El recurso de la inteligencia artificial en el ámbito laboral de forma ética y su repercusión en los derechos y libertades de los trabajadores.

Facultad de Derecho ICADE

Tutor: José María Lassalle Ruíz
Autor: Alfonso de Aguilar de la Concha
Curso 5º E3 A

Índice (PROVISIONAL)

La IA, ¿ha llegado para añadir valor a las funciones de un trabajador o para anular las mismas? La IA llega para formar parte de toda la actividad empresarial y su implantación está suponiendo un cambio drástico en la forma de mantener las relaciones laborales. Esta implementación puede provocar tensiones éticas y albergar riesgos de desigualdad que carecen de regulación.

LISTADO DE ABREVIATURAS

……

INTRODUCCIÓN

1. PLANTEAMIENTO, JUSTIFICACIÓN DE LA CUESTIÓN Y OBJETIVOS DE LA INVESTIGACIÓN

En este momento, nos enfrentamos a la oportunidad de diseñar una IA que acelere la inspiración humana, la creatividad y promueva la equidad en el mundo. Pero esa elección no ocurre por defecto. Está hecho por gente como tú y como yo. Y una vez construidos, no podemos confiar simplemente en que los sistemas de IA tengan la obligación de actuar de forma ética. En cambio, la ética es y siempre será una función humana.

La IA generativa nos ha envuelto. Cada mes, se lanzan nuevos modelos con nuevas características, con nuevas capacidades emergentes, y los humanos han aprendido a usar estas herramientas de manera cada vez más efectiva. Desde los aspectos positivos, el uso en nuestros lugares de trabajo y cómo nos comunicamos entre nosotros incluso de alguna manera. Y, por otro lado, estamos viendo que la gente empieza a utilizar estas herramientas de formas menos maravillosas, utilizando la ***tecnología deepfake*** para afectar a las elecciones y al discurso político, o a los robots que llaman y a los estafadores que utilizan GenAI para intentar estafarnos.

Ahora bien, nada de esto es inesperado, pero a medida que vemos estas nuevas aplicaciones, tenemos que volver al núcleo de cómo pensamos sobre la IA generativa a través de una lente ética. ¿Promueven estas herramientas nuestros intereses humanos? ¿Se aseguran de que tengamos más agencia en el mundo? Y a medida que se nos ocurren formas nuevas e innovadoras de utilizar las herramientas de GenAI, ¿nos estamos basando en una comprensión de la dignidad, la equidad y la justicia? A medida que avanzamos en un mundo donde la GenAI está en todas partes, debemos fortalecer esa capacidad de análisis ético. Cuando veas una nueva noticia o aplicación de GenAI, espero que te preguntes: "¿A quién ayuda esto? ¿A quién defiende esto? ¿A quién apoya esto? ¿Cómo nos hace a todos más humanos?"

Soy Alfonso Aguilar y he dedicado este trabajo de fin de grado para **delimitar las fronteras de la innovación tecnológica** y, al mismo tiempo, a **abogar por un mundo más justo y equitativo** porque nuestras decisiones de hoy darán forma al futuro de la IA durante décadas.

3. METODOLOGÍA Y ESTRUCTURA DE LA INVESTIGACIÓN .....

CAPÍTULO I. MARCO TEÓRICO

1. DELIMITACIÓN CONCEPTUAL

1.1 *¿Qué es la inteligencia artificial?*

La inteligencia artificial (IA) es un campo de la informática que busca desarrollar sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como aprender, razonar y tomar decisiones. Estos sistemas utilizan algoritmos y aprenden de los datos para mejorar su desempeño. A diferencia de las personas, los dispositivos basados en IA no necesitan descansar y pueden procesar grandes cantidades de información. Estas tecnologías están siendo utilizadas en diversos campos para mejorar la eficiencia, pero también debemos ser conscientes de los posibles efectos negativos que pueden tener.

La innovación tecnológica altera la organización de la sociedad, la política y la economía, ya que modifica las condiciones de funcionamiento de la sociedad. La idea de la destrucción creativa como un motor del cambio económico sostiene que el progreso económico surge a través de la creación de nuevas tecnologías y productos, que destruyen las antiguas formas de producción y crean nuevas oportunidades de negocio. Se puede aplicar a la forma en que la IA está cambiando la forma en que las empresas operan y compiten en el mercado, así como revolucionando la forma en que se hacen negocios, desde la automatización de tareas rutinarias hasta la creación de nuevos productos y servicios. Por lo tanto, el impacto de la IA en la economía puede ser disruptivo, destruyendo antiguas formas de producción y creando nuevas oportunidades de negocio, pero también puede tener un impacto en los ciclos económicos al aumentar la productividad y la eficiencia, lo que a su vez puede estimular el crecimiento económico.

1.2 ¿Qué es la ética y la moral?

