



Facultad de Ciencias Económicas y Empresariales

La Inteligencia Artificial y la empresa: dilemas éticos generados por su aplicación en un futuro automatizado

Clave: 201504218

Autor: Alejandro Valencia García

Tutor: Prof. Javier Fuertes Pérez

MADRID | Junio 2021

Abstract:

La inteligencia artificial en la actualidad no está falta de problemas. Para las empresas del presente y el futuro, los dilemas con carga ética que surgen de su aplicación son problemas de difícil solución. El rápido avance de estas tecnologías hace de estos problemas desafíos que deben combatirse en el presente para poder lograr la aplicación de estos sistemas de forma ética y segura para todos.

Artificial intelligence presents several problems nowadays. For businesses in the present and future, ethical dilemmas caused by practical applications are a tough problem to face. The speed these systems are being developed at makes it necessary to solve these problems now to allow for the safe and ethical implantation of AI systems.

Palabras clave: Negocios, Inteligencia Artificial, Dilemas Éticos, Aplicaciones Prácticas, Sistemas Inteligentes, Desarrollo.

Key words: Business, Artificial Intelligence, Ethical Dilemmas, Practical Applications, Intelligent Systems, Development.

Índice de contenidos

1. Introducción.....	4
2. Objetivos y metodología.....	6
3. ¿Qué es la inteligencia artificial?: orígenes y clasificación de la inteligencia artificial.....	7
2.1 Orígenes de la inteligencia artificial.....	7
2.2 Clasificación de la inteligencia artificial:.....	10
2.2.1 Inteligencia artificial fuerte e inteligencia artificial débil	12
2.2.2 Clasificación de Arend Hintze de sistemas de inteligencia artificial.....	13
a) Tipo I: Máquinas reactivas.....	14
b) Tipo II: Máquinas de memoria limitada.....	15
c) Tipo III: Teoría de la mente	16
d) Tipo IV: Autoconciencia.....	17
3. Aplicaciones actuales de la inteligencia artificial en los negocios	17
3.1 Automatización de tareas	18
3.2 Gestión y análisis de datos	19
3.3 Análisis de patrones de clientes y personalización de servicios	20
3.4 Análisis textual y de redes sociales	21
3.5 Ciberseguridad y detección de fraude	22
3.6 Atención a clientes e interfaces inteligentes	23
4. Problemas éticos de la inteligencia artificial en los negocios: actualidad y futuro	23
4.1 Desempleo por automatización	24
4.2 Desigualdad digital.....	25
4.3 Privacidad y control de datos	27
4.4 Parcialidad y prejuicios en el desarrollo de inteligencia artificial	31
4.5 Responsabilidad y moral de la inteligencia artificial	33
4.6 Interacción entre humanos e inteligencias artificiales.....	36
5. Conclusiones.....	38
6. Bibliografía	43

1. Introducción

La inteligencia artificial es una de las tecnologías que más avances ha experimentado en las últimas décadas. Desde los orígenes de la computación como tecnología y como disciplina, muchos visionarios han imaginado las posibilidades que podría ofrecer un ordenador capaz de pensar por sí mismo. Dadas las limitaciones que la programación y los ordenadores tienen al enfrentarse a problemas sencillos para la mente humana, resulta fácil ver por qué estas personas lo consideraban un concepto tan atractivo.

La combinación de la capacidad de razonamiento que ofrece la mente humana y la potencia de cálculo que ofrecen los ordenadores es un premio tentador que ha guiado durante muchos años el desarrollo de la inteligencia artificial. Hoy en día podemos ver los resultados de este desarrollo y observar directamente las posibilidades que ofrece esta tecnología a la sociedad y los negocios.

Los grandes avances en el ámbito de las telecomunicaciones, junto con el desarrollo de procesadores cada vez más potentes han convertido en realidad conceptos que anteriormente solo se habían planteado como teoremas. El avance en la potencia de procesamiento ha convertido en no solo factible, sino práctico, el uso del procesamiento de datos masivo para el desarrollo de la inteligencia artificial.

Lo que hoy en día se conoce como Big Data, o análisis masivo de datos forma la base sobre la que el desarrollo de la inteligencia artificial actual se ha construido. La estrecha interrelación entre estas dos tecnologías hace necesario referir a una al hablar de la otra. Los algoritmos que se entienden como inteligencia artificial dependen de su entrenamiento para avanzar, es decir, mejoran al recibir mayor cantidad de datos para refinar su procesamiento e interpretación.

Estas tecnologías ofrecen grandes beneficios a la sociedad en su aplicación práctica en una gran variedad de ámbitos. Desde planificación, análisis o sanidad hasta la aplicación directa en los negocios. Es esta última categoría en la que nos centraremos a lo largo de este proyecto.

La inteligencia artificial y todas las disciplinas hermanas como el Big Data tienen aplicaciones directamente útiles para la empresa. La mayoría de los modelos de negocio existentes en la actualidad pueden aprovechar las ventajas que ofrecen estas tecnologías para mejorar su eficiencia, reducir costes o ahorrar tiempo de alguna forma. Independientemente del sector en el que se

encuentre la actividad principal de la empresa, es altamente probable que pueda sacar algún beneficio de una de las múltiples aplicaciones que tiene la inteligencia artificial.

Sin embargo, todos estos avances presentan problemas potenciales que es necesario abordar a la hora de poder aplicar estas tecnologías en el mundo de los negocios. Algunos de los problemas principales que plantea la inteligencia artificial son de naturaleza ética (Governance Newsletter, 2018). Por este motivo, son difíciles de abordar y de solucionar. La ética en los negocios es un tema polémico desde su origen, con múltiples ideas y teorías que buscan ofrecer un marco ético aplicable a todos los negocios. La ética, por naturaleza es un tema en el que se ven involucrados muchos factores, desde la experiencia personal hasta la cultura o la zona geográfica.

Como se ha mencionado, la inteligencia artificial tiene una relación muy estrecha con otras disciplinas, como la computación o el análisis masivo de datos. Estos datos, que conforman las bases para el entrenamiento de algoritmos de inteligencia artificial, pueden representar un problema ético muy grave para las empresas que los recogen y los usan.

La privacidad es otro tema con estrecha relación con la inteligencia artificial, uno de los límites que condicionan su desarrollo. Se trata de un aspecto con gran importancia en la aplicación uso seguro de estos sistemas, como podremos comprobar más adelante. Las posibilidades de extrapolación que ofrece el uso de Big Data y la inteligencia artificial se han usado en numerosas ocasiones para invadir gravemente la privacidad de las personas (Brown, 2018).

Dado el avance de la globalización y la difusión de las tecnologías de la comunicación por todo el mundo, es importante tener en cuenta el impacto que puede tener la cultura en los dilemas éticos que plantea la inteligencia artificial. El problema de la privacidad es uno que tiene distintas connotaciones dependiendo del país, ya que las opiniones en cuanto a privacidad y recogida de datos de forma masiva varían en gran medida.

La actividad económica y la cultura de los negocios varían en gran medida dependiendo de la ubicación geográfica, por lo que es lógico pensar que las aplicaciones de inteligencia artificial en los negocios deberán adaptarse a todas estas peculiaridades para ser eficaces y éticas.

La inteligencia artificial es una tecnología que ya presenta efectos sobre la sociedad. Por esta razón, es importante que los problemas éticos que conlleva se

2. Objetivos y metodología

La creciente importancia que tienen los sistemas de inteligencia artificial en el mundo de la empresa los convierte en un caso de estudio clave para el futuro de las empresas actuales. El potencial de la tecnología viene acompañado con desafíos de muchos tipos que las empresas del futuro se verán obligadas a solventar.

La aplicación práctica de sistemas inteligentes presenta problemas de naturaleza ética para las empresas. El objetivo fundamental de este trabajo es identificar las líneas generales que los desarrolladores y reguladores deberán tener en cuenta a la hora de plantear soluciones eficaces para estos problemas. La pregunta de investigación que guiará este trabajo de investigación será ¿Qué líneas generales existen que puedan servir para solucionar los problemas de la aplicación práctica de la inteligencia artificial por parte de las empresas?

Para poder identificar estas líneas, lo primero será investigar cómo han evolucionado estos sistemas desde su concepción hasta sus aplicaciones actuales. Esto permitirá entender como el desarrollo del campo ha permitido llegar hasta los sistemas de inteligencia artificial que se emplean en la actualidad.

En segundo lugar, se examinarán las distintas formas en que se clasifican estos sistemas según sus capacidades para identificar aquellos tipos de sistemas cuya aplicación práctica es posible en el mundo empresarial. Estas categorías servirán para examinar el posible futuro de la tecnología de la inteligencia artificial. Dado que algunas clasificaciones son aún teóricas, permiten identificar las capacidades que estos sistemas podrían tener en el futuro, algo relevante para identificar futuros dilemas.

En tercer lugar, se examinarán de forma general las aplicaciones de inteligencia artificial más importantes en la actualidad y para el futuro. En base a estas aplicaciones se identificarán los problemas con naturaleza ética que se derivan de ellas. Estos problemas se examinarán analizando ejemplos existentes en el mundo empresarial que ilustren los dilemas que se presentan en el uso de estos sistemas.

De las aplicaciones extraídas de la investigación se derivarán los problemas éticos con mayor importancia para las empresas en el presente y de cara al futuro. Estos problemas se analizarán de forma general, empleando ejemplos que ilustren la relevancia que tienen para la aplicación ética de sistemas inteligentes en la empresa.

Por último, se plantearán las conclusiones en forma de las líneas que se considere que sean más importantes para guiar el desarrollo e implantación de la inteligencia artificial en la empresa. Estas conclusiones no pretenden ser exhaustivas, sino que buscan encontrar los puntos comunes que pueden servir para aplicar estos sistemas de forma ética y segura.

3. ¿Qué es la inteligencia artificial?: orígenes y clasificación de la inteligencia artificial

2.1 Orígenes de la inteligencia artificial

Para poder analizar el impacto que ha tenido la tecnología de la inteligencia artificial en la sociedad, y más concretamente en el mundo empresarial, es necesario resumir brevemente la historia de su desarrollo. Los orígenes de esta tecnología, junto con los problemas y desafíos que planteó su desarrollo enmarcan las capacidades actuales de esta tecnología y sirven para explicar los usos actuales que se le da a la inteligencia artificial.

Los orígenes de la inteligencia artificial se remontan hasta los filósofos griegos, que concebían el pensamiento humano como un sistema lógico e interpretable racionalmente. La mitología griega ya contenía menciones de autómatas, máquinas creadas para imitar al ser humano en todas sus capacidades. Aunque el concepto de inteligencia artificial no surgió allí, sus ideas sirvieron para inspirar las generaciones futuras que desarrollarían el concepto y lo llevarían a la práctica en cierta medida.

En el año 1950, el matemático Alan Turing presentó un ensayo titulado “*Computing Machinery and Intelligence*” o “Maquinaria de Computación e Inteligencia”, donde sentaba las bases de lo que él consideraría un sistema de inteligencia artificial. No con ese término, ya que el término “inteligencia artificial” se estableció de forma oficial unos años después. Una de sus primeras preguntas de investigación, y una que lanzó al resto del mundo reflejaba su curiosidad sobre esta tecnología: ¿Pueden las máquinas pensar? (Turing, 1950).

Este documento estableció las bases de la inteligencia artificial como campo, y sirvió para la creación del test de Turing. El test de Turing se consideró durante muchos años la prueba más fiable de la capacidad de un ordenador de imitar a un ser humano en su forma de pensar. El objetivo principal de esta prueba era el de evaluar las capacidades de un ordenador de imitar el comportamiento de un ser humano. Es decir, el objetivo que debían cumplir los sistemas evaluados para superar la prueba es lograr ser indistinguibles de un ser humano en su conversación.

