



Universidad Pontificia Comillas, ICADE

ÉTICA APLICADA A LA INTELIGENCIA ARTIFICIAL DE LOS COCHES AUTÓNOMOS

Autor: Patricia Guilleuma González

Director: Javier Fuertes Pérez

ÍNDICE

1. RESUMEN Y ABSTRACT, PALABRAS CLAVE Y KEYWORDS	5
2. INTRODUCCIÓN	6
2.1 Objetivos:.....	6
2.2 Metodología	6
2.3 Principales partes del TFG	7
3. INTELIGENCIA ARTIFICIAL	9
3.1 ¿Qué es la inteligencia artificial?	9
3.2 Tipos de inteligencia artificial	10
3.3 Modelos de investigación:	11
3.3.1 Modelo de investigación simbólico	11
3.3.2 Modelo de investigación conexionista	12
3.3.3 Modelo de investigación evolutivo.....	13
3.3.4 Modelo de investigación corpóreo y situacional	13
4. INTELIGENCIA ARTIFICIAL APLICADA AL SECTOR DE LA AUTOMOCIÓN	14
4.1 ¿Dónde encontramos inteligencia artificial en automoción?.....	14
4.2 Grados de automatización en los coches autónomos	15
4.2.1 Nivel 0: La conducción no tiene asistencia alguna.....	15
4.2.2 Nivel 1: Hay conducción asistida para el conductor.....	16
4.2.3 Nivel 2: Automatización parcial.	16
4.2.4 Nivel 3: Automatización condicional.	17
4.2.5 Nivel 4: Alto nivel de automatización para la conducción.....	17
4.2.6 Nivel 5: Nivel de automatización completo para la conducción.	18
4.3 Beneficios que traerán los coches autónomos	19
5. LA ÉTICA EN EL DISEÑO DE LA INTELIGENCIA ARTIFICIAL	20
5.1 Principios éticos.....	20
5.2 Beneficios	20
5.3 Riesgos.....	21
5.4 ¿Deberían ser las máquinas consideradas agentes morales?	21
6.GRANDES PROBLEMÁTICAS ÉTICAS	24

6.1 Privacidad	25
6.3 Seguridad	28
6.3.1 Primer dilema moral: Niño pequeño Vs persona mayor	29
6.3.2 Segundo dilema moral: Una persona Vs varias personas	30
6.3.3 Tercer dilema moral: Animales	31
6.3.4 Cuarto dilema moral: Personas que cumplen la ley vs personas que no cumplen la ley	31
6.3.6 Moral Machine	34
6.4 RESPONSABILIDAD	38
6.5 REFLEXIÓN ÉTICA	40
7. CONCLUSIONES	43
8. BIBLIOGRAFÍA:	44

ÍNDICE DE IMÁGENES

<i>Imagen 1: LEVELS OF DRIVING AUTOMATION</i>	18
<i>Imagen 2: Moral Machine Results</i>	35
<i>Imagen 3: The Moral Machine Experiment</i>	36

ÍNDICE DE TABLAS

<i>Tabla 1: Diferencias entre IA fuerte e IA débil</i>	11
--	----

1. RESUMEN Y ABSTRACT, PALABRAS CLAVE Y KEYWORDS

Resumen:

Este trabajo trata de una reflexión sobre la inteligencia artificial que llevan implementada los coches autónomos, ya que cambiarán de forma radical la forma que conocemos actualmente de transportarnos. Para establecer un marco teórico es necesario conocer cuales son los dilemas éticos a los que se enfrentaría y las distintas respuestas que se pueden dar.

Tras analizar el caso de estudio se concluye que: Es una industria poco desarrollada y que todavía quedan muchos aspectos por matizar para que puedan verse circulando por las vías, por otro lado, hay que tener en mente qué papel se quiere que asuman los seres humanos ya que supondrá un cambio muy grande para la sociedad que conocemos hoy en día. Por ello este trabajo muestra que es necesario seguir investigando de qué manera se puede adaptar la IA de la manera más justa y segura en la sociedad.