La ética es aquella parte de la filosofía que se dedica a la reflexión sobre la moral. Es un tipo de saber que intenta construirse racionalmente, utilizando para ello el rigor conceptual, los métodos de análisis y explicación propios de la Filosofía. Pretende desplegar los conceptos y argumentos que permitan comprender la dimensión moral de la persona humana en cuanto tal dimensión moral, es decir, sin reducirlas a sus componentes psicológicos, sociológicos, económicos o de cualquier otro tipo.

También la moral es un saber que ofrece orientaciones para la acción, pero mientras ésta última, propone acciones concretas en casos concretos, la Ética, como Filosofía moral, se remonta a la reflexión sobre las distintas morales y sobre los distintos modos de justificar racionalmente la vida moral, de modo que su manera de orientar la acción es indirecta.

Ética y moral en el contexto de la Inteligencia Artificial

Ahora bien, la Filosofía Jurídica, la cual será el centro de nuestro estudio, tiene por objetivo regular la disciplina y realidad jurídica existente, realizando lo anterior mediante la investigación y análisis de las causas últimas y propias del Derecho, llegando a los fundamentos últimos, fines, utilidad y esencia de los fenómenos jurídicos. La ética jurídica se encuentra, en este orden de ideas dentro de la Filosofía Jurídica, pues tiene por objeto de estudio verificar en qué medida las normas jurídicas se ajustan o no a los valores morales. Por ello que la ética jurídica desempeña un papel fundamental en la evaluación y regulación del desarrollo y uso de la Inteligencia Artificial (IA). En un contexto donde la IA impacta cada vez más en la sociedad, la dimensión ética se vuelve indispensable para garantizar que estas tecnologías se implementen de manera justa, transparente y alineada con los valores humanos fundamentales.

Por consiguiente, la moral, entendida como el conjunto de normas y principios que guían nuestras acciones concretas, encuentra un campo de aplicación inmediata en la IA. Por ejemplo, los sistemas de IA utilizados en decisiones críticas, como diagnósticos médicos, contratación laboral o justicia predictiva, deben operar bajo principios morales que garantizan la equidad, la no discriminación y el respeto a los derechos humanos.

Desde sus orígenes en la Grecia clásica, la ética ha tenido un propósito normativo: guiar las acciones humanas. En el ámbito de la IA, esta función se manifiesta en la creación de marcos éticos que orienta su diseño, implementación y regulación. Por ejemplo:

- La elaboración de principios éticos globales para la IA, como los propuestos por la UNESCO, establece valores como la transparencia, la responsabilidad y la protección de la privacidad.
- Normas específicas, como las desarrolladas en la Unión Europea sobre IA confiable, buscan garantizar que los sistemas sean seguros, explicables y no causen daño.

Por ello, es necesario que la ética, como reflexión filosófica más profunda, evalúe críticamente estas aplicaciones morales para identificar y cuestionar los posibles sesgos y limitaciones inherentes a los sistemas de IA. Por ejemplo, si un algoritmo discrimina a ciertos grupos sociales, la ética puede cuestionar las bases de su entrenamiento y exigir cambios para alinearse con valores universales.

2. TRABAJADORES VS ROBOTS INTELIGENTES

“La tecnología responsable del comportamiento humano”

La inteligencia artificial generativa está revolucionando todos los aspectos de la experiencia humana. Me considero una persona optimista que cree firmemente que estas herramientas nos harán más humanos, más creativos, inspirados y conectados. Por ejemplo, en el sector bancario, la IA y el aprendizaje automático están ayudando a identificar personas que podrían beneficiarse de mayores niveles de seguridad financiera, fomentando el ahorro, aumentando las contribuciones a la jubilación y facilitando mejores caminos hacia oportunidades económicas.

En el ámbito de la agricultura, los modelos de IA están transformando las prácticas agrícolas al predecir eventos climáticos importantes. Esto permite a los agricultores y productores tomar decisiones más informadas, como determinar el momento idóneo para contratar un seguro adicional o saber cuándo plantar y cosechar para maximizar sus rendimientos económicos. Dentro de las organizaciones, la IA también está reformulando los recursos humanos, ayudando a los gerentes a inspirar un mejor desempeño en sus equipos, identificar y corregir posibles sesgos o discriminación, y garantizar que los ascensos se basen en el mérito.

Sin embargo, aunque estoy entusiasmado con las oportunidades que la IA nos presenta, también soy consciente de la necesidad de construir estas herramientas con intención positiva, fundamentadas en principios éticos y responsables. Esta preocupación no es exclusiva mía. Desde sus inicios, los desarrolladores de inteligencia artificial han sido conscientes del enorme poder de estas tecnologías y de los dilemas éticos que pueden surgir al implementarlas. En los últimos años, estas preocupaciones han cobrado mayor relevancia. Investigadores y responsables políticos han expresado inquietudes sobre cómo la IA podría perpetuar sesgos, aumentar la desigualdad y hacerlo de formas que podrían pasar desapercibidas.