Aunque posteriormente se ha reconocido que las observaciones que provee no son completamente objetivas, ha servido como un referente en el desarrollo de la inteligencia artificial. Miles de sistemas de inteligencia artificial se han enfrentado a lo largo de los años a esta prueba histórica con el objetivo de lograr el distintivo de haberla superado. Cada año más sistemas intentan superar la barrera que Turing estableció en competiciones para desarrolladores.

El primer sistema en superar el test de Turing lo logró en 2014. Se trataba de un programa llamado “Eugene Goostman” que pretendía simular a un niño ucraniano de 13 años. Posteriormente se ha argumentado que la prueba favorecía al programa demasiado, ya que partía con la ventaja de alegar que el inglés no era su idioma nativo. A pesar de la polémica que generó en los círculos de desarrolladores de inteligencia artificial, su éxito fue un evento histórico para este campo y refleja los avances que han ocurrido en el desarrollo de estos sistemas desde que Turing concibió su prueba (BBC News, 2014).

El término “inteligencia artificial” se empleó por primera vez en 1956 durante una conferencia en el Dartmouth College, en los Estados Unidos. Los científicos que formaron parte de aquella conferencia se mostraron optimistas respecto al futuro de la inteligencia artificial. En algunos casos, estimaron que la inteligencia artificial sería real en una generación.

Desgraciadamente, unos años después el desarrollo de inteligencia artificial experimentó un estancamiento tras varios informes que criticaron el progreso del desarrollo de la inteligencia artificial. La retirada de fondos para el desarrollo de estas tecnologías provocó una parada en la investigación y desarrollo de estos sistemas entre los años 1974 a 1980.

El interés por el desarrollo de la inteligencia artificial volvió a avivarse a partir del año 1980, cuando la competencia entre occidente y Japón en cuanto a sistemas informáticos atrajo nuevo interés al desarrollo. Aunque durante un tiempo parecía que el desarrollo de la inteligencia artificial

llegaría a desarrollarse hasta una aplicación práctica, nuevas circunstancias volvieron a retrasar su desarrollo.

Entre 1987 y 1993 volvió a producirse un estancamiento en el desarrollo de las tecnologías de inteligencia artificial, provocado por un colapso en el mercado de ordenadores que se dio en esta época. Con el interés reducido en el mercado de la computación, muchos gobiernos retiraron la financiación necesaria a los programas de desarrollo de inteligencia artificial, dejando el campo estancado hasta entrados los años 90 (Lewis, 2014).

A partir del año 1993 la inteligencia artificial recupera su importancia en el mundo del desarrollo. Los avances cada vez mayores en procesadores informáticos, sistemas de almacenamiento de memoria y en programación hacen que cada vez sea más fácil obtener ordenadores capaces de procesar un gran volumen de datos a altas velocidades.

Este procesamiento de grandes cantidades de información es justo lo que los campos de la inteligencia artificial y la ciencia de los datos necesitan para despegar y comenzar una etapa de desarrollo cada vez más rápido.

Durante este período se van marcando hitos en el mundo de la inteligencia artificial, como la derrota del campeón de ajedrez Garry Kasparov por parte del ordenador Deep Blue de IBM en 1997. Este evento es de gran importancia para los desarrolladores de sistemas de inteligencia artificial, ya que sirve para atraer nuevo interés y financiación hacia sus campos.

Desde este primer momento histórico los sistemas de inteligencia artificial han competido contra maestros de ajedrez y otros juegos clásicos de mayor complejidad. A lo largo de este periodo han logrado muchas victorias, pero sobre todo han logrado refinar cada vez más sus algoritmos hasta sobrepasar la capacidad humana en estos juegos (Anyoha, 2017).

Estas competiciones reflejan muy bien las diferencias entre el razonamiento de los sistemas de inteligencia artificial y la forma de pensar que logra la mente humana. Tras años de desarrollo los ordenadores tienen algoritmos refinados al máximo posible para la tarea o desafío que se les plantea.

Dado suficiente entrenamiento es fácil que un ordenador supere en juego a un maestro humano. Todo esto no hace más que ilustrar las posibles ventajas que los sistemas de inteligencia artificial

ofrecen a la sociedad, aunque es cierto que las reglas en los juegos no demuestran bien las limitaciones que sufren estos sistemas ante otros problemas.

El repentino avance de estos sistemas ha hecho que la sociedad esté poco preparada para enfrentarse a los problemas que puede plantear la aplicación práctica de estos sistemas. Como veremos en adelante, no hay lugar más apropiado para observar los problemas que presentan estas tecnologías que la regulación gubernamental.

El desarrollo de la tecnología en múltiples campos y especialmente en las tecnologías relacionadas con la comunicación y procesamiento de datos se ha disparado desde el año 2000. Esto ha puesto de manifiesto los muchos problemas en los que las leyes y políticas gubernamentales tienen a la hora de afrontar los escenarios que plantea la tecnología de la información.

Desde temas como la privacidad hasta la destrucción de puestos de trabajo por automatización, está claro que los gobiernos deberán estar involucrados en el nuevo paradigma que representa la cuarta revolución industrial. Para ello es muy probable que tengan que adaptarse al panorama económico y social que plantea el futuro. Mas aun, deberán encontrarse en un proceso de continua adaptación si quieren afrontar los desafíos que plantea la inteligencia artificial en el futuro cercano.

2.2 Clasificación de la inteligencia artificial:

Como hemos mencionado anteriormente, la definición más común de la inteligencia artificial es la capacidad de un sistema informático de simular el comportamiento de la mente humana. No se trata de una definición oficial, ya que los expertos aun discuten la definición de esos sistemas. El rápido avance de la tecnología dificulta aún más el acuerdo entre los expertos a la hora de establecer una definición formal para la inteligencia artificial.

Alan Turing, uno de los primeros teóricos de la inteligencia artificial, consideraba que la inteligencia artificial era una máquina capaz de aprender de sus experiencias y crecer hasta alcanzar el nivel intelectual de un humano. Para Turing, un sistema inteligente era aquel que podía simular el comportamiento humano de forma que fuera indistinguible de una persona (Turing, 1950). Con el desarrollo del campo, han surgido cada vez más definiciones de lo que se considera inteligencia artificial. El informático John McCarthy la definió como “la ciencia e ingeniería de

crear maquinas inteligentes, en especial de programas informáticos” (McCarthy, 2004). El rápido avance de la tecnología no ha hecho más que generar más definiciones y clasificaciones para ordenar el campo.

Estos nuevos avances permiten a los sistemas obtener capacidades que hasta hace poco solo podíamos observar en el razonamiento humano. Capacidades como la creatividad o el análisis complejo de situaciones en base a datos incompletos son características básicas de la mente humana, pero los sistemas informáticos tienden a encontrar dificultades al simularlas. Sin embargo, algunas características como la flexibilidad sí que están presentes en sistemas actuales.

Un punto importante a la hora de hablar de inteligencia artificial es la respuesta a la pregunta que Alan Turing lanzó en su conocido documento. Los ordenadores, por sí mismos, son incapaces de pensar. Un sistema informático se limita a seguir las instrucciones que su programación le indica. Por ello es importante entender la diferencia clave que marca el desarrollo de la inteligencia artificial. Cuando un sistema de inteligencia artificial aprende, está siguiendo las instrucciones y algoritmos que lo componen para llevar a cabo el objetivo indicado. Por eso decimos que la inteligencia artificial emula el comportamiento de la mente humana (Tablado, 2020).

Debido a los problemas que presenta lograr que un ordenador emule el razonamiento humano, los algoritmos de inteligencia artificial se clasifican según su función y capacidades. Existen múltiples formas de clasificar un algoritmo de inteligencia artificial. Generalmente estas clasificaciones se realizan según el proceso de su desarrollo, los algoritmos empleados para hacerlo y los objetivos que se persiguen con ella.

Para realizar este trabajo examinaremos dos formas principales de clasificar los sistemas de inteligencia artificial. Una clasificación es la de sistemas de inteligencia artificial fuerte y sistemas de inteligencia artificial débil. La segunda clasificación, obra de Arend Hintze establece cuatro categorías de inteligencia artificial según las capacidades de los programas.

2.2.1 Inteligencia artificial fuerte e inteligencia artificial débil

La primera clasificación que examinaremos es la de la inteligencia artificial fuerte y la inteligencia artificial débil. Se trata de una forma de clasificación que surgió en los años 80 para diferenciar los sistemas inteligentes de la época en función de sus capacidades para realizar tareas. Se trataba de una clasificación amplia, aunque una de las categorías solo es teórica hasta la fecha.

La inteligencia artificial débil, también conocida como inteligencia artificial estrecha, se compone de aquellos sistemas desarrollados y entrenados para realizar una sola tarea. Las tareas que llevan a cabo pueden ser muy variadas, desde jugar al ajedrez hasta responder a preguntas según los inputs de usuarios.

Estos sistemas son muy eficaces a la hora de realizar una única tarea para la que han sido programados y entrenados. Los problemas con este tipo de sistemas aparecen al presentar tareas diferentes a un sistema ya entrenado para realizar otra. Estos sistemas son incapaces de adaptarse a tareas distintas a las que se les ha asignado.

Un ejemplo claro de los problemas que pueden encontrarse se puede ver en algo tan aparentemente simple como el reconocimiento de una imagen. Esta es una capacidad de la mente humana que parece sencilla por lo instintiva que nos resulta, pero que tiene muchas dificultades para su realización por parte de sistemas informáticos. Si un algoritmo ha sido entrenado para reconocer caras de seres humanos en imágenes y distinguir si se trata de hombres o mujeres, no puede identificar otros objetos. Si a este sistema teórico se le presentan imágenes de perros y gatos, no será capaz de distinguir la diferencia.

Prácticamente todos los sistemas de inteligencia artificial que existen en la actualidad son sistemas de inteligencia artificial débil. Las herramientas que emplean las empresas hoy en día en su aplicación de sistemas como el machine learning o las redes neuronales se incluyen en esta categoría. Por este motivo muchos de los problemas que examinaremos en este proyecto surgen de las limitaciones que tienen estos sistemas.

El que se las conozca como inteligencias artificiales débiles no implica que sean menos útiles. Un sistema de inteligencia artificial con una tarea bien definida y el entrenamiento adecuado puede superar en rendimiento a un ser humano en la misma de forma significativa.

La inteligencia artificial fuerte es un concepto teórico hasta la fecha. Otro término empleado para describirla es el de inteligencia artificial general. El nombre indica lo que se esperaría de un sistema de inteligencia artificial para merecer incluirlo en esta categoría.

Un sistema de inteligencia artificial fuerte debería ser consciente de sí mismo y ser capaz de solucionar problemas, aprender y planificar de cara al futuro. Para ello, un sistema de inteligencia artificial fuerte debería tener las capacidades que hoy en día solo asociamos con los seres humanos. Teóricamente, una inteligencia artificial de esta clase tendría que aprender poco a poco de la información que se le proporcione para que se desarrolle a sí misma progresivamente (IBM Cloud Education, 2020).

Se trata de uno de los objetivos principales de muchos proyectos de desarrollo de inteligencia artificial. Visto lo que se ha logrado con sistemas de inteligencia artificial débil, los sistemas de inteligencia artificial fuertes podrían revolucionar el campo. En su aplicación práctica, un sistema de inteligencia artificial fuerte podría tener aplicaciones inimaginables en el mundo de los negocios. Como consecuencia, los problemas que podría causar sin precauciones son igualmente difíciles de imaginar.