Palabras clave: Inteligencia artificial, ética, coche autónomo, agente moral, dilema moral, problemática ética

Abstract:

This paper is a reflection on the artificial intelligence implemented in autonomous cars, as they will radically change the way we currently know how to transport ourselves. In order to establish a theoretical framework it is necessary to know what are the ethical dilemmas that would be faced and the different answers that can be given.

After analyzing the case study it is concluded that: It is a little developed industry and that there are still many aspects to be nuanced so that they can be seen circulating on the roads, on the other hand, it is necessary to keep in mind what role we want human beings to assume since it will mean a very big change for the society we know today. Therefore, this work shows that it is necessary to continue researching how AI can be adapted in the fairest and safest way in society.

Keywords: *artificial intelligence, ethics, autonomous car, autonomous car, moral agent, moral dilemma, ethical problems*

2. INTRODUCCIÓN

2.1 Objetivos:

El objetivo de este trabajo de investigación se centra en analizar tanto los criterios racionales supuestamente éticos que se implementan en los algoritmos y que de alguna manera comandan las decisiones éticas de los vehículos llamados autónomos como las problemáticas éticas que se pueden llegar a derivar del uso de la IA aplicada a la conducción.

Para ello se pretenden conseguir los siguientes objetivos específicos:

- Comprender lo que el desarrollo de la inteligencia artificial ha supuesto y en especial en el sector de la automoción.
- Analizar las problemáticas éticas que supone la integración de la inteligencia artificial en un coche autónomo y hacer una reflexión ética sobre la repercusión que tendrán en la sociedad, viendo en qué aspectos supondrá una mejora y en cuales supondrá un agravamiento.

2.2 Metodología

Para conseguir alcanzar los objetivos mencionados anteriormente se ha empleado el método de estudio deductivo. La información recogida en el trabajo ha sido extraída de la revisión de bibliografía proporcionada por expertos en la materia, en el que se han analizado artículos científicos, libros, artículos periodísticos y de revista.

Una de las desventajas que proporciona esta metodología al estar aplicada al sector de la ética en la Inteligencia artificial de los coches autónomos es la limitación de la información, ya que interfieren variables como los sentimientos, motivaciones y no se puede generalizar.

Aun con lo explicado anteriormente, esta metodología parece la más adecuada para este trabajo de investigación ya que gracias a la bibliografía utilizada se pretende analizar los criterios éticos que transformarán el sector de la automoción que se conoce hasta la fecha, posibilitando y favoreciendo un mejor entendimiento de los lectores de esta investigación.

2.3 Principales partes del TFG

El trabajo está estructurado en tres apartados, el primero trata el tema de la inteligencia artificial y como es aplicada en el sector de la automoción, el segundo de ellos trata la ética que lleva implementada la Inteligencia artificial y el tercer bloque trata sobre el coche autónomo y las problemáticas éticas que le acompañan.

En el primer apartado se incluye el capítulo 3 y 4. En este apartado se explica qué es la inteligencia artificial, cuál es su objetivo final y se desarrollan los dos principales tipos de IA y sus diferencias, además se explican los distintos modelos de investigación que existen. La IA se emplea en muchos sectores por eso nos hemos centrados en la que se aplica en el sector de la automoción, en concreto en el coche autónomo y se explican los diferentes niveles de conducción autónoma y los beneficios que a la larga supondrán para todos cuando el vehículo sea autónomo al 100%, ya que hasta la fecha no existe el coche automatizado al completo.

En el segundo apartado se incluye el capítulo 5. Este capítulo habla de la ética que se emplea en la IA, en qué principios está basada y cuales son sus consecuencias en la sociedad, tanto positivas como negativas.

En el tercer apartado se incluye el capítulo 6. En este capítulo se plantean una serie de problemáticas éticas que la implementación de la IA en los coches autónomos supone. Se estudian casos en los que la ética es fundamental ya que el vehículo se encuentra con dilemas morales. Se analizan las diferentes respuestas que se podrían dar a esos dilemas y se observan las preferencias a nivel mundial para posteriormente poder proponer unos principios éticos. En este tercer apartado también se hace una reflexión

ética sobre lo que la automatización al 100% de los coches supondría para la sociedad y cuáles serían sus repercusiones.