Estas preocupaciones éticas se están manifestando de manera tangible. Por ejemplo, la aparición de los **deepfakes** plantea riesgos significativos, como la posibilidad de crear avatares que suplanten a personas de confianza para difundir información fraudulenta, engañar a clientes o incluso incitar acciones peligrosas. Los **chatbots** también representan un desafío, ya que, sin un diseño ético adecuado, pueden proporcionar consejos erróneos a médicos, estudiantes u otros usuarios, basados en datos inexactos o falsos que no han sido auditados por humanos.
Por último, también surgen cuestiones sobre propiedad intelectual, derechos de autor y responsabilidad legal. Es imprescindible abordar estos retos con seriedad para garantizar que la IA beneficie a la sociedad de manera justa y segura, permitiendo que su potencial transformador se desarrolle de manera ética y responsable.

2.1 El marco ético de la IA.

Tenemos que asegurarnos de que estamos diseñando herramientas que respalden el futuro que queremos crear, siendo éstas equitativas, sostenibles y prósperas. Y para ello, vamos a tener que idear nuevos marcos para la creación ética con la misma rapidez con la que avanzamos en la frontera de la innovación. Los tres pilares del marco ético a considerar son las prácticas responsables en materia de datos, unos límites bien definidos sobre el uso seguro y adecuado y una transparencia sólida.

La **práctica responsable de los datos** es el pilar central sobre el cual se debe construir cualquier herramienta ética de inteligencia artificial (IA). Este principio enfatiza que la ética de una tecnología depende en gran medida de la calidad y representatividad de los datos con los que ha sido entrenada. Por ello, es crucial adoptar enfoques rigurosos para garantizar que los datos utilizados sean diversos, imparciales y reflejen de manera justa las realidades y necesidades de todas las poblaciones a las que la tecnología podría impactar. Si el conjunto de datos de entrenamiento está sesgado o carece de diversidad, los modelos de IA podrían perpetuar o incluso amplificar esos sesgos. Por ejemplo, si la mayoría de los datos provienen de consumidores de una raza o género específico, los resultados podrían estar diseñados para satisfacer principalmente las necesidades de esa población, dejando de lado a otros grupos. Esto no solo resulta éticamente problemático, sino también ineficaz desde una perspectiva de inclusión y accesibilidad.

Por lo tanto, al considerar la creación o implementación de cualquier herramienta nueva, debe preguntarse: ¿cuál es la fuente de los datos de entrenamiento? ¿Qué se ha hecho para reducir el sesgo explícito e implícito en ese conjunto de datos? ¿Cómo podrían los datos que estamos utilizando perpetuar o aumentar el sesgo histórico? ¿Y qué oportunidades hay para prevenir la toma de decisiones sesgadas en el futuro? Y, por consiguiente, para reducir el riesgo de perpetuar inequidades, es esencial implementar prácticas proactivas que promuevan la equidad en todas las etapas del desarrollo y despliegue de la IA. Esto incluye:

1. Auditorías periódicas: Evaluar regularmente el rendimiento del modelo para identificar resultados injustos o inconsistentes entre diferentes grupos.
2. Colaboración interdisciplinaria: Trabajar con expertos en ética, ciencias sociales y derechos humanos para identificar posibles áreas de riesgo.
3. Transparencia en los datos: Documentar claramente los procesos de recolección y tratamiento de los datos, así como las decisiones tomadas para mitigar el sesgo.

Adoptar una práctica responsable de los datos no solo mejora la calidad y la equidad de los sistemas de IA, sino que también fortalece la confianza pública en estas tecnologías. Esto permite construir herramientas que no solo sean

innovadoras, sino también justas e inclusivas, generando un impacto positivo en la sociedad.

La segunda parte del marco es la importancia de crear límites bien definidos para usos seguros y éticos. Cualquier nueva herramienta o aplicación de IA debe comenzar con una declaración de intenciones centrada en los objetivos de la organización y una identificación de la población a la que intentamos servir. Establecer límites claros sobre cómo se puede utilizar la IA en el entorno laboral ayuda en primer lugar a proteger los derechos de los trabajadores garantizando que la IA no sea utilizada para prácticas discriminatorias, invasión de la privacidad o evaluación injusta del desempeño laboral. Y, en segundo lugar, ayuda a delimitar el alcance de las decisiones automatizadas definiendo qué decisiones pueden ser tomadas por sistemas de IA y cuáles deben permanecer bajo control humano, especialmente aquellas que afectan significativamente la vida de las personas.

La tercera parte del marco es una **transparencia sólida**. La transparencia es un componente esencial para fomentar la confianza de los trabajadores y otras partes interesadas en el uso de la IA. Esto implica proporcionar claridad sobre:

- Cómo funciona la IA: Explicar de manera accesible los algoritmos y procesos detrás de las decisiones automatizadas, incluyendo los criterios utilizados.
- Qué datos se utilizan y cómo se protegen: Informar a los empleados sobre los datos recopilados, su propósito y las medidas implementadas para garantizar su seguridad y privacidad.
- Responsabilidades y rendición de cuentas: Establecer quién es responsable de supervisar el uso de la IA y cómo se manejarán los errores o impactos negativos derivados de su implementación.