Algunos expertos se muestran optimistas frente al desarrollo de inteligencias artificiales fuertes, y consideran que es posible que se desarrollen en un futuro cercano. Otros expertos son más pesimistas y consideran que necesitan mucho tiempo aun para convertirse en realidad, o incluso que no es posible desarrollar un sistema con estas características.

2.2.2 Clasificación de Arend Hintze de sistemas de inteligencia artificial

La clasificación de sistemas de inteligencia artificial fuerte y débil se ha empleado en el desarrollo de estos sistemas desde los años 80. A lo largo de los años, de ella se han derivado sistemas de clasificación más concretos que reflejan el progreso en el campo. Entre ellos, el sistema de clasificación de Arend Hintze establece subcategorías derivadas de la inteligencia artificial fuerte y la débil.

Arend Hintze es profesor de Biología Integrada y Ciencias de la Computación de la Universidad de Michigan. A lo largo de su carrera se ha dedicado al desarrollo de la inteligencia artificial

tratando de superar las barreras que existen hoy en día. Sus categorías de la inteligencia artificial sirven hoy en día como referente para el campo (Hintze, 2016).

Cabe destacar que algunas de las categorías que establece Hintze se consideran incluso hoy en día teóricas. Se podría decir que estas últimas categorías son preparatorias para los futuros avances en inteligencia artificial. Al igual que la categoría de inteligencia artificial fuerte descrita anteriormente, estas categorías sirven para delimitar las características que un futuro sistema debería cumplir para considerarse apto para ellas.

Las categorías de Hintze son cuatro: las máquinas reactivas, las máquinas de memoria limitada, la teoría de la mente y la autoconciencia.

a) Tipo I: Máquinas reactivas

Las máquinas reactivas son la categoría que comprende los sistemas de inteligencia artificial más simples. Estos sistemas son puramente reactivos en sus funciones y no tienen capacidad de mirar al pasado ni de planificar para el futuro. Sus capacidades se limitan a tareas concretas y analizar el momento presente.

Este tipo de sistemas de inteligencia artificial solo son capaces de percibir el mundo que les rodea en el momento actual para realizar su tarea. Su mundo se ve reducido al campo en el cual el sistema deba ejecutar sus funciones. No es capaz de formar memorias y solo actúa con la información que recibe directamente.

Un ejemplo de este tipo de sistemas es el ordenador Deep Blue de IBM mencionado con anterioridad. El sistema de Deep Blue está completamente enfocado al ajedrez, y no es capaz de analizar los movimientos que un contrincante ha hecho anteriormente. Su función se limita a estudiar el estado de un tablero en un momento concreto y a realizar el movimiento que calcule óptimo en el momento (Lewis, 2014).

Estos sistemas son muy potentes cuando se aplican correctamente, pero su utilidad fuera de las tareas para las que se han especializado son muy limitadas. Los sistemas de memoria reactiva son incapaces de interactuar con el entorno que los rodea más allá de recibir los inputs por el canal que se le haya asignado. Además, son incapaces de evolucionar y adaptarse a nuevas tareas o

circunstancias. En términos simples, cuando una máquina reactiva se encuentra con una situación que ha visto con anterioridad, actuará de la misma forma.

La mayoría de las aplicaciones prácticas que existen en los negocios de sistemas de inteligencia artificial pueden englobarse en esta categoría. Enfocadas a tareas concretas con poca desviación en los parámetros son capaces de superar con creces el rendimiento humano. Estos sistemas se emplean en la actualidad en miles de empresas automatizando tareas repetitivas que anteriormente requerían personal humano.

b) Tipo II: Máquinas de memoria limitada

Este tipo de sistemas de inteligencia artificial se caracteriza por tener la capacidad de mirar hacia el pasado. Sus sistemas y algoritmos permiten la formación de recuerdos almacenados en la memoria del sistema. La capacidad de estos sistemas de inteligencia artificial de almacenar información del pasado tiende a ser limitada y temporal.

Estos sistemas se emplean en ocasiones en las que una inteligencia artificial debe monitorizar variables a lo largo del tiempo. Para poder hacer esto, es necesario que sean capaces de almacenar y revisar información en el pasado.

Estas observaciones se recogen y se procesan según la programación de los desarrolladores, que deben crear una representación del mundo para el sistema. Un sistema de inteligencia artificial de memoria limitada solo es capaz de almacenar información si sus desarrolladores indican que es necesaria para su tarea. En este aspecto es igual de dependiente del entrenamiento que los desarrolladores le proporcionen que las máquinas reactivas. Una vez programada la representación de su entorno, es capaz de modificarla ligeramente con la información que almacena para mejorar su rendimiento.

Aunque el mejor ejemplo de los usos de estos sistemas es el de los coches autónomos, existen muchos ámbitos en el mundo de los negocios que emplean esta clase de inteligencia artificial. Las empresas dedicadas a la inversión financiera dependen en gran medida de la recogida y análisis de datos. Hoy en día los sistemas de inteligencia artificial son cada vez más capaces de automatizar la búsqueda y recolección de datos relevantes del mercado de valores para su análisis.

Un sistema dedicado al seguimiento del mercado financiero debe ser capaz de comparar el estado de las posiciones que le indiquen sus usuarios con los valores del pasado. Por ello estos sistemas tienen mucho éxito en tareas relacionadas con el rastreo y seguimiento de valores. Servicios como los rastreadores de precios o los asistentes de inversiones digitales emplean hoy en día sistemas de memoria limitada (Fourie & Bennett, 2019).

c) Tipo III: Teoría de la mente

Esta tercera categoría representa una línea importante que divide los sistemas de inteligencia artificial que existen en la actualidad y aquellos que podrían ser posibles en un futuro. Los sistemas que comprenden las dos últimas categorías pertenecen más al mundo de la investigación y desarrollo teóricos que a la práctica por el momento.

Se ha mencionado anteriormente la representación del mundo que un sistema de inteligencia artificial requiere. Al hablar de la representación se hace referencia al entorno que un desarrollador de inteligencia artificial debe programar para que su sistema opere en él. La representación del mundo debe ser lo más parecida posible a la realidad en la que el sistema desarrollará sus funciones. Por este motivo los programadores deben dedicar mucho esfuerzo a programar un entorno lo más fiel posible al mundo real.

Un sistema de este tipo debería ser capaz de analizar su entorno y elaborar una representación de lo que le rodea. Lo que la distingue de otras clases es que además debería ser capaz de elaborar representaciones de los agentes que la rodean.

Esto en psicología se denomina la teoría de la mente, y engloba la capacidad humana de entender que otros agentes del entorno tienen sus propios factores que afectan al comportamiento. Esta capacidad que es instintiva en los seres humanos es muy difícil de modelar para un sistema informático. Es un aspecto clave de la mente humana, ya que es el que les ha permitido construir sociedades y cooperar unos con otros.

En el futuro, si el desarrollo de este tipo de sistemas tiene éxito, podrían suponer una revolución en la economía y sociedad del mundo. Estos sistemas serían capaces no solo de analizar el comportamiento de las personas, sino de predecirlo partiendo de información incompleta. En caso de aplicarse en el mundo de los negocios podría suponer una revolución en muchos sectores. Un

sistema informático capaz de analizar el mercado de valores y de comprender el efecto que el pánico tiene en la economía podría realizar operaciones financieras respondiendo ante incertidumbre más rápido que un sistema automático actual.

d) Tipo IV: Autoconciencia

Esta es la última categoría que establece Hintze y, al igual que la anterior, solo es teórica por el momento. Los sistemas que engloba esta categoría tienden hacia la idea que la ciencia ficción ha popularizado sobre la inteligencia artificial.

Para considerarse parte de esta categoría, una inteligencia artificial debería ser capaz de establecer representaciones del entorno que la rodea y de sí misma. En términos simples, este sistema sería consciente de sí mismo y de aquellos que lo rodean de la misma forma que lo son los seres humanos.

Esta categoría es una extensión del tipo III de inteligencias artificiales, ya que profundiza en capacidades de la mente humana aplicadas a sistemas informáticos. Si su desarrollo se realizara con éxito, supondría un cambio radical en la sociedad y la economía globales. Aunque su desarrollo supondría una revolución, es interesante examinar las consecuencias que tendría en los negocios.

3. Aplicaciones actuales de la inteligencia artificial en los negocios

La inteligencia artificial es una tecnología que brilla por la variedad de aplicaciones que presenta a los usuarios. Miles de negocios utilizan en el presente alguna aplicación de la inteligencia artificial en sus operaciones. En muchos casos, las aplicaciones de inteligencias artificiales de Tipo I y Tipo II permiten a las empresas automatizar tareas repetitivas o interpretar rápidamente datos de su negocio.

Aunque existen miles de aplicaciones posibles, y cada vez más se desarrollan cada año, para centrarnos en los objetivos de este documento se centrará en las aplicaciones dedicadas

directamente a las empresas. Sería fácil encontrar ejemplos de aplicaciones de inteligencia artificial en áreas concretas como la sanidad, la banca o los seguros, pero el objetivo de este documento es realizar un análisis general de las aplicaciones más comunes.

3.1 Automatización de tareas

Una de las aplicaciones más extendidas de sistemas de inteligencia artificial de Tipo I y Tipo II es la automatización de tareas repetitivas. Como se ha mencionado anteriormente, estos sistemas son capaces de un rendimiento significativamente superior al de un ser humano en tareas concretas.

La automatización de tareas es algo que muchas empresas emplean en la actualidad en todos los niveles de su estructura. Desde el servicio directo a clientes hasta la logística, las formas de aplicar la inteligencia artificial han cambiado el paradigma para muchos tipos de negocios.

Un punto importante para aclarar es que los sistemas de automatización actuales aun requieren supervisión humana en muchos casos. Aunque el desarrollo de inteligencia artificial avanza cada año, aún existen limitaciones que impiden la independencia total de estos sistemas en sus tareas. En muchos casos, estos programas funcionan de forma principalmente independiente, alertando a un operario cuando surgen problemas que el sistema es incapaz de solucionar.

En el futuro, la inteligencia artificial promete poseer la capacidad de operar en entornos complejos que puedan reducirse a tareas repetitivas. Un ejemplo muy claro de usos de inteligencia artificial de esta clase se encuentra en la logística. En este campo cada vez más la inteligencia artificial organiza la distribución de productos a consumidores de forma masiva.

Uno de los principales referentes en la aplicación de sistemas de inteligencia artificial en los negocios es Amazon. Desde sus orígenes Amazon ha apostado por la tecnología como forma de crear valor y potenciar sus servicios. Su uso de inteligencia artificial en la logística y la distribución es uno de los ejemplos más impresionantes que existen de automatización mediante inteligencia artificial.

Amazon dispone de centros de distribución regionales completamente informatizados. Cuando un cliente realiza una compra en su tienda online, una serie de programas y sistemas basados en la inteligencia artificial comienza el proceso de envío. En muchos centros de distribución, miles de

robots guiados por inteligencia artificial ejecutan órdenes para enviar millones de productos a su destino final.

Todo esto es posible con un mínimo de supervisión humana, logrando crear un sistema de distribución capaz de responder instantáneamente a los pedidos de los clientes. Otro factor que refleja la importancia de la inteligencia artificial en estos centros de distribución es el inventario. El inventario en estos centros es controlado por un sistema de inteligencia artificial que lo organiza según los pedidos actuales y predice los posibles pedidos de futuros clientes (Dash, McMurtrey, & Kar, 2019).

Este último punto es clave para una empresa como Amazon que ofrece a sus clientes envíos rápidos en miles de productos. Los centros de distribución locales almacenan los productos necesarios para cubrir los pedidos de su zona de distribución. Los sistemas de inteligencia artificial permiten solicitar productos de antemano para cubrir demandas futuras con una precisión sorprendente.

