Number of AI-related regulations in the United States, 2016–23

Source: AI Index, 2024 | Chart: 2024 AI Index report

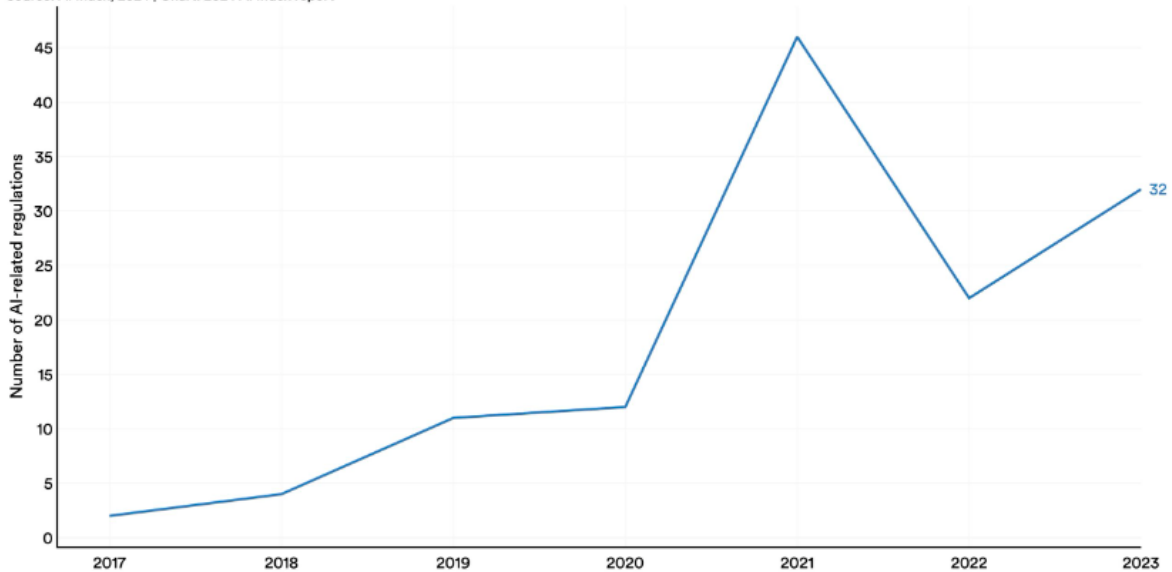


Tomado de: (Stanford University, cap.7, 2024)

#11. Regulaciones de IA en UE.

Number of AI-related regulations in the European Union, 2017–23

Source: AI Index, 2024 | Chart: 2024 AI Index report



Tomado de: (Stanford University, cap.7, 2024)

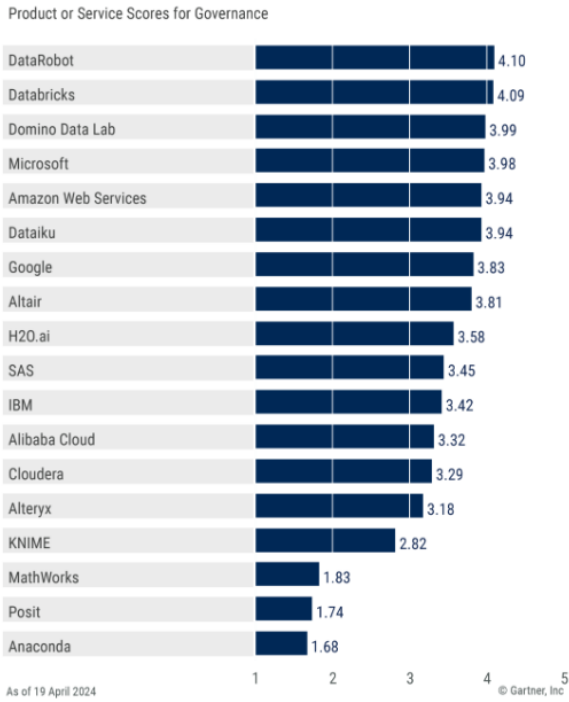
#12. Beneficios del ML.