Hacer más transparente el uso de la IA también es esencial para identificar posibles efectos discriminatorios generados por los algoritmos, así como para validar si su aplicación responde al interés público, así como para atribuir responsabilidades sobre las decisiones tomadas por sistemas de IA.

Para fomentar esta transparencia, se recomienda que los desarrolladores y quienes implementen sistemas de IA compartan información clave, como el propósito del uso, detalles del código fuente, los datos utilizados, las limitaciones de los sistemas y su posible impacto.

Otro enfoque importante es proporcionar explicaciones en un lenguaje claro y accesible para todas las partes interesadas. Esto incluye, por ejemplo, aplicar principios como el derecho a la explicación reconocido en el ***Reglamento General de Protección de Datos (RGPD)***, que otorga a las personas la posibilidad de comprender, cuestionar y desafiar decisiones automatizadas que las afecten.

2.2 Impacto de la Inteligencia Artificial en los Derechos Laborales

El uso de inteligencia artificial puede vulnerar los derechos fundamentales de las personas, incluso de manera no intencionada, cuando esta tecnología se usa sin aplicar o tener en cuenta valores éticos, sin prever realmente los derechos que pueden verse afectados o las consecuencias de ello (es decir, sin haber hecho un análisis de impacto previo).

Cabe recordar que los derechos fundamentales derivan, en gran medida, de la ***Declaración Universal de Derechos Humanos*** y teniendo en cuenta cómo se entrenan, funcionan y la finalidad que tienen algunos sistemas de IA, es innegable que siempre existirá una relación entre el uso de inteligencia artificial y derechos fundamentales, donde la tecnología tendrá un potencial doble: por un lado, podrá contribuir a la mejora de la sociedad, y, por otro lado, también podrá afectar negativamente a estos derechos, incluso violarlos, como veremos en los siguientes puntos.

- *Derecho a la intimidad, al honor y a la propia imagen*

El derecho a la intimidad, al honor y a la propia imagen es uno de los derechos fundamentales más amenazado por la inteligencia artificial, puesto que para conseguir que determinados modelos de IA sean eficientes y eficaces, es necesario recabar grandes cantidades de datos personales, datos que pueden incluir todo tipo de información personal, médica, financiera, de comportamiento, imágenes, etc.

Estos datos son recogidos de forma masiva, a veces sin el consentimiento explícito e informado de los interesados, lo que ya de por sí es una violación de la privacidad personal.

Además, de usar los datos personales para el entrenamiento de los modelos de IA, no son pocas las empresas que ya emplean herramientas de IA para analizar patrones de comportamiento a partir de esos datos, en muchas ocasiones sin el conocimiento de los interesados o sin informarles debidamente.

Así mismo, ya hay herramientas y aplicaciones de IA que su solo uso puede vulnerar el derecho a la imagen; por ejemplo, las aplicaciones de IA que pueden usarse para «desnudar» personas y que se emplean para crear deepfakes pornográficos, suponen una vulneración del derecho de imagen y del derecho al honor, o las IA capaces de clonar voz vulneran nuestros derechos de voz (que son parte de nuestro derecho a la imagen).

No podemos dejar de mencionar también que el reconocimiento facial ha avanzado mucho gracias al uso de la inteligencia artificial, de manera que cámaras y programas con capacidad para el reconocimiento de rostros pueden

emplearse para seguir a las personas y vigilarlas y, sin las salvaguardas legales y éticas adecuadas, acabar vulnerando el derecho a la intimidad y a la libertad de movimientos, incluso de manifestación.

- *Derecho a la protección de datos*

Estrechamente ligado al derecho a la intimidad, el derecho a la protección de datos es otro derecho fundamental que puede verse vulnerado por el uso de la inteligencia artificial en los siguientes aspectos:

I. La falta de transparencia en la IA implica que los interesados no sepan para qué y cómo se usan sus datos personales en estos sistemas, más allá de las explicaciones, normalmente insuficientes, que aparecen en la política de privacidad.
II. Esa falta de transparencia también implica que los procesos de toma de decisiones automatizadas pueden resultar incomprensibles para el interesado, que no entiende por qué un sistema ha tomado una decisión concreta que puede tener consecuencias jurídicas o legales para él.
III. Riesgos de seguridad que pueden poner en peligro la confidencialidad, integridad y disponibilidad de los datos, puesto que los sistemas de IA también son susceptibles de sufrir ciberataques; si los datos personales que usan o almacenan no están debidamente protegidos, podrían filtrarse y usarse con fines maliciosos.
IV. Pérdida de control sobre los datos personales.

- *Derecho a la no discriminación y a la igualdad de oportunidades*

Aunque la inteligencia artificial tiene como fin ser objetiva, puede perpetuar y amplificar la discriminación existente, cuando los algoritmos han sido entrenados con datos que reflejan sesgos humanos históricos o se diseñan sin tener en cuenta la diversidad de la población.