Estas predicciones se emplean para gestionar toda la cadena de distribución de Amazon desde sus proveedores hasta la llegada de los productos a los clientes. En cierto sentido, los sistemas de inteligencia artificial de Amazon son de los más extensos del mundo, ya que regulan un flujo de mercancías a escala global.

Aunque el ejemplo de Amazon es uno de los más impactantes en escala y amplitud, no se deben descartar las ventajas que la automatización ofrece a las empresas en otros ámbitos. Las posibilidades que ofrece son tan variadas como el número de tareas repetitivas o claramente definidas que puedan establecerse dentro de una empresa. Otras aplicaciones importantes son la automatización de tareas peligrosas para seres humanos, como el trabajo con materiales dañinos para la salud.

[3.2 Gestión y análisis de datos](#)

Las empresas hoy en día dependen de la información para poder realizar sus operaciones. La gestión eficiente de los datos es clave en muchas empresas de todos los sectores para poder mantener un funcionamiento óptimo. El análisis e interpretación rápida y fiable de esa información es algo que permite a los directivos tomar las mejores decisiones para sus empresas, por lo que es crucial hoy en día.

La importancia de los datos se ha marcado en numerosos estudios sobre la privacidad y análisis de datos. Shoshana Zuboff denominó el fenómeno que marca la comercialización de estos datos como “capitalismo de vigilancia”. En la economía actual los datos se han convertido en bienes comercializables que permiten a las empresas perfilar a sus clientes con gran precisión. Los numerosos medios de recogida de datos de los que disponen las empresas y la utilidad que suponen han llevado a una invasión casi completa de la vida de las personas (Zuboff, 2015). Esta combinación de volumen de datos y análisis cada vez más preciso da a las empresas un gran poder sobre sus clientes. Dichos clientes en muchos casos ceden inadvertidamente información que se puede emplear para manipular su opinión e incitar al consumo.

En este ámbito los sistemas de inteligencia artificial son de gran utilidad. Una vez más, los sistemas de máquinas reactivas y de memoria limitada son los que las empresas emplean actualmente en este campo. Desde realizar análisis complejos de datos según los inputs de nueva información por parte de trabajadores hasta el análisis y resumen de textos, hoy en día las tareas de gestión de datos que pueden hacer estos sistemas son de gran variedad (Bakken, 2019).

Este control y análisis de datos puede tomar muchas formas dentro de una empresa. Un ejemplo claro es el del control de tareas de mantenimiento. En muchas empresas cuyas actividades se centran en la industria, hoy en día el análisis del estado de herramientas está automatizado. En muchos casos, sistemas de inteligencia artificial monitorizan las condiciones dentro de un centro de producción para asegurar que el producto final es satisfactorio.

En otros casos el propio mantenimiento se ve regulado por sistemas de inteligencia artificial encargados de controlar el estado de robots industriales o herramientas. De esta forma los operarios humanos solo tienen que centrarse en el mantenimiento y revisión de los problemas que resalta el programa para mantener el buen funcionamiento.

[3.3 Análisis de patrones de clientes y personalización de servicios](#)

Como consecuencia de los avances en tecnologías de comunicación y de transportes, los comercios tradicionales han dejado de ser la única alternativa para los consumidores. El sector del comercio online es hoy en día uno de los más importantes para el comercio mundial. La evolución de los sistemas informáticos y de comunicación en las últimas décadas han llevado a la creación de un

comercio internacional. Sin embargo, la proliferación de ordenadores personales y la expansión de internet hasta lo que existe hoy en día ha permitido el acceso al comercio online a más personas que nunca. La facilidad de acceso a internet y la mejora de la conectividad ha logrado que el sector de las compras online gane cada vez más peso en el comercio en todos los niveles.

Uno de los grandes beneficios de la inteligencia artificial para las empresas de comercio online es el uso de datos. Como se ha mencionado anteriormente, el análisis de datos y la inteligencia artificial son disciplinas hermanas que confluyen en muchas aplicaciones. El análisis de clientes y la personalización es uno de estos campos donde ambas ofrecen grandes beneficios.

Los programas de inteligencia artificial que emplean muchas empresas permiten analizar los patrones de compra de clientes. De esta forma establecen perfiles de sus clientes que se emplean para personalizar las recomendaciones de compras según sus intereses percibidos. El uso de datos y el análisis de clientes mediante inteligencia artificial permite a las empresas conocer mejor a sus clientes y saber qué clase de productos buscan (Klubnikin, 2016).

Esta información es muy valiosa para las empresas, ya comercialicen sus productos online o no. Para las empresas con presencia online, la información fiable de sus clientes es una forma importante de mejorar sus productos y servicios o adaptarse a nuevas tendencias.

[3.4 Análisis textual y de redes sociales](#)

En el mundo de los negocios hoy en día la presencia en las redes sociales es muy importante para la imagen de la empresa. En un mundo en el que estas redes tienen tanta influencia es importante para las empresas mantener un control constante de su imagen y su percepción por parte de sus clientes.

Para ello, muchas empresas tienen un departamento dedicado al marketing en redes sociales. Muchos de estos departamentos hoy en día emplean servicios de inteligencia artificial que ayudan a monitorizar las redes sociales en que está presente la empresa (Bakken, 2019). Estos sistemas son capaces de analizar el texto de comentarios de redes sociales o posts de blogs e interpretar las opiniones que representan. De esta forma si surge una tendencia en las redes sociales relevante para la empresa, ya sea de forma positiva o negativa, la empresa puede aprovecharla para posicionarse como prefiera frente a sus clientes.

Estas herramientas son muy potentes en sus aplicaciones, y ofrecen grandes ventajas a empresas cuyas actividades pueden depender en gran medida de las opiniones de los clientes. Uno de los sectores donde esto cobra gran importancia es en el entretenimiento, ya que permite adaptarse a la volatilidad de la opinión pública con mucha rapidez.

3.5 Ciberseguridad y detección de fraude

En un mundo cada vez más conectado, las empresas están realizando grandes inversiones en ciberseguridad. Los datos de todo tipo son cada vez más valiosos, por lo que existen muchas amenazas en la red dispuestas a obtenerlos. La ciberseguridad es una de las grandes prioridades de empresas y gobiernos en el presente por los peligros que el crimen online presenta.

En este ámbito los sistemas de inteligencia artificial están cobrando cada vez más importancia. Aunque es cierto que no es la aplicación más extendida de inteligencia artificial, el objetivo de este proyecto es analizar las aplicaciones presentes que mayor relevancia pueden tener en el futuro. Por este motivo, la aplicación de inteligencia artificial en ciberseguridad es un caso relevante para los objetivos del proyecto.

En las comunidades de ciberseguridad siempre se asume que, en un sistema de ciberseguridad, el punto más débil para cualquier ataque es el usuario humano. La mayoría de las técnicas existentes para superar protecciones informáticas dependen de la acción inadvertida de un usuario que abre una vulnerabilidad.

La inteligencia artificial promete detectar y contrarrestar problemas de seguridad informática con mucha más velocidad que técnicos humanos. Con un sistema siempre vigilante y capaz de identificar ataques al instante, muchas de las vulnerabilidades típicamente empleadas son difíciles de explotar. Muchas empresas han visto el daño que un ciberataque puede causar a sus operaciones, su reputación y a sus clientes.

Otra aplicación relacionada con la seguridad de las operaciones de la empresa es la de detección de fraude. Esto es algo que concierne principalmente al sector financiero o de los seguros, pero con aplicaciones posibles en muchos otros campos.

La combinación de técnicas de análisis de datos y la inteligencia artificial han creado programas capaces de identificar intentos de cometer fraude con un alto grado de precisión. Esto representa una ventaja clara para las empresas, ya que permite reducir el tiempo invertido en investigar y determinar si existe la posibilidad de fraude. Muchas empresas del sector de los seguros y de la banca han invertido en sistemas inteligentes capaces de identificar posibles fraudes entre sus clientes. Esto permite recortar costes en investigaciones y agilizar los procesos necesarios para solucionar esta clase de problemas (Balasubramanian, Libarikian, & McElhaney, 2021).

3.6 Atención a clientes e interfaces inteligentes

Tradicionalmente, las empresas dedicadas al sector de servicios debían mantener una inversión importante en personal dedicado a la atención de sus clientes. Esto suponía cubrir los costes de operarios humanos para las líneas telefónicas y los equipos de comunicación necesarios. Con los avances en tecnologías de comunicación muchos de estos servicios han evolucionado, pero siguen siendo una parte importante de las operaciones de la empresa.

La inteligencia artificial permite automatizar la gestión de llamadas telefónicas al departamento de atención al cliente. Mediante un sistema inteligente que permite filtrar llamadas, se consigue una mayor fluidez en la atención a los clientes.

Este sistema es aplicable en más campos que en la atención telefónica. En muchas empresas se ha popularizado la posibilidad de chatear online con un sistema que asista al cliente. En caso de ser incapaz de determinar y solucionar el problema, el sistema conecta al cliente con un representante humano.

4. Problemas éticos de la inteligencia artificial en los negocios: actualidad y futuro

Como se puede observar, la inteligencia artificial ofrece grandes beneficios a aquellas empresas que deciden emplearla en sus operaciones. Sin embargo, estas aplicaciones pueden conllevar problemas de carácter ético si no se toman precauciones. A continuación, se van a analizar algunos de los problemas éticos de la inteligencia artificial aplicada a negocios más relevantes. Estos problemas comprenden en líneas generales los dilemas a los que se enfrentan las empresas en el uso de inteligencia artificial en el presente y de cara al futuro.

Generalmente, estos problemas están relacionados con tres categorías principales: la progresiva invasión de la privacidad que supone el uso de inteligencia artificial, la transformación de la sociedad y la economía por su uso generalizado y los dilemas de la creación de sistemas que buscan aproximarse a la mente humana. Cada uno de ellos deriva de alguna de las aplicaciones actuales de sistemas inteligentes analizadas, y pueden aumentar en magnitud al mismo ritmo que ellas en el futuro.

4.1 Desempleo por automatización

Al hablar de dilemas éticos del uso de inteligencia artificial en los negocios, el desempleo es uno de los temas más polémicos. Esto se debe en parte a que con los avances en tecnología estas últimas décadas, muchos empleos tradicionales han sido suplantados por sistemas automáticos. Para muchos, la inteligencia artificial representa un grave peligro para las economías y sociedades de todo el mundo. Esto se debe a que pone en peligro muchos puestos de empleo de los cuales las personas dependen para ganarse la vida (Sutton, Vicky, & Holt, 2018).

Como se ha mencionado anteriormente, la automatización de tareas es una de las grandes ventajas que ofrece la inteligencia artificial a las empresas. Muchas ya emplean esta clase de sistemas para agilizar trámites administrativos que antes requerían personal humano. Únicamente en los departamentos administrativos, la inteligencia artificial tiene un enorme potencial para agilizar el desarrollo de las actividades de la empresa. Como consecuencia, muchos expertos consideran posible e inevitable la destrucción de millones de puestos de empleo de esta clase.

Esto representa un dilema muy importante para las empresas y la sociedad. Si los sistemas de inteligencia artificial permiten automatizar estas tareas requiriendo un mínimo de supervisión, ¿Cuál es el futuro de los trabajadores cuyos puestos dejan de ser necesarios?

Muchos expertos argumentan que estos trabajadores podrían reentrenarse para poder desempeñar otra función en el mercado laboral. En muchos casos afirman que la automatización de estas tareas abriría puestos de trabajo destinados a la supervisión y control de sistemas de inteligencia artificial.