Al finalizar el tercer apartado se ha realizado un apartado de conclusiones.

3. INTELIGENCIA ARTIFICIAL

3.1 ¿Qué es la inteligencia artificial?

Según Lasse Rouhiainen, experto en inteligencia artificial y tecnologías disruptivas define la inteligencia artificial como “La capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano” (Rouhiainen, 2018)

El objetivo final de la inteligencia artificial (en adelante IA) es crear máquinas que sean capaces de pensar, responder y defenderse de la misma manera que lo hacen los humanos, aunque hay investigadores que defienden que la IA no tiene que pensar de la misma manera que lo hace la mente humana.

La IA nos plantea una revolución en el derecho, en la economía, en la medicina... Aunque lo veamos como algo que todavía está en desarrollo, la inteligencia artificial ya convive entre nosotros todos los días y más de lo que somos capaces de imaginar.

El desarrollo de la IA tiene que cooperar con la sociedad, ayudándola a progresar, a conseguir una igualdad entre todos. Esto supone asumir riesgos y aceptar que habrá equivocaciones que nos harán aprender. (Oliver, 2020)

La inteligencia artificial es capaz de aprender por sí misma, ya que emplea un sistema de aprendizaje basado en respuestas y acciones que ha realizado con anterioridad, además de las propias reglas que han sido implantadas por los programadores. La inteligencia artificial nos ha demostrado que no hay límites, el único límite existente es el de nuestra imaginación, además, esta IA se tiene que adaptar a nosotros y no al revés, no nos podemos adaptar nosotros a la inteligencia.

La integración de la Inteligencia Artificial, la automatización, el internet de las cosas, el *big data* y la computación, entre otras, en la sociedad está ocurriendo de forma muy acelerada y a gran escala, esta integración ha supuesto una transformación conocida como Cuarta Revolución Industrial (Perasso, 2016)

Que la IA sea capaz de aprender por sí misma es todo un logro, ya que gracias a ella se conseguirá revolucionar el mundo que conocemos y solucionar muchos problemas para los que ahora mismo no tenemos solución, pero también nos plantea dudas e incluso miedo ¿Llegará la IA a ser superior a nosotros y llegar a controlarnos?

3.2 Tipos de inteligencia artificial

Dentro de la IA encontramos dos tipos de inteligencia principalmente según John Searle, el primero de ellos es la IA fuerte y el segundo de ellos es la IA débil.

La IA fuerte también conocida como general es aquella “*Ciencia e ingeniería que permitirá replicar la inteligencia humana mediante máquinas*” (López de Mántaras Badia, 2015) mientras que la IA débil o también conocida como estrecha es aquella “*Ciencia e ingeniería que permiten diseñar y programar ordenadores de forma que realicen tareas que requieren inteligencia*” (López de Mántaras Badia, 2015)

Toda la IA que conocemos hasta ahora es débil o estrecha, es decir, los ordenadores están programados para realizar tareas específicas, simulando una mente humana, sin conseguirlo. Por ejemplo, un programa que juega al ajedrez a alto nivel es incapaz de jugar a otros juegos, aunque sean más sencillos. Para esto sería necesario diseñar e implementar un programa distinto para que ese ordenador fuera capaz de jugar también a las damas. Es decir que no puede aprovechar su capacidad para jugar a un juego y adaptarlo a otro, únicamente muestran comportamiento inteligente en el ámbito específico para el que han sido diseñadas.

En los seres humanos no sucede así, ya que cualquier persona que juegue al ajedrez podría aprovechar sus conocimientos de este juego para aprender, en pocos minutos, a jugar a las damas perfectamente.

Esto no quita que se ha demostrado que los avances de la IA débil son capaces de realizar tareas mejor que los propios seres humanos, como en los casos en los que haya que realizar cálculos de fórmulas lógicas con muchas variables, juegos de lógica, diagnósticos médicos y muchos otros aspectos relacionados con la toma de decisiones.