Esta discriminación causada por el sesgo algorítmico está bien documentada y conduce a que determinados grupos de población y demográficos sean discriminados por el sistema, especialmente cuando se confía en él para la toma de decisiones. Y esto es algo que puede ocurrir en diferentes ámbitos, desde el empleo hasta la seguridad y la justicia.

Por ejemplo, un banco puede usar un sistema automático para la concesión de hipotecas basadas en los datos del solicitante; si ese sistema basado en IA ha sido entrenado con conjuntos de datos sesgados (por ejemplo, el histórico de clientes y solicitantes, en el que la mayoría son hombres blancos, de una edad entre los 30 y 50 años), discriminará a cualquier solicitante que se salga de esos datos (en el ejemplo, mujeres, personas de otras etnias u hombres menores de 30 años).

Una solución fundamental es entrenar los algoritmos con conjuntos de datos diversos y representativos que reflejan equitativamente factores como género, etnia y edad, disminuyendo así la posibilidad de que se perpetúen patrones discriminatorios. Además, las auditorías regulares, tanto éticas como técnicas, son esenciales para identificar y corregir sesgos de manera oportuna. Estas auditorías deben ser realizadas por equipos multidisciplinarios que incluyan tanto expertos técnicos como especialistas en derechos humanos.

El uso de métricas de equidad es otra clave estratégica, ya que permite evaluar si los resultados de los sistemas de inteligencia artificial se benefician de manera proporcional a todos los grupos. En decisiones críticas, como las relacionadas con contrataciones o concesiones de préstamos, se debe garantizar la intervención humana, asegurando que un profesional capacitado supervise o revise el proceso para evitar errores o discriminaciones no intencionadas.

La transparencia y la explicabilidad de los sistemas también son fundamentales. Diseñar algoritmos que puedan justificar sus decisiones permite que las personas afectadas comprendan el razonamiento detrás de un resultado y, si es necesario, lo cuestionen o lo impugnen. Paralelamente, es crucial capacitar a desarrolladores, ingenieros y tomadores de decisiones en temas de sesgos y diversidad, fomentando así una mayor conciencia sobre la importancia de diseñar tecnologías inclusivas.

- *Derecho a la justicia*

Ya se estudia la aplicación de la inteligencia artificial en el sistema judicial, lo que, unido a los citados sesgos, puede vulnerar el derecho a un juicio justo de las personas, así como provocar sentencias injustas y, en el caso más extremo, aplicar la ley de forma preventiva (es decir, usar la IA para predecir qué posibilidades tiene una persona de cometer un delito y actuar antes de ello; esto puede sonar ahora a ciencia ficción, pero los sistemas de puntuación social que ya se emplean en China, así como el monitoreo por las cámaras de seguridad, son un paso en esa dirección).

- *Derecho a la información y a la libertad de expresión*

La inteligencia artificial también puede usarse para manipular la opinión pública al crear de manera más sencilla bulos creíbles y hacerlos virales en mucho menos tiempo, empleando para su distribución y difusión granjas de bots y «medios» online sin respeto por la ética ni la deontología periodística. Esto puede vulnerar el derecho a la información al saturar internet de contenidos falsos o desinformación, haciendo que al final los usuarios duden de la veracidad de cualquier contenido, incluidos los veraces, lo que puede acabar dañando también a la democracia

Así mismo, la IA ya se emplea para moderar contenidos en redes sociales, de manera que las publicaciones que vulneren los términos y condiciones de la

plataforma sean retiradas de forma automática. Sin embargo, estos algoritmos no son infalibles y pueden acabar censurando contenido legítimo.

Automatizar la censura en internet puede restringir la capacidad de los individuos para expresarse libremente, algo que se agrava cuando no se explica debidamente por qué se retira dicho contenido y los mecanismos para impugnar esas censuras no son claros o resulta engorroso usarlos.

- *Derecho a la privacidad*

La monitorización en el lugar de trabajo puede ser una herramienta valiosa para mejorar la eficiencia y la seguridad, pero debe equilibrarse cuidadosamente con el derecho a la privacidad de los empleados. Para garantizar que no se vulneren derechos fundamentales, estas prácticas deben regirse por principios éticos, legales y transparentes, cumpliendo con límites y consideraciones clave.

En primer lugar, es indispensable respetar los límites legales y regulatorios establecidos por normativas como el Reglamento General de Protección de Datos (RGPD) en Europa. Estas leyes especifican requisitos claros sobre la recopilación, uso y almacenamiento de datos. Por ejemplo, los empleados deben ser informados de manera clara y previa sobre qué datos serán recopilados y con qué propósito, asegurándose de que únicamente se recaben los datos estrictamente necesarios para cumplir con objetivos legítimos. Además, estas normativas otorgan a los empleados el derecho de acceder, corregir o eliminar sus datos personales. Un componente esencial de este marco legal es el consentimiento informado, que debe ser explícito, voluntario y no condicionado a la continuidad del empleo, garantizando así que los empleados estén plenamente conscientes de las prácticas de monitorización.