El problema que surge es la gran diferencia en el número de puestos de trabajo que se crearían frente a los que desaparecerían. El objetivo de la automatización es aumentar la eficiencia y reducir los costes que suponen trabajadores humanos. Si unas tareas como los procesos de producción en

una fábrica se automatizan, la supervisión de los sistemas requiere menos trabajadores de los que servían para realizar las tareas anteriormente.

En un informe de 2018 la empresa PwC realizó un estudio de más de 200.000 puestos de trabajo en 29 países diferentes. El objetivo del estudio era analizar el impacto que la automatización podría tener en el mercado laboral global en los próximos diez a quince años.

Al cabo del estudio concluyeron que un 30% de los puestos de empleo globales podrían verse eliminados por la automatización a mediados de la década de 2030. El estudio encontró que las mujeres podrían verse afectadas en mayor medida por la automatización a corto y medio plazo. En el largo plazo, los empleos manuales y físicos en que los hombres tienen mayor presencia serían los más afectados (PwC, 2018).

Este estudio consideraba los efectos de la automatización mediante la aplicación de sistemas similares a los que se encuentran en uso actualmente. Inteligencias artificiales de tipo I y tipo II aplicadas de forma generalizada serían capaces de producir estos efectos en el mercado laboral global.

Es importante tener en cuenta que la inteligencia artificial es un campo en constante desarrollo. Con el interés que ha suscitado en gobiernos de todo el mundo, los avances se producen más rápido que nunca. Por este motivo, es posible que los sistemas que se produzcan en los próximos diez años alteren significativamente las conclusiones de este informe.

4.2 Desigualdad digital

La desigualdad es un problema que afecta a todo el mundo desde el origen de las civilizaciones. Desde la Revolución Industrial las diferencias entre las clases sociales no han hecho más que crecer en todo el mundo. La Cuarta Revolución Industrial podría continuar esta tendencia si no se toman precauciones con los avances que promete. La desigualdad es un problema en que intervienen factores de todo tipo, desde económicos a sociales, pero es un problema ético muy grave que hay que considerar.

Al reflexionar sobre los problemas económicos y sociales generados por la inteligencia artificial, la desigualdad no suele ser uno de los principales dilemas. Sin embargo, el problema de la desigualdad generada por la aplicación de inteligencia artificial por parte de las empresas es muy

importante para la economía del futuro. Este problema deriva en parte de la automatización de muchas tareas de gestión que las empresas antes gestionaban mediante trabajadores humanos. En sectores como la logística, es habitual que algoritmos guíen las actividades de los repartidores sin necesidad de supervisión humana. En empresas con modelos de negocio basados en inteligencia artificial existe una tendencia hacia una división tecnológica entre los trabajadores de bajo nivel y los puestos de gestión.

En los modelos de empresa tradicionales, casi siempre existía algún tipo de movilidad vertical para los trabajadores. Un trabajador tenía posibilidades de ascender y adquirir un mejor puesto de trabajo dentro de una empresa. Ejemplos como el de Charlie Bell, el antiguo CEO de McDonald's que comenzó su carrera haciendo hamburguesas, eran pocos, pero existían.

El modelo de muchas nuevas compañías pone muchas dificultades al movimiento vertical para trabajadores en puestos de bajo nivel. Un conductor trabajando para la empresa Uber hallaría muchas dificultades a la hora de acceder a un empleo de mayor grado en la empresa, por no hablar de alcanzar puestos en la dirección.

Los trabajadores de empresas con este modelo ejecutan sus funciones siguiendo las indicaciones de algoritmos. Tienen poco contacto con el resto de la empresa en general, limitándose a seguir las pautas de un sistema dedicado a maximizar la eficiencia. En estos puestos de empleo el contacto o la cooperación de otros trabajadores del mismo nivel son mínimos o inexistentes. Además, la constante monitorización del rendimiento puede llevar a penalizar a los trabajadores si un algoritmo considera que su trabajo no es satisfactorio (Walsh, 2020).

El empleo se encuentra en un proceso de polarización cada vez más prevalente en todo el mundo. Por una parte, existen los trabajadores con los conocimientos y capacidades de trabajar con los algoritmos que rigen estas empresas. Por otra se encuentran los trabajadores sin experiencia o conocimientos que trabajan siguiendo las indicaciones de dichos algoritmos. La diferencia en los niveles de educación necesaria para acceder a los puestos de mayor nivel representa un obstáculo para el ascenso dentro de la propia empresa. Por la propia estructura de las empresas cuyos negocios se basan en algoritmos, existe una brecha entre niveles de la empresa prevalente en muchos negocios en todo el mundo. La diferencia con la estructura tradicional de la empresa es clara, ya que las estructuras de poder de ambas se encuentran separadas por sistemas de

comunicación indirecta. Esta situación convierte a los trabajadores de bajo nivel en piezas fácilmente reemplazables en caso de mal rendimiento para la empresa.

Entre ambos existe una brecha en las capacidades y conocimiento que impide a muchos trabajadores progresar en su carrera profesional. Estos puestos de empleo no educan al trabajador ni le preparan para poder desempeñar funciones a un nivel más alto. Las funciones que llevan a cabo son simples, repetitivas y están monitorizadas por algoritmos con la capacidad de penalizar al trabajador por mal rendimiento.

4.3 Privacidad y control de datos

Como se ha mencionado en varias ocasiones, muchos de los avances de la inteligencia artificial y el análisis de datos se deben a las similitudes entre ambas tecnologías. La combinación de técnicas de análisis de datos y el uso de programas de inteligencia artificial han permitido en la actualidad el análisis masivo de datos.

Una consecuencia de este rápido desarrollo es que la información es cada vez un bien más valioso. La información personal de millones de personas de todo el mundo se ha convertido en un bien comercializable, con aplicaciones cada vez más importantes para las empresas. Muchas empresas han aprovechado rápidamente el análisis de datos para mejorar sus actividades perfilando con mayor precisión a sus clientes.

Uno de los problemas más importantes de la tecnología de la comunicación en general es el de la privacidad de las personas. Desde el inicio de las bases de datos, las empresas han acumulado información sobre sus clientes relevante para el desarrollo de la actividad. En sectores como los seguros, donde el análisis estadístico es clave, estas bases de datos eran importantes fuentes de información.

Con el desarrollo de las técnicas de análisis masivo de datos y la inversión en inteligencia artificial por parte de las empresas, estas bases de datos tienen más versatilidad que nunca. El desarrollo de las redes sociales y el comercio online han ayudado a generar registros sobre las acciones y opiniones de billones de personas en el planeta. Todos estos usuarios de servicios online en el mundo ceden su información personal a aquellas empresas que proveen los servicios. En la mayoría de las ocasiones los usuarios no saben qué se va a hacer con esa información, y las consecuencias que puede tener.

Desde el punto de vista ético, la privacidad es algo a lo que todas las personas tienen derecho. La intrusión de la tecnología en la vida de las personas es cada vez más prevalente, y es algo que debe tenerse en cuenta para el uso ético de inteligencia artificial en los negocios. En su teoría del capitalismo de vigilancia, Zuboff afirma que el conjunto de capacidades de recogida y análisis de información da a las empresas una forma de poder sobre sus clientes difícil de contrarrestar. Zuboff denomina las sociedades actuales “sociedades de la información”, recalcando lo importantes que son los datos en la actualidad.

El uso de técnicas de inteligencia artificial como el machine learning permiten a las empresas procesar un gran volumen de datos y segmentar a la población con precisión increíble. Esto implica que un usuario puede estar utilizando un servicio de redes sociales y una inteligencia artificial es capaz de extrapolar rasgos personales como gustos u opinión política.

La cuestión en este tema es hasta qué punto estarían dispuestos los usuarios de estos servicios a seguir usándolos, sabiendo que las empresas pueden extrapolar información privada que no se les ha proporcionado. En muchos casos, el limitar los datos personales que se ceden es difícil o imposible al usar estos servicios. A esto se combina el problema de que muchas bases de datos se venden a empresas, permitiéndoles combinarlos para crear perfiles precisos de sus posibles clientes.

Casi todas las acciones del día a día dejan un rastro digital. Desde comprar un café mediante un pago con tarjeta de crédito hasta emplear el navegador de teléfono móvil, muchas acciones aparentemente inocuas quedan registradas y catalogadas. Estos sistemas funcionan constantemente, recogiendo información en segundo plano que se condensa en enormes bases de datos.

En su teoría del capitalismo de vigilancia, Zuboff denomina este conjunto de sistemas y capacidades “El Gran Otro”. Con esto se refiere al poder de recogida, análisis y uso de la información que poseen empresas como Google o Facebook. Según Zuboff, estas capacidades tienen un impacto muy importante en la economía y sociedad actuales. Llegan incluso a afectar a áreas como la política o la ideología, algo que en el capitalismo tradicional se consideraba separado de la economía. (Zuboff, 2015)

El dilema en cuanto a la privacidad tiene dos partes principales. En primer lugar, se encuentra el dilema de que las empresas registren datos personales de clientes con la justificación de emplearlos para mejorar sus servicios. Es cierto, pero en muchos casos esto no es lo único que se hace con esos datos personales. Con la proliferación de asistentes personales del hogar como Alexa de Amazon, espacios que tradicionalmente y por ley se consideran privados en muchos países han dejado de serlo. Estos asistentes representan el último paso hacia la recolección total de datos en todas las esferas a la que tienden los avances tecnológicos de los últimos años.

La segunda parte del dilema tiene que ver con el compromiso que una sociedad está dispuesta a aceptar entre privacidad y utilidad. Los servicios de empresas que emplean inteligencia artificial son útiles, pero tienen un precio que muchas veces pasa inadvertido en la cesión de espacios privados (Brown, 2018). Los límites que la empresa debería respetar en cuanto a la recogida de datos son ambiguos en muchos casos, algo que dificulta mucho el regular su recogida y comercialización.

Para solucionar estos problemas existe cierta regulación en cuanto al tipo de datos personales que pueden recogerse y las formas de procesarlos que pueden emplearse. Desgraciadamente, en muchas ocasiones estos son insuficientes. La mayoría de las leyes de protección de datos aseguran que un individuo no puede ser identificado mediante sus datos personales. Esto se consigue anonimizando los datos mediante técnicas de análisis de datos (Mutschelknaus, 2020).

El principal problema de estas medidas es el equilibrio que las empresas tienen que mantener para poder emplear los datos. Si un dato se anonimiza excesivamente, pierde utilidad en cuanto a la información que permite extraer de él. Por el contrario, si un conjunto de datos personales no se anonimiza lo suficiente, puede ser posible identificar a individuos concretos en un conjunto de datos.

Aquí vuelve a aparecer una vez más el dilema que marca los problemas de la privacidad de datos. El equilibrio entre privacidad y utilidad, aplicado esta vez desde la perspectiva de las empresas que recogen dichos datos. Este equilibrio marca el centro del dilema del derecho a la privacidad frente a las utilidades y ventajas de las nuevas tecnologías.

Para combatir estos problemas el método más eficaz sería del de la regulación por parte de los gobiernos. En general, la regulación es lenta en comparación con la velocidad de desarrollo de las

nuevas tecnologías. Por este motivo, en la actualidad aún se discuten los particulares de marcos legales relativos a la privacidad y el uso de datos por parte de los negocios.