Todavía no se ha conseguido implementar la IA fuerte o general, e incluso el filósofo Searle en un artículo crítico con la IA publicado en 1980 (Searle, 1980) asegura que nunca se llegará a conseguir implementar esta misma. La IA fuerte implicaría que un ordenador convenientemente diseñado no simularía una mente humana, sino que sería una mente y por tanto debería ser capaz de tener una inteligencia igual o incluso superior a la humana.

Las principales diferencias entre la Inteligencia Artificial fuerte y la Inteligencia Artificial débil se resumen en la siguiente tabla:

Tabla 1: Diferencias entre IA fuerte e IA débil

Inteligencia Artificial Fuerte	Inteligencia artificial Débil
Actúa como una mente humana	Simula una mente humana
No existe en la vida real	Existe en la vida real
Se enfrenta a problemas abiertos	Se enfrenta a problemas concretos
Actitud proactiva	Actitud reactiva
Razonan	Computan
Capacidades similares a las de un ser humano: Pueden aprender	No pueden aprender, solo pueden repetir
Capacidad de autoprogramarse	Necesitan programación de un ser humano
Muchas redes neuronales	Pocas redes neuronales
Adecuación a nuevas circunstancias	No pueden adecuarse a nuevas circunstancias

Fuente: Huawei

3.3 Modelos de investigación:

La inteligencia artificial cuenta con distintos modelos de investigación: Simbólico, conexionista, corpóreo y situacional y finalmente evolutivo.

3.3.1 Modelo de investigación simbólico

El modelo de investigación simbólico ha sido el paradigma dominante hasta finales de los años 80, también se le conoce ahora como GOF AI (Good Old-Fashioned Artificial Intelligence). Se basa principalmente en el razonamiento lógico y actúa bajo las representaciones simbólicas de tareas cognitivas, como la toma de decisión (Puede tomar resoluciones basadas en modelos como el árbol de decisiones). Estas representaciones simbólicas están basadas primeramente en la lógica matemática y su proceso es determinista, es decir, las mismas entradas producirán las mismas salidas, no se contempla el azar o la incertidumbre (Wikipedia, 2020).

Al tomar decisiones basadas en modelos como el árbol de decisiones, es fácil saber para los seres humanos cómo se ha llegado a una conclusión ya que el procedimiento es transparente e interpretable, esto permite que la conclusión pueda ser revisable y evaluable (Coeckelbergh, 2021)

El modelo de investigación simbólico puede tomar decisiones complejas, está basado en la hipótesis de Sistema de Símbolos Físicos (SSF) Consiste en un “Conjunto *de entidades denominadas símbolos que, mediante relaciones, pueden ser combinados formando estructuras más grandes y que pueden ser transformados aplicando un conjunto de procesos*” (López de Mántaras Badia, 2015)

El modelo simbólico no necesita interactuar con el entorno real, actuar sobre él y formar parte de un cuerpo. Es decir, la IA simbólica opera con representaciones abstractas del mundo real que se modelan mediante lenguajes de representación basados principalmente en la lógica matemática y sus extensiones. (López de Mántaras, 2019)

3.3.2 Modelo de investigación conexionista

El modelo de investigación conexionista surgió como una alternativa al modelo simbólico en los años 80 y a diferencia de este último, mediante una serie de datos es capaz de aprender de manera autónoma, aunque no es incompatible con la hipótesis SSF. Se basa en la hipótesis de que a partir de la creación de un gran número de redes neuronales que están conectadas entre sí y procesan la información paralelamente surge la inteligencia artificial.

Se usan a menudo para el aprendizaje automático, y a diferencia del modelo simbólico, no es posible saber cómo ha sido el procedimiento hasta llegar a la conclusión y que la conexión de las redes neuronales tiene varias capas por lo que hace que no se pueda ver lo que pasa entre las capas intermedias entre el input y el output. Que no se pueda saber cómo ha sido el procedimiento no

permite la revisión, interpretación y evaluación de las conclusiones por parte de los seres humanos (Coeckelbergh, 2021)

El modelo conexionista al igual que el modelo simbólico tampoco necesita interactuar con el entorno real, actuar sobre él y formar parte de un cuerpo.