Desde un enfoque ético, la proporcionalidad juega un papel fundamental. Esto significa que las medidas de monitorización deben ser proporcionales al objetivo que se desea alcanzar. Por ejemplo, no se justifica emplear cámaras o software de seguimiento altamente invasivos si con informes generales de productividad es suficiente. Además, la transparencia es clave, las empresas deben comunicar de manera clara las políticas de monitorización, explicando cómo y por qué se recopilan los datos, así como las medidas de protección implementadas para garantizar la privacidad de los empleados.

Por lo tanto, después de este recorrido por todos los derechos que pueden ser vulnerados, no cabe duda que la Inteligencia artificial y los derechos humanos deben convivir, y el desarrollo y uso de la primera debe respetar a los segundos, por ello es fundamental no solo que los gobiernos legislen y promulguen leyes como la Ley de Inteligencia Artificial europea, sino que las propias empresas que crean modelos y sistemas de IA y aquellas que los usan y los ponen a disposición de sus clientes, adopten un enfoque ético, responsable y transparente.

2.3 La Deshumanización del Trabajo: Un Riesgo de la IA (PROVISIONAL)

……

Introducción
La automatización impulsada por la inteligencia artificial (IA) está reemplazando progresivamente a los trabajadores en sectores que dependen de tareas repetitivas o altamente estructuradas. Este cambio plantea un dilema ético sobre cómo equilibrar los beneficios económicos de la tecnología con el impacto social y humano de la pérdida de empleos. Desde una perspectiva ética, la implementación de IA en el ámbito laboral debe considerar no solo la eficiencia empresarial, sino también el bienestar de los trabajadores y la equidad en el acceso al trabajo.

Problemáticas éticas clave

1. **El dilema de la eficiencia frente a la justicia social**
    - **Problema ético:**
        - La automatización mediante IA permite a las empresas reducir costos y aumentar la productividad, pero esto a menudo se logra a expensas de los trabajadores más vulnerables. Despedir empleados para sustituirlos por sistemas automatizados puede generar desigualdades sociales al dejar a muchos sin sustento económico.
    - **Perspectiva ética:**
        - Desde el punto de vista de la ética utilitarista, los beneficios de la automatización para la sociedad (eficiencia, menor costo de bienes y servicios) deben contrapesarse con el daño causado a los trabajadores afectados.
        - Desde una ética basada en derechos, los trabajadores tienen derecho a un trato digno y justo, lo cual incluye acceso al empleo y protección frente a cambios drásticos no anticipados.
2. **Desigualdad de oportunidades**
    - **Problema ético:**
        - La IA favorece la automatización de trabajos de baja cualificación, afectando desproporcionadamente a los trabajadores con menos acceso a educación o formación tecnológica. Esto refuerza ciclos de desigualdad social.
    - **Perspectiva ética:**
        - Una sociedad justa debe garantizar igualdad de oportunidades. Implementar la IA sin un plan para capacitar a los trabajadores desplazados contradice el principio ético de justicia distributiva, que busca una distribución equitativa de recursos y oportunidades.
3. **Deshumanización del trabajo**
    - **Problema ético:**

        - Sustituir humanos por máquinas no solo elimina empleos, sino que también deshumaniza el concepto de trabajo, reduciéndolo a métricas de productividad y ganancias económicas, sin considerar el valor intrínseco del empleo para la dignidad y el desarrollo personal.
    - **Perspectiva ética:**
        - Según la ética kantiana, los trabajadores deben ser tratados como fines en sí mismos y no meros medios para obtener beneficios económicos. El uso de IA que prioriza la eficiencia sobre el bienestar humano viola este principio.
4. **Responsabilidad social de las empresas**
    - **Problema ético:**
        - Las empresas que implementan IA a menudo priorizan los beneficios a corto plazo sobre sus responsabilidades hacia los empleados y las comunidades afectadas por el desempleo.
    - **Perspectiva ética:**
        - Según la ética de la responsabilidad social corporativa, las empresas tienen la obligación moral de contribuir al bienestar de la sociedad, lo que incluye mitigar los efectos negativos de sus decisiones tecnológicas sobre los trabajadores.

Propuestas éticas para abordar la precarización laboral

1. **Transición justa hacia la automatización:**
    - Las empresas y gobiernos deben garantizar que los trabajadores afectados por la IA reciban apoyo para reinsertarse en el mercado laboral.
    - Propuesta ética: Implementar programas de formación y reentrenamiento que permitan a los trabajadores desarrollar nuevas habilidades.
2. **Regulación para mitigar el impacto social:**
    - Establecer leyes que limiten la automatización en sectores sensibles hasta que se implementen políticas de transición laboral.
    - Propuesta ética: Diseñar regulaciones que consideren el impacto social y garanticen la consulta previa a los trabajadores.
3. **Renta básica universal como medida ética:**
    - La implementación de una renta básica podría compensar el impacto del desempleo tecnológico y garantizar un nivel mínimo de subsistencia para todos.
    - Propuesta ética: Desde el punto de vista de la justicia distributiva, esta medida asegura que los beneficios económicos de la IA se repartan de manera equitativa entre toda la sociedad.
4. **Promoción de empleos centrados en lo humano:**
    - Fomentar sectores donde el trabajo humano tenga un valor añadido que no puede ser reemplazado por máquinas, como la creatividad, el cuidado de personas o la educación.
    - Propuesta ética: Reconocer el valor intrínseco del trabajo humano y su contribución al desarrollo personal y social.