Un ejemplo de los esfuerzos para combatir estos problemas es el de la Unión Europea. La Comisión Europea presentó en abril de 2021 una propuesta de regulación conjunta para la inteligencia artificial dentro de la Unión Europea. Se trata del primer esfuerzo por parte de la Unión Europea para establecer un marco regulatorio conjunto para el desarrollo y empleo de inteligencias artificiales (Comisión Europea, 2021). Este marco recogía su base ética de una propuesta del Parlamento Europeo a la Comisión Europea para la creación de un marco ético para las cuestiones de la robótica, la inteligencia artificial y la tecnología. La propuesta, presentada y adoptada en 2020, contiene los principios éticos que deberían regir estas tecnologías según el Parlamento Europeo (Parlamento Europeo, 2020). Los principios establecidos son:

- El desarrollo de sistemas de inteligencia artificial centrados en humanos, creados por humanos y controlados por humanos.
- Evaluaciones obligatorias del cumplimiento de normas éticas en inteligencias artificiales de alto riesgo por parte de organismos supervisores regionales si se requiriera.
- Seguridad, transparencia y responsabilidad. En esto se incluye informar a usuarios de que están interactuando con sistemas de inteligencia artificial, garantizar su uso seguro y disponer de métodos seguros de desactivación en caso de incumplimiento.
- Protección y soluciones contra problemas de discriminación o parcialidad
- Derecho a compensación por parte de los usuarios
- Sostenibilidad en el desarrollo o implantación de sistemas de inteligencia artificial.
- Respeto a la privacidad y limitaciones en el reconocimiento con datos biométricos.
- Buena gestión de la inteligencia artificial y los datos recogidos durante su operación.

Este marco ético refleja los principios y valores que rigen las instituciones y acciones de la Unión Europea. Dado que la Unión Europea reconoce una serie de valores comunes para las naciones que la componen, se trata de un importante punto de partida para la regulación común de estas tecnologías.

De cara al futuro, la ética deberá ser uno de los puntos principales donde se base la regulación del uso de sistemas inteligentes. Dada la rapidez de los avances de estas tecnologías, es necesario que

los gobiernos preparados para adaptarse a los cambios que puedan suponer para la economía y sociedad.

4.4 Parcialidad y prejuicios en el desarrollo de inteligencia artificial

En apartados anteriores se ha recalcado que los llamados sistemas de inteligencia artificial no son realmente inteligentes. Con esto nos referimos a que, en su base, un sistema de inteligencia artificial se compone de algoritmos de mayor o menor complejidad. Estos algoritmos solo ejecutan su función según la información de entrenamiento que le proporcione su desarrollador.

El proceso de entrenamiento es similar para todos los tipos de inteligencia artificial empleados en la actualidad (tipo I y tipo II). Una vez delimitados los objetivos y la tarea del sistema, se introducen conjuntos de datos relevantes para entrenar los algoritmos. El sistema de inteligencia artificial procesa este volumen de datos y reacciona en función de su tarea. Este proceso de entrenamiento se repite hasta que los desarrolladores están satisfechos con el rendimiento del sistema.

Un problema que ha surgido a lo largo del desarrollo de la inteligencia artificial es el de la parcialidad en el desarrollo de sus funciones. En el campo de la inteligencia artificial, la parcialidad se refiere a alteraciones en la interpretación de la información que presenta una inteligencia artificial (Howard & Borenstein, 2018). Se habla de la parcialidad cuando una inteligencia artificial refleja una falta de imparcialidad u objetividad en la interpretación de datos o ejecución de sus funciones.

Al simular el funcionamiento de la mente humana, los sistemas de inteligencia artificial pueden adquirir prejuicios debido a problemas en la información de entrenamiento o errores en la programación. Para mantener la imparcialidad en el funcionamiento, los desarrolladores deben tener cuidado en las fases de programación y entrenamiento iniciales.

Durante la programación de los algoritmos es posible que errores de programación produzcan un sistema que no procesa la información con total imparcialidad. Igualmente, durante el entrenamiento los conjuntos de datos deben reflejar la realidad lo mejor posible. Los problemas de parcialidad por mal entrenamiento surgen cuando los conjuntos de datos preparados por desarrolladores reflejan opiniones preconcebidas. Los programas procesan estos conjuntos y observan patrones que pueden no ser representativos de la realidad.

En los negocios, la parcialidad representa un problema importante en la aplicación de inteligencia artificial. Los sistemas más vulnerables a este problema son aquellos que tienen que evaluar a personas según rasgos personales. Esto es clave en sectores de la banca como los préstamos, o para las compañías de seguros. De cara al futuro, se trata de un problema más importante de lo que parece. Una de las funciones en las que se está empleando la inteligencia artificial es en la de recursos humanos de la empresa. La automatización de procesos de selección puede presentar problemas si los algoritmos no son completamente imparciales.

Un ejemplo importante de la importancia de la imparcialidad en el uso de inteligencia artificial es el que sucedió en Amazon. Amazon es una de las principales compañías del mundo, y es líder mundial en la aplicación de sistemas de inteligencia artificial en la empresa.

Un proyecto de Amazon iniciado en 2014 consistió en la automatización de su proceso de selección de recursos humanos. El objetivo del sistema en que se basaba el proyecto era filtrar las solicitudes de candidatos y asignar una puntuación a los candidatos. Esta puntuación funcionaba de forma similar a las evaluaciones de la tienda de Amazon, de una a cinco estrellas. De esta forma pretendían crear un sistema que recibiera cientos de solicitudes y presentara las más atractivas.

En 2015 la empresa observó que las recomendaciones del algoritmo favorecían los perfiles de los hombres para empleos de desarrolladores de sistemas. Tras investigar el problema se encontró que el entrenamiento de los algoritmos se había realizado introduciendo todas las solicitudes de un periodo de diez años. Durante el entrenamiento el algoritmo observó patrones en las solicitudes de los candidatos recibidas por la empresa. La parcialidad a los perfiles masculinos reflejaba el estado del mercado laboral del sector tecnológico, en el que los hombres tenían mayor presencia (Dastin, 2018).

Desde el punto de vista ético, la parcialidad y los prejuicios en los algoritmos son un dilema que puede afectar a las empresas y a las sociedades. Los seres humanos tenemos prejuicios que condicionan nuestro razonamiento. Nuestras opiniones y valores se ven afectadas en muchas ocasiones por estos prejuicios.

Para mantener la imparcialidad de los sistemas de inteligencia artificial en el futuro, es necesario supervisar su desarrollo para evitar la introducción involuntaria de prejuicios. Una posible solución para el problema es establecer unas líneas generales en cuanto al desarrollo de inteligencia artificial

imparcial. Esto podría conseguirse mediante regulación por parte de instituciones gubernamentales, aunque requeriría que expertos sobre el tema asesoraran a los órganos regulatorios. De no ser así se correría el riesgo de obstaculizar el desarrollo de sistemas inteligentes con un exceso de regulación.

4.5 Responsabilidad y moral de la inteligencia artificial

Un punto importante en la ética de la inteligencia artificial es el de la moralidad. Un sistema de inteligencia artificial no es capaz de comprender la moralidad de la misma forma que los humanos. Todas las culturas del planeta tienen una serie de normas subyacentes que rigen la sociedad. Estas normas no escritas se ven reflejadas en la ley y la sociedad que rigen. Lo que se considera moral y ético puede variar en gran medida en función de muchos factores. Diferencias como la etnia, la cultura o la zona geográfica causan grandes variaciones en la moralidad de las personas.

Los sistemas de inteligencia artificial no tienen normas morales internas, sino que las reciben de sus desarrolladores. De la misma forma en que los prejuicios pueden heredarse de los desarrolladores, las normas que rigen la conducta de estos sistemas se basan en las de sus creadores.

Uno de los dilemas de la inteligencia artificial en el presente es el de la responsabilidad de la inteligencia artificial. Como se ha mencionado, estos sistemas no tienen conciencia y se limitan a seguir su programación. Si una inteligencia artificial causa daño a una persona, ¿Es responsable de sus acciones?

En la actualidad la responsabilidad de las acciones de sistemas de inteligencia artificial recae sobre los propietarios y desarrolladores. Los desarrolladores en especial son aquellos que más importancia tienen en este problema. Durante su trabajo tienen la responsabilidad de considerar todos los posibles impactos de su sistema en la sociedad, y limitar el daño todo lo posible (Bonime-Blanc, 2018).

La cuestión del posible daño causado por sistemas de inteligencia artificial es importante en la actualidad. Sin embargo, los avances de la tecnología indican que la ética en su desarrollo y aplicación va a cobrar mayor importancia en el futuro. Existe un campo del desarrollo de la inteligencia artificial en que la cuestión de la moralidad de las acciones tiene un papel central: el coche autónomo.

El desarrollo de coches autónomos es un caso en el que la moralidad de la inteligencia artificial cobra gran importancia en el presente. Se trata de una tecnología que promete grandes beneficios para campos como la logística por tierra. En el futuro, podría ser posible que los transportes por tierra se realicen mediante camiones autónomos. Estos vehículos podrían funcionar sin descanso y evitarían los accidentes causados por el sueño en carreteras.

Sin embargo, la tecnología se encuentra aún en desarrollo. Aunque existen prototipos de coches autónomos, la tecnología debe desarrollarse más para garantizar su seguridad. Durante los últimos años se han producido accidentes en los que estaban involucrados vehículos autónomos, resaltando la falta de madurez de la tecnología. Ejemplos como el ocurrido en mayo de 2018, cuando un prototipo de coche autónomo de la empresa Uber atropelló a una mujer que cruzaba la carretera de forma inapropiada en Tempe, Arizona.

Tras meses de investigación la *National Transportation Safety Board* atribuyó el accidente a una “cultura de seguridad inadecuada”. Según sus investigaciones, el accidente podría haberse evitado con mayores precauciones y conciencia de la seguridad por parte del operario y de Uber. Un problema del coche que surgió a la luz fue que el ordenador de a bordo era incapaz de identificar a un ser humano si no se encontraba cerca de un paso de cebra. Al cruzar la carretera de forma inapropiada, la víctima no fue identificada como una persona, por lo que no accionó los frenos (U.S. National Transportation Safety Board, 2019).

Este accidente demuestra los peligros que pueden surgir si no se toman las suficientes precauciones en el desarrollo de la inteligencia artificial. Errores indetectables en condiciones de prueba pueden tener graves consecuencias al presentarse en el mundo real. En este caso, el programa era incapaz de detectar seres humanos cruzando de forma inapropiada, ya que en su experiencia todas las personas seguían las normas viales.

La moralidad en situaciones como los accidentes de coche es un tema que los desarrolladores de coches autónomos deben tener en cuenta. Cuando se produce un accidente, un sistema de inteligencia artificial tiene segundos para tomar decisiones que pueden causar víctimas de forma inevitable. Las preguntas entonces son ¿Cuál es la acción moral en estos casos? ¿A quiénes se les debería priorizar la supervivencia?

Para responder a estas preguntas un grupo de investigadores del MIT comenzó un experimento conocido como *The Moral Machine*. El experimento se llevó a cabo en una plataforma online a la que podían acceder visitantes de todo el mundo. La plataforma presentaba una serie de situaciones basadas en el dilema ético del tranvía. En este dilema se plantea una situación con dos alternativas en que la muerte de alguna víctima es inevitable, y la persona debe decidir si prefiere causar la muerte a una para salvar a varias o no actuar.

El experimento planteaba un escenario en que un coche autónomo perdía el control aproximándose a un paso de cebra. Las situaciones del experimento también presentaban dos alternativas, aunque complicaban los escenarios con más variables. A lo largo de una serie de escenarios, los usuarios debían escoger la mejor de dos opciones según su opinión. En los escenarios la elección era entre niños y ancianos, los pasajeros y los peatones, ricos y pobres o combinaciones de estos y otras categorías.

El experimento cobró fama en las redes sociales y pronto se convirtió en el mayor experimento sobre la moralidad de la historia. La plataforma recogió más de cuarenta millones de decisiones, en diez idiomas diferentes de personas de 223 países y territorios. Con estos datos, los investigadores lograron identificar variaciones en la moral de las personas según la demografía, zona geográfica y la cultura entre otros (Awad, y otros, 2018).