Tomado de: (Iberdrola, s.f.)

#13. DataRobot y gobernanza

Figure 3: Vendors' Product Scores for Governance Use Case



Tomado de: (DataRobot, 2023)

ANEXO: Objetivos de Desarrollo Sostenible (ODS)



ANEXO I. RELACIÓN DEL TRABAJO CON LOS OBJETIVOS DE DESARROLLO SOSTENIBLE DE LA AGENDA 2030

Anexo al Trabajo de Fin de Grado y Trabajo de Fin de Máster: Relación del trabajo con los Objetivos de Desarrollo Sostenible de la agenda 2030.

Grado de relación del trabajo con los Objetivos de Desarrollo Sostenible (ODS).

Objetivos de Desarrollo Sostenibles	Alto	Medio	Bajo	No Procede
ODS 1. Fin de la pobreza.				X
ODS 2. Hambre cero.				X
ODS 3. Salud y bienestar.				X
ODS 4. Educación de calidad.				X
ODS 5. Igualdad de género.				X
ODS 6. Agua limpia y saneamiento.				X
ODS 7. Energía asequible y no contaminante.				X
ODS 8. Trabajo decente y crecimiento económico.		X		
ODS 9. Industria, innovación e infraestructuras.		X		
ODS 10. Reducción de las desigualdades.			X	
ODS 11. Ciudades y comunidades sostenibles.				X
ODS 12. Producción y consumo responsables.		X		
ODS 13. Acción por el clima.				X
ODS 14. Vida submarina.				X
ODS 15. Vida de ecosistemas terrestres.				X
ODS 16. Paz, justicia e instituciones sólidas.		X		
ODS 17. Alianzas para lograr objetivos.				X

Descripción de la alineación del TFG/TFM con los ODS con un grado de relación más alto.

***Utilice tantas páginas como sea necesario.

Anexo al Trabajo de Fin de Grado y Trabajo de Fin de Máster: Relación del trabajo con los Objetivos de Desarrollo Sostenible de la agenda 2030. (Numere la página)

ODS 8: La tesis habla del desafío que las plataformas de crowdsourcing tienen para mantener la calidad del trabajo y la remuneración justa de las campañas. Esto está directamente relacionado con el ODS 8, que promueve el trabajo decente y el crecimiento económico. La necesidad de garantizar condiciones laborales justas en las plataformas de crowdsourcing es crucial para cumplir con este objetivo.

ODS 9: La tesis expone que las plataformas de crowdsourcing pueden mejorar la calidad y diversidad de datos para entrenar modelos de inteligencia artificial. Esto está relacionado con el ODS 9, que promueve la construcción de infraestructuras resilientes, la promoción de la industrialización inclusiva y sostenible, y el fomento de la innovación. El desarrollo y uso de IA y ML pueden impulsar la innovación en varios sectores industriales.

ODS 10: Aunque no se mencione el tema en la tesis, las plataformas de crowdsourcing pueden contribuir a la reducción de desigualdades al ofrecer oportunidades de trabajo a personas en distintas partes del mundo, incluyendo aquellas en regiones menos desarrolladas. Este aspecto está vinculado con el ODS 10, que busca reducir la desigualdad dentro y entre los países.

ODS 12: El manejo de datos y el uso ético de la inteligencia artificial están relacionados con la producción y consumo responsables. Asegurar que los datos utilizados en IA y ML sean recopilados y manejados de manera ética y sostenible contribuye a los objetivos del ODS 12.

ODS 16: La tesis habla de la importancia de implementar gobernanza y políticas relacionadas con el manejo de datos y la inteligencia artificial. La promoción de instituciones eficaces, responsables y transparentes en todos los niveles está alineada con el ODS 16. Garantizar la privacidad, la seguridad de los datos y la ética en la IA son fundamentales para la construcción de sociedades justas y pacíficas.