CAPÍTULO II. RETOS JURÍDICOS.

1. FALTA DE REGULACIÓN ESPECÍFICA

1.1 Leyes y regulación: Ley de IA de la UE, primera normativa sobre inteligencia artificial

1.2 Posición de la Unión Europea en materia de IA

2. ¿QUIÉN ES EL PROPIETARIO DE LAS IDEAS Y PRODUCTOS CREADOS POR LA IA GENERATIVA? (PROVISIONAL)

……

El propietario de las ideas y productos creados por la IA generativa abre un debate importante sobre derechos de autor, ética y responsabilidad. En principio, las ideas y productos generados por la IA suelen ser considerados propiedad de quien posee o controla la herramienta de IA, ya que ésta opera como un instrumento. Sin embargo, la cuestión se complica cuando el contenido creado refleja sesgos o toma decisiones que afectan a las personas de manera tangible.

Imaginemos que su empresa ha implementado un sistema de IA para respaldar la función de recursos humanos, escaneando currículums de solicitantes para identificar posibles entrevistados. A primera vista, esta herramienta parece funcionar excepcionalmente bien, procesando información rápidamente y proporcionando una cantidad similar de candidatos a los que seleccionaría un equipo humano. Sin embargo, un análisis más profundo revela patrones preocupantes: la herramienta da prioridad a ciertos géneros, vecindarios o patrones en los historiales laborales. Estos sesgos son una réplica de los prejuicios humanos presentes en el conjunto de datos que entrenaron el sistema.

Aquí surge un reto importante. Si bien el algoritmo está proporcionando recomendaciones sesgadas, las decisiones finales sobre a quién entrevistar recaen en los humanos. En este contexto, debemos cuestionarnos:

***1*. ¿Cuál es el estándar más alto de comportamiento humano responsable?**

Esto implica reflexionar sobre cómo garantizar que las herramientas de IA no solo sean eficientes, sino también éticas y justas. El estándar más alto debe incluir la revisión constante de los resultados de la IA, la transparencia en sus procesos y un compromiso con la diversidad y la equidad. En este sentido, no basta con confiar ciegamente en las recomendaciones algorítmicas; es crucial validar las decisiones a través de un enfoque crítico y consciente.

2. ¿Qué medidas podemos tomar para promover mejor la equidad y la dignidad?

- **Implementar auditorías regulares de los sistemas de IA** para identificar y mitigar sesgos que puedan surgir en las recomendaciones.
- **Entrenar la IA con conjuntos de datos diversos y representativos** que reflejen una mayor inclusión y eviten replicar inequidades históricas.
- **Capacitar a los usuarios de la herramienta** para que puedan identificar posibles sesgos y tomar decisiones conscientes e informadas. Este enfoque educa a las personas a cuestionar las recomendaciones de la IA y evaluar su impacto desde una perspectiva humana y ética.

3. ¿Es posible que hayamos entrenado a una IA para que proporcione una respuesta por debajo de ese estándar más alto?

Sí, y esta posibilidad nos obliga a reevaluar el ciclo de entrenamiento de la IA. Si los datos originales reflejan inequidades humanas, es probable que la IA perpetúe esas mismas fallas. La responsabilidad recae en los desarrolladores y usuarios de la tecnología para garantizar que estas herramientas no perpetúen errores históricos. Además, esto exige un compromiso activo con la mejora continua y la supervisión humana para prevenir consecuencias no deseadas.

4. ¿Qué podemos hacer para ayudar a los humanos a tomar mejores decisiones basadas en las recomendaciones del algoritmo?

- **Proveer transparencia sobre cómo el sistema de IA genera sus recomendaciones**, ofreciendo explicaciones claras y comprensibles.
- **Asegurarse de que las decisiones finales sean tomadas con supervisión humana informada**, priorizando la capacidad crítica de los tomadores de decisiones.
- **Crear sistemas que incluyan mecanismos para cuestionar o modificar las recomendaciones del algoritmo**, permitiendo la identificación y corrección de patrones problemáticos.

En resumen, el propietario de las ideas y productos generados por la IA debe asumir la responsabilidad ética de su implementación, especialmente en situaciones donde los resultados de la IA tienen implicaciones directas en la equidad y la dignidad de las personas. La clave está en combinar la capacidad tecnológica con un compromiso inquebrantable con la justicia social y la ética profesional. Solo así podremos garantizar que la IA sea una herramienta que promueva no solo eficiencia, sino también humanidad y equidad.