El experimento identificó diferencias significativas en las decisiones dependiendo de la zona geográfica y la cultura. Las implicaciones en el desarrollo de inteligencia artificial de este experimento son muy importantes. Implica que es casi imposible programar un programa capaz de tomar decisiones moralmente aceptables para todo el mundo, ya que la moralidad y la ética pueden variar significativamente.

Este planteamiento parece justificar la teoría ética del relativismo moral. Esta teoría considera que existen diferencias insalvables entre lo que las personas consideran moral en función de distintos factores. Sin embargo, existen teorías opuestas como el universalismo moral que se oponen a este razonamiento. Estas teorías son importantes porque suponen la mejor oportunidad para crear sistemas de inteligencia artificial capaces de decisiones morales en el futuro.

Esto es un punto muy importante para el desarrollo de sistemas de inteligencia artificial más avanzados que los existentes actualmente. Su utilidad en el mundo empresarial implica que estarán

al cargo de la toma de decisiones en muchos ámbitos. Si en el futuro se logra el desarrollo de inteligencias artificiales fuertes (tipo III y tipo IV), la toma de decisiones morales y la responsabilidad que tienen de dichas decisiones será un aspecto clave de su uso en la economía.

Este problema coincide con el debate ético de la universalidad de la razón moral. Este debate ha enfrentado a filósofos durante cientos de años, y aparece nuevamente reflejado en las aplicaciones de nuevas tecnologías. La universalidad de la moral es un tema que Kant desarrolla en su pensamiento, afirmando que las acciones se consideran morales mientras que sea posible aplicarlas como ley universal para todas las personas. De esta forma, Kant pretendía establecer la existencia de algún principio moral que se aplica de forma común a todos los seres racionales.

La Declaración Universal de los Derechos Humanos podría entenderse como una forma de universalismo moral, y sirve como ejemplo de que es posible lograr acuerdos en aspectos fundamentales de la ética en el mundo. Para muchos autores, este evento representa un punto clave en el estudio de la ética universal, y una base sobre la que se pueden construir normas morales aceptables de forma global (Carchidi, 2020).

Si existen diferencias significativas en los valores heredados por estos sistemas, pueden llevar a conflictos entre distintos tipos de algoritmos con diferentes prioridades. Aunque es muy probable que no veamos estos sistemas antes de 2050, es importante tener en cuenta estos problemas en el camino hacia su desarrollo. Es de especial interés dado el énfasis que están poniendo los gobiernos de grandes potencias en ponerse a la cabeza del desarrollo de la inteligencia artificial (Geist, 2016).

4.6 Interacción entre humanos e inteligencias artificiales

Los dilemas éticos mencionados anteriormente ya tienen cierto impacto en el presente, pero existen muchos que pueden afectar al futuro de estas tecnologías. Uno de los dilemas emergentes más importantes es el problema de los derechos de la inteligencia artificial (Andreotta, 2021). El campo de la investigación y desarrollo de la inteligencia artificial tiene un objetivo subyacente en todos sus ámbitos. Este es la creación de sistemas capaces de emular completamente el funcionamiento de la mente humana.

Los sistemas de inteligencia artificial fuertes son aún un concepto teórico, pero es importante empezar a plantear el marco en el que se desarrollarían. Los sistemas de los tipos III y IV

categorizados por Hintze serían sistemas equivalentes en inteligencia a ciertos animales y seres humanos respectivamente.

Esta situación plantearía problemas morales muy importantes en cuanto a la definición de lo que se considera un organismo consciente e inteligente. La conciencia es uno de los aspectos de los seres humanos que los convierte en únicos en el planeta. Un sistema de inteligencia artificial está compuesto por sus algoritmos, programación y su entrenamiento. Todo ello generalmente es propiedad de las empresas o laboratorios involucrados en su desarrollo. Si un sistema de esta clase adquiriera la conciencia ¿Sería aún considerado como un objeto que pertenece a estas entidades?

La Declaración Universal de los Derechos Humanos de la O.N.U. ofrece a toda persona una serie de derechos por el simple hecho de ser humano. Si un sistema de inteligencia artificial llega a emular la conciencia humana de forma indistinguible de una persona ¿No debería recibir los mismos derechos y protecciones?

Estas preguntas tendrán mucha influencia en la relación entre seres humanos e inteligencias artificiales en el futuro. Está visto que la investigación sobre estos dilemas suele aparecer una vez se han convertido en problemas propiamente dichos. En un caso como el de sistemas inteligentes será vital adelantarse en la creación de un marco común en cuanto a la evaluación y tratamiento de los distintos tipos de inteligencia artificial. Si en el futuro se busca la cooperación estrecha entre humanos e inteligencias artificiales, será clave que existan unas líneas generales que guíen la relación.

Para las empresas del futuro, estas cuestiones pueden suponer la diferencia entre adaptarse a un mundo conectado o desaparecer. La regulación que surja en torno a las inteligencias artificiales del futuro condicionará en gran medida su uso en los negocios. Por este motivo las empresas no solo deben pensar en las ventajas que les ofrecen estos sistemas, sino planificar para un futuro en que sean más que algoritmos y programas.

5. Conclusiones

Este proyecto de investigación se ha centrado principalmente en las aplicaciones de la inteligencia artificial en la empresa de la actualidad. Para ello, el primer paso ha sido examinar lo que se considera inteligencia artificial, y las clases generales en que se categorizan. De esta forma es posible examinar los tipos de inteligencia artificial empleados en la actualidad y los que podrían surgir en el futuro. Cada uno de estos tipos tiene distintas capacidades que presentan ventajas e inconvenientes para las empresas.

Debido a la variedad de aplicaciones concretas existentes, el objetivo ha sido examinar las aplicaciones más prevalentes en la actualidad. Para cada una de estas aplicaciones se han examinado las ventajas e inconvenientes que presentan, con énfasis en los problemas a medio y largo plazo.

De estas aplicaciones y de la investigación realizada se han identificado y analizado los problemas éticos más importantes que afectan al uso de la inteligencia artificial en la empresa. Cada uno de estos dilemas representa un problema que es necesario solucionar para el futuro de la inteligencia artificial en el mundo empresarial. Algunos ya han tenido impacto en las aplicaciones prácticas de la tecnología, pero existen otros que representan problemas a largo plazo. Este proceso de investigación ha servido para extraer una serie de conclusiones sobre el futuro del uso de la inteligencia artificial de forma ética en forma de posibles líneas generales para la solución de los dilemas que presenta.

En primer lugar, es necesario que la respuesta a los dilemas planteados se realice en un contexto global. La cooperación a la hora de establecer normas que regulen el desarrollo y uso de inteligencia artificial en los negocios es clave para el éxito. En la situación actual las grandes potencias están realizando grandes inversiones en inteligencia artificial. En cierto modo, existe en la actualidad una carrera de la inteligencia artificial en la que varias naciones compiten por el primer puesto (Geist, 2016). Esto representa un problema para el desarrollo ético de inteligencia artificial. El clima de competición incita a mantener en secreto los avances en la investigación. Además, la falta de transparencia en el desarrollo de estos sistemas representa un problema a la hora de evitar problemas éticos en los algoritmos desarrollados.

Otro problema es que el desarrollo de estos sistemas se realiza de forma regional para los avances más importantes. Esto representa un problema ya que, como se observó en el experimento de *The Moral Machine*, la moralidad humana varía en cuanto a culturas y regiones. Estos sistemas desarrollados por equipos de una región luego realizarán operaciones de forma global. Los sistemas de logística inteligentes de empresas como Amazon son un ejemplo de esta situación. Es muy posible que los valores heredados de los desarrolladores choquen con los de otras regiones en las que opere el algoritmo.

Si en el futuro se quiere lograr la implantación de forma ética de inteligencia artificial en los negocios la mejor solución es lograr un compromiso entre naciones. Establecer unas líneas éticas comunes para el desarrollo e implantación de estos sistemas en la economía global es vital para el futuro. Aunque lograr el compromiso en algo como la ética es una tarea difícil, es necesario para prevenir problemas a nivel global.

En este sentido, las teorías éticas de la universalidad de la razón moral son un punto de partida importante para la creación de un marco que regule la ética en el desarrollo de la inteligencia artificial de forma global. El ejemplo de la Declaración Universal de los Derechos Humanos debería servir como aliciente de que el acuerdo entre naciones en cuanto a principios éticos es algo posible.

Para crear este marco, es imprescindible que los legisladores preparen la regulación colaborando con expertos en estos sistemas, atendiendo siempre a los principios de la ética como base. En el campo de la inteligencia artificial aplicada el exceso de regulación podría reducir las ventajas que ofrece a las empresas y desincentivar el desarrollo. Esta tecnología representa demasiadas ventajas para las empresas, por lo que limitarla excesivamente podría afectar negativamente a sus actividades en el futuro. Sin embargo, la falta de regulación eficaz sería incluso peor que su exceso, ya que dejaría vía libre a muchos problemas de naturaleza legal y ética.

El ejemplo de la reciente propuesta de regulación de nuevas tecnologías de la Comisión Europea representa un paso adelante en la integración de estos sistemas en la sociedad según valores definidos por aquellos a los que afectarán. El hecho de que el desarrollo e implementación de sistemas inteligentes en la Unión Europea se realice siguiendo principios éticos de la Unión implica que es posible lograr regular estas tecnologías en base a la ética (Comisión Europea, 2021).

Otro punto importante para combatir estos dilemas es la velocidad de desarrollo de la tecnología. Los gobiernos e instituciones deberán enfrentarse al problema como una situación en continua evolución. Como se ha visto en otras tecnologías, si los gobiernos buscan una única solución general para el problema está quedará obsoleta en cuestión de años. Por este motivo, se puede concluir que los gobiernos deberán estar preparados para adaptarse continuamente a los desafíos que presente la tecnología.

Una conclusión importante a raíz de esta investigación es la importancia de la seguridad. En el desarrollo de inteligencia artificial se crean sistemas que llevarán a cabo tareas en el mundo real. Esto conlleva un riesgo importante para las personas en caso de fallos. Los riesgos son tan variados como las aplicaciones de los sistemas. La responsabilidad de los desarrolladores y dueños es velar por garantizar la seguridad de las personas durante las operaciones de su sistema.

La transparencia durante el desarrollo de un sistema de inteligencia artificial es un buen instrumento para evaluar los riesgos de un sistema. La evaluación externa de algoritmos puede permitir a empresas reducir los riesgos a los que exponen a las personas en caso de fallos en el funcionamiento. Como se ha mencionado, los sistemas de inteligencia artificial dependen de la representación creada por sus programadores para aprender a moverse en un entorno. Las consecuencias cuando el escenario no se corresponde completamente con la realidad pueden ser muy graves. Esto es muy importante para el desarrollo de sistemas cuyas tareas puedan acarrear riesgos para personas. Ejemplos de estos sistemas son los coches autónomos o los sistemas que controlan maquinaria pesada en fábricas. En estos casos un error en la programación o entrenamiento pueden tener consecuencias funestas.

En el desarrollo de las líneas éticas de la inteligencia artificial para el futuro, las empresas deben considerar a fondo el impacto de la tecnología en el entorno. En cualquier compromiso para la regulación del desarrollo de estos sistemas, la seguridad de las personas debe ser la prioridad, ante todo.