3.3.3 Modelo de investigación evolutivo

El modelo de investigación evolutivo es aquel que como él mismo propio nombre indica puede evolucionar y desarrollarse. Tuvo su desarrollo en los años 60 y para ello utilizaron algoritmos genéticos. A través de estos algoritmos genéticos se intenta imitar la evolución del ser humano y que el dispositivo sea capaz de aprender y mejorar automáticamente.

El modelo de investigación evolutivo para poder aprender automáticamente necesita interactuar con el entorno real formando parte de un cuerpo. Interactuando con el entorno conseguirá obtener representaciones abstractas. (López de Mántaras Badia, 2015)

3.3.4 Modelo de investigación corpóreo y situacional

El modelo de investigación corpóreo y situacional se centra en las interacciones con el entorno y en las tareas motoras para poder dar soluciones. En los años 90 el robótico Rodney Brooks creó Cog, un robot humanoide que resolvía los problemas interactuando con el entorno y no mediante representaciones simbólicas. (Koekelberg, 2021)

Muchos investigadores consideran esencial la representación interna para poder llegar a alcanzar la inteligencia artificial fuerte o general. Son defensores de que *“el cuerpo da forma a la inteligencia”* (López de Mántaras Badia, 2015)

4. INTELIGENCIA ARTIFICIAL APLICADA AL SECTOR DE LA AUTOMOCIÓN

El sector de la automoción se ha caracterizado por ser uno de los sectores más dinámicos y con un alto componente tecnológico, que, en los últimos años ha sufrido una gran transformación haciéndose mucho más competitivo (Carbajal, 2010). Esta nueva competitividad ha hecho que se busquen nuevas formas de desarrollar la tecnología, llegando a reestructurar y cambiar la industria tal como se conocía introduciendo los coches autónomos.

En el sector de la automoción la IA todavía no está muy implantada, solo el 10% de las empresas que trabajan en este sector la han implantado, y el avance de esta misma implantación es de forma lenta y pausada, desde el 2017 solo se ha incrementado el número de empresas que utilizan la IA en un 3%. (Capgemini, 2019)

Además, no se está implantando en todos los países de la misma manera, el país con más IA implementada en el sector de la automoción es EE. UU. con un 25% de las empresas, seguido de Reino Unido con un 14% y finalmente Alemania con un 12%. (Capgemini, 2019)

4.1 ¿Dónde encontramos inteligencia artificial en automoción?

Las tecnologías basadas en IA que ya podemos encontrar en el sector de automoción están enfocadas a prestar ayudas a la conducción asistida y serán los precursores, hasta poder llegar al coche autónomo al 100%:

- Aparcamiento automático
- Alertas por salirse del carril o pisar línea
- Advertencia de riesgo de colisión
- Sistemas de frenado automático
- Autollamada en caso de accidente a emergencias
- Sistemas de navegación

- Asistente conversacional por voz - este tipo de inteligencia artificial es una de las aplicaciones más usadas, son capaces de reconocer el lenguaje natural humano para identificar lo que el conductor solicita

Para que los vehículos autónomos sean capaces de operar por sí mismos o de ayudar a la conducción es necesario que cuenten con:

1) Sensores por todo el vehículo para que sea capaz de detectar lo que pasa a su alrededor y en su entorno y a su vez sea capaz de detectar el propio movimiento. Estos sensores pueden ser cámaras, sensores de ultrasonido y ondas de radio (sirven para la medición de distancias entre objetos)

2) Ordenador de a bordo, este ordenador permitirá tener acceso a la información relevante del vehículo como la ubicación de este, un GPS, ofrece una orientación y un conocimiento de las rutas. A través de este ordenador de a bordo la IA puede aprender de manera automática. (Li, Díaz, Morantes, & Dorati, 2018)

4.2 Grados de automatización en los coches autónomos

La Sociedad de Ingenieros de la automoción en inglés Society of Automotive Engineers o SAE clasificó en el año 2014 los grados de automatización de conducción que pueden llegar a tener los vehículos. El SAE clasifica estos grados en 6, según el nivel de guiado automático del vehículo.