3. RESPONSABILIDAD EN CASO DE DAÑOS CAUSADOS POR LA IA. (PROVISIONAL)

……

La responsabilidad por los daños ocasionados por sistemas de inteligencia artificial (IA) es un tema complejo, influido por la naturaleza tecnológica de estos sistemas y la intervención de múltiples actores en su diseño, desarrollo y utilización. A continuación, se detalla un marco general para determinar quién es responsable y cómo proceder en estos casos.

Marco Normativo Europeo

1. **Responsabilidad de los operadores de IA:** La Resolución del Parlamento Europeo de octubre de 2020 sobre el régimen de responsabilidad civil en materia de IA establece que:
   - La responsabilidad recae sobre la persona que **crea, mantiene, controla o explota** el sistema de IA.
   - La Directiva 85/374/CEE sobre responsabilidad por productos defectuosos también puede aplicarse en caso de reclamaciones contra el productor de un sistema de IA defectuoso.
2. **Diligencias preliminares:**
   - En sistemas de "alto riesgo", dada la dificultad de identificar al responsable entre los numerosos participantes, los potenciales demandantes pueden solicitar a los tribunales la **exhibición de pruebas** antes de presentar una demanda formal.
   - Esta solicitud debe estar respaldada por hechos suficientes que demuestren la viabilidad de la reclamación y, en primera instancia, debe dirigirse al proveedor o usuario del sistema de IA que se sospeche que causó el daño.

Tipos de Responsabilidad

1. **Responsabilidad basada en la culpa:**
   - En los países de la Unión Europea, el criterio predominante se basa en la culpa, como establece el artículo 1902 del Código Civil.
   - En este caso, el operador no será responsable si puede demostrar que actuó sin negligencia o imprudencia en el manejo del sistema de IA.
2. **Responsabilidad objetiva:**
   - Para sistemas de IA de "alto riesgo" que operan de forma autónoma, se aboga por un régimen de responsabilidad objetiva del operador.
   - Este enfoque elimina la necesidad de demostrar culpa y se aplica cuando la actividad desarrollada crea un riesgo evidente que excede lo habitual, similar a lo que ocurre con actividades como la conducción de vehículos a motor.
3. **Responsabilidad por producto defectuoso:**

- La responsabilidad también puede extenderse al productor (fabricante del sistema de IA) y al distribuidor que pone el sistema en manos del usuario final.
- En este esquema, el perjudicado podrá reclamar tanto al productor como al operador, dejando abierta la posibilidad de que estos actores se repitan entre sí según corresponda.

Presunción de Causalidad y Distribución de la Carga de Prueba

1. **Sistemas de IA de alto riesgo:**
   - Se introduce una **presunción refutable de causalidad**, siempre que:
     - Se demuestre que el demandado incumplió un deber de diligencia establecido por normas nacionales o de la Unión.
     - Sea razonablemente probable que dicho incumplimiento influyó en los resultados del sistema de IA.
     - Se pruebe que la salida del sistema de IA causó el daño.
2. **Diligencias preliminares:**
   - El demandante puede solicitar la exhibición de pruebas para determinar al responsable y el nexo causal antes de presentar una demanda formal.
   - Si el demandado se niega a proporcionar las pruebas requeridas, se presume que hubo un incumplimiento de sus deberes de diligencia.
3. **Exclusión de la presunción:**
   - En ciertos casos, como cuando el demandante dispone de pruebas suficientes para establecer el nexo causal, la presunción de causalidad no será aplicable.

CAPÍTULO IV. CONCLUSIONES Y RECOMENDACIONES.

.....

BIBLIOGRAFÍA (PROVISIONAL)

- El Derecho. (n.d.). *La responsabilidad civil como consecuencia de daños causados por la inteligencia artificial*. Recuperado de https://elderecho.com/la-responsabilidad-civil-como-consecuencia-de-danos-causados-por-la-inteligencia-artificial#planteamiento
- Impulsa Empresa. (n.d.). *Ética en la inteligencia artificial: Garantizar equidad y responsabilidad*. Recuperado de https://www.impulsa-empresa.es/etica-en-la-inteligencia-artificial/#Garantizar
- La Barca de Teseo. (n.d.). *Revista: artículo académico*. Recuperado de https://labarcadeteseo.org/index.php/revista/article/view/5/14
- Parlamento Europeo. (2023). *Ley de IA de la UE: primera normativa sobre inteligencia artificial*. Recuperado de https://www.europarl.europa.eu/topics/es/article/20230601STO93804/ley-de-ia-de-la-ue-primera-normativa-sobre-inteligencia-artificial
- Parlamento Europeo. (2024). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que* Protección de Datos. (n.d.). *Derechos sobre la inteligencia artificial*. Recuperado de https://protecciondatos-lopd.com/empresas/derechos-inteligencia-artificial/
- Universitat Oberta de Catalunya. (n.d.). *Cinco principios de la inteligencia artificial ética*. Recuperado de https://blogs.uoc.edu/informatica/es/cinco-principios-inteligencia-artificial-etica/