Una conclusión importante en cuanto a la seguridad es la prevención del uso de estos sistemas con fines criminales o inmorales. Los gobiernos, desarrolladores y empresas deberán hacer todo lo que esté en su mano para impedir estas posibilidades. El abaratamiento de la tecnología hace posible que en el futuro estos sistemas puedan emplearse con malos fines. Es un problema que es necesario estar preparados para combatir desde el principio. La inteligencia artificial en la actualidad es una

herramienta más y, como se ha mencionado, no entiende de ley, ética, o moralidad. Vistos los riesgos que tiene en sus aplicaciones normales, resulta fácil imaginar el daño que podría causar si se empleara de forma criminal.

Para finalizar este proyecto se presentan las limitaciones a las que este se ve sujeto y algunas líneas de investigación posibles que surgen a raíz de él. Estas limitaciones son importantes en el alcance de la investigación, por lo que hay que definir las claramente.

En primer lugar, el proyecto se ha centrado principalmente en los tipos de inteligencia artificial que existen en la actualidad. Esto se limita a sistemas de inteligencia artificial débil, categorizados como tipo I y tipo II por Hintze. Estos sistemas son los más simples, aunque no por ello se debe subestimar la potencia que tienen. La falta de ejemplos prácticos de inteligencias artificiales fuertes hace que cualquier análisis de su impacto en el mundo empresarial se limite a la teoría. En este proyecto se han analizado sistemas con aplicaciones prácticas que presentan distintos problemas que pueden crecer en el futuro.

El enfoque del estudio se ha centrado de forma general en las aplicaciones prácticas en empresas de la tecnología, mostrando ejemplos de cada una de ellas. El análisis de cada aplicación no es excesivamente profundo, ya que sirve como marco para reflejar los problemas éticos que conllevan. Un análisis detallado de cada aplicación requeriría un importante estudio relacionado con las capacidades técnicas de los sistemas y algoritmos actuales y métodos por los cuales se realiza.

Otra limitación concierne al estudio de los dilemas éticos analizados. Los dilemas analizados se centran en su mayoría en aquellos que ya son relevantes para la aplicación práctica de estos sistemas. Los dilemas relacionados con temas como la conciencia y los derechos de la inteligencia artificial tienen menor importancia en el corto y medio plazo. La literatura relacionada con la inteligencia artificial contiene más investigación en problemas actuales. Presumiblemente por su mayor relevancia en la actualidad. Por esta razón, el cuerpo principal del análisis se dedica a problemas cuyas consecuencias afectan al futuro cercano.

Por último, este estudio se enfoca en el uso de inteligencia artificial en los negocios. Esta tecnología ha tenido un gran impacto en todo tipo de sectores, pero el impacto social o en la política son algunos que este estudio no investiga con profundidad. Los cambios que ha causado, y aquellos

que pueda causar, afectaran fuertemente a todas las esferas. Aunque hemos analizado el impacto de problemas como la automatización, esta investigación profundiza más en el lado empresarial y económico que en el impacto social que puede causar.

De cara a investigación futura sobre estos temas, existen algunas líneas de investigación que parecen relevantes para el futuro. En primer lugar, la integración en el futuro de inteligencias artificiales fuertes dentro de las empresas es una línea de investigación que podría cobrar mucha relevancia en función de los avances del campo en los próximos años.

Por otra parte, la cuestión de los problemas que podrían causar algoritmos con diferentes valores heredados de sus creadores sería interesante en el contexto de las operaciones financieras. ¿Cómo afectarían los valores heredados al rendimiento de estos sistemas en función de la región?

6. Bibliografía

- Ananny, M. (2017). Boards Need to Keep an Eye on the Ethics of AI: Balancing algorithms, business and humanity. *Directors & Boards*, 26-27.
- Andreotta, A. J. (2021). The hard problem of AI rights. *Ai & Society*, 36, 19-32.
- Anyoha, R. (17 de Agosto de 2017). *The History of Artificial Intelligence*. Obtenido de Harvard University Blog: Special Edition on Artificial Intelligence: <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., . . . Rahwan, I. (2018). The Moral Machine experiment. *Nature*, 59-64.
- Azoulay, A. (2018). Towards an Ethics of Artificial Intelligence. *UN Chronicle*, 24-25.
- Bakken, R. (18 de Marzo de 2019). *Business Applications for Artificial Intelligence: An Update for 2020*. Obtenido de Professional Development: Harvard Division of Continuing Education: <https://professional.dce.harvard.edu/blog/business-applications-for-artificial-intelligence-an-update-for-2020/>
- Balasubramanian, R., Libarikian, A., & McElhaney, D. (12 de Marzo de 2021). *Insurance 2030—The impact of AI on the future of insurance*. Obtenido de McKinsey & Company: <https://www.mckinsey.com/industries/financial-services/our-insights/insurance-2030-the-impact-of-ai-on-the-future-of-insurance>
- BBC News. (9 de Junio de 2014). *Computer AI passes Turing test in 'world first'*. Obtenido de BBC News: <https://www.bbc.com/news/technology-27762088>
- Bonime-Blanc, A. (2018). Technology, Trust and Ethics: An Actionable Governance Toolkit for a Disruptive Time. *Carrier Management*, 31-35.
- Brown, J. (18 de Diciembre de 2018). *The Amazon Alexa Eavesdropping Nightmare Came True*. Obtenido de Gizmodo: <https://gizmodo.com/the-amazon-alexa-eavesdropping-nightmare-came-true-1831231490>
- Carchidi, V. J. (2020). The Nature of Morals: How Universal Moral Grammar Provides the Conceptual Basis for the Universal Declaration of Human Rights. *Human Rights Review*, 65-92.
- Carter, D. (2018). How real is the impact of artificial intelligence? The business information survey 2018. *Business Information Review*, 99-115.
- Comisión Europea. (21 de Abril de 2021). *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS*. Obtenido de Eur-Lex: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0206>

- Cosgrove, B. (16 de Enero de 2020). *8 ways to ensure your company's AI is ethical*. Obtenido de World Economic Forum: <https://www.weforum.org/agenda/2020/01/8-ways-to-ensure-your-companys-ai-is-ethical/>
- Dash, R., McMurtrey, M. E., & Kar, U. (2019). Application of artificial intelligence in automation of supply chain management. *Journal of Strategic Innovation and Sustainability*, 14(3), 43-53.
- Dastin, J. (8 de Octubre de 2018). *Amazon scraps secret AI recruiting tool that showed bias against women*. Obtenido de Reuters: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
- Fourie, L., & Bennett, T. K. (2019). Super intelligent financial services. *Journal of Payments Strategy & Systems*, 151-164.
- Geist, E. M. (2016). It's already too late to stop the AI arms race—We must manage it instead. *Bulletin of the Atomic Scientists*, 72(5), 318-321.
- Governance Newsletter. (2018). Business ethics and AI. *Governance Newsletter*, 3.
- Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, 99-120.
- Hintze, A. (14 de Noviembre de 2016). *Understanding the four types of AI, from reactive robots to self-aware beings*. Obtenido de The Conversation: <https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616>
- Howard, A., & Borenstein, J. (2018). The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity. *Science and Engineering Ethics*, 1521-1536.
- IBM Cloud Education. (31 de Agosto de 2020). *Strong AI*. Obtenido de IBM Cloud Learn Hub: <https://www.ibm.com/cloud/learn/strong-ai#toc-what-isstr-kGAqO4bV>
- Kantarci, A. (8 de Junio de 2021). *131 Myth-Busting Statistics on Artificial Intelligence (AI) in 2021*. Obtenido de AIMultiple: <https://research.aimultiple.com/ai-stats/#does-ai-replace-human-stats>
- Kantarci, A. (2 de Febrero de 2021). *AI Ethics in 2021: Top 9 Ethical Dilemmas of AI*. Obtenido de Aimultiple: <https://research.aimultiple.com/ai-ethics/#unemployment-due-to-automation>
- Kirsten, M. (2019). Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics*, 835-850.
- Klubnikin, A. (4 de Noviembre de 2016). *Top 4 Applications of Artificial Intelligence in Business*. Obtenido de Andrei Klubnikin: <https://andrei-klubnikin.medium.com/top-4-applications-of-artificial-intelligence-in-business-7804e3cf9bf0>

- Lewis, T. (4 de Diciembre de 2014). *A Brief History of Artificial Intelligence*. Obtenido de Livescience: <https://www.livescience.com/49007-history-of-artificial-intelligence.html#:~:text=The%20beginnings%20of%20modern%20AI,%22artificial%20intelligence%22%20was%20coined.>
- McCarthy, J. (24 de Noviembre de 2004). *What is Artificial Intelligence*. Obtenido de John McCarthy's Home Page: <http://www-formal.stanford.edu/jmc/whatisai.html>
- Munoko, I., Brown-Libur, H., & Vasarhelyi, M. (2020). The Ethical Implications of Using Artificial Intelligence in Auditing. *Journal of Business Ethics*, 209-234.
- Mutschelknaus, J. E. (23 de Julio de 2020). *Top Five Data Privacy Issues that Artificial Intelligence and Machine Learning Startups Need to Know*. Obtenido de InsideBigData: <https://insidebigdata.com/2020/07/23/top-five-data-privacy-issues-that-artificial-intelligence-and-machine-learning-startups-need-to-know/>
- Ng, A. (2 de Julio de 2019). *Amazon Alexa keeps your data with no expiration date, and shares it too*. Obtenido de Cnet.com: <https://www.cnet.com/news/amazon-alexa-keeps-your-data-with-no-expiration-date-and-shares-it-too/>
- Parlamento Europeo. (20 de Octubre de 2020). *Framework of ethical aspects of artificial intelligence, robotics and related technologies*. Obtenido de Legislative Observatory - European Parliament: <https://oeil.secure.europarl.europa.eu/oeil/popups/printsummary.pdf?id=1636985&l=en&t=D>
- Pazzanese, C. (26 de Octubre de 2020). *Great promise but potential for peril: Ethical concerns mount as AI takes bigger decision-making role in more industries*. Obtenido de The Harvard Gazette: <https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/>
- PwC. (Febrero de 2018). *Will robots really steal our jobs? An international analysis of the potential long term impact of automation*. Obtenido de PwC: <https://www.pwc.co.uk/economic-services/assets/international-impact-of-automation-feb-2018.pdf>
- Rossi, F. (2018). Building Trust in Artificial Intelligence. *Journal of International Affairs Vol. 72, No. 1, THE FOURTH INDUSTRIAL REVOLUTION*, 127-134.
- Rowthorn, M. (2019). How Should Autonomous Vehicles Make Moral Decisions? Machine Ethics, Artificial Driving Intelligence, and Crash Algorithms. *Contemporary Readings in Law and Social Justice*, 9-14.
- Satell, G. (2016). Teaching an Algorithm to Understand Right and Wrong. *Harvard Business Review*, 2-5.

- Sutton, S. G., Vicky, A., & Holt, M. (2018). How Much Automation Is Too Much? Keeping the Human Relevant in Knowledge Work. *Journal of Emerging Technologies in Accounting*, 15-25.
- Tablado, F. (18 de Junio de 2020). *Inteligencia artificial: Definición, tipos y aplicaciones*. Obtenido de Grupo Atico34: <https://protecciondatos-lopd.com/empresas/inteligencia-artificial/>
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 433-460.
- U.S. National Transportation Safety Board. (19 de Noviembre de 2019). *'Inadequate Safety Culture' Contributed to Uber Automated Test Vehicle Crash - NTSB Calls for Federal Review Process for Automated Vehicle Testing on Public Roads*. Obtenido de National Transportation Safety Board: <https://www.nts.gov/news/press-releases/Pages/NR20191119c.aspx>
- Walsh, M. (22 de Octubre de 2020). *Algorithms Are Making Economic Inequality Worse*. Obtenido de Harvard Business Review: <https://hbr.org/2020/10/algorithms-are-making-economic-inequality-worse>
- Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30, 75-89.