4.2.1 Nivel 0: La conducción no tiene asistencia alguna.

El nivel 0 es el nivel inicial en el que el conductor es el que se encarga de hacer todas las labores como frenar, acelerar, aparcar... En este nivel no hay asistencia y el conductor tiene todo el control de las acciones, aunque puede tener sistemas de detección de objetos que ayuden a la conducción. La mayoría de los vehículos que se encuentran en movimiento y en el mercado pertenecen a este nivel.

4.2.2 Nivel 1: Hay conducción asistida para el conductor.

El nivel 1 ya cuenta con sistemas de conducción asistida que pueden ayudar al conductor. Esta automatización puede verse en forma de sistemas que toman el control sobre la dirección o sobre la velocidad, aunque no pueden darse de forma simultánea.

También pueden contar con piloto automático adaptativo (El vehículo frena o acelera en función del tráfico de forma automática) y con sistemas de centrado de carril, que controlan la invasión del carril sin que esta sea premeditada. Aunque estos sistemas de conducción asistida actúen en escenario específicos el conductor es el que tiene el control total sobre las acciones.

4.2.3 Nivel 2: Automatización parcial.

En el nivel 2 el vehículo puede tomar el control, siempre bajo la supervisión del conductor, de los sistemas de control del volante, de la aceleración y del frenado de manera simultánea. Esta toma de control es solo posible en determinadas ocasiones y durante periodos de tiempo muy cortos. El conductor podrá tomar el control del vehículo en cualquier momento ya que este no es capaz de identificar aspectos como el tráfico, accidentes, etc....

Los primeros tres niveles cuentan con características de soporte al conductor, aunque éste es responsable de controlar todo lo que sucede a su alrededor, como pueden ser otros vehículos, condiciones climatológicas, estado de las carreteras...

A partir del 3 nivel ya se encuentran funciones automatizadas y a partir de ese nivel empiezan los problemas éticos y desde la perspectiva legal ya que en caso de accidente no está regulado si es el responsable sería el conductor o la compañía que haya fabricado el coche y hay que empezar a dotar de un marco teórico.

4.2.4 Nivel 3: Automatización condicional.

En el nivel 3 el vehículo ya cuenta con funciones automatizadas y ya empieza a poder conducirse solo durante largos periodos de tiempo en la gran mayoría de situaciones. Entre estas funciones automatizadas se encuentran aquellas como la del asistente de conducción, poder ir a una velocidad constante, asistencia en el aparcamiento, etc. Pero en este nivel todavía se necesita que el conductor esté atento a lo que pasa a su alrededor para poder tomar el control de la conducción si hay algún problema o imprevisto que el coche no controla.

4.2.5 Nivel 4: Alto nivel de automatización para la conducción.

A partir de este nivel el vehículo ya es capaz de conducirse completamente por sí solo y no será necesaria la supervisión del conductor ante imprevistos exceptuando cuando las condiciones climatológicas no sean las adecuadas, es decir, abundante niebla, tormentas, lluvias torrenciales o nieve.

En este nivel hay dos tipos de automóviles: los primeros, son aquellos que cuentan con mandos para la posible conducción manual si así se requiere y en el caso de que se quiera volver a la conducción autónoma se tendrá que activar la conducción autónoma; los segundos, son aquellos tipos de vehículos que no cuentan con mandos para la posible conducción manual, en el caso de que haya un imprevisto de condiciones en las que haya poca visibilidad el vehículo se detendrá y cuando las condiciones sean más favorecedoras se reanudará la marcha.

Todavía no hay vehículos en el mercado disponibles para la compra que cuenten con un nivel 4 de automatización, aunque sí que existen prototipos como el modelo S de Tesla y el modelo Waymo de Google, aunque se han estado haciendo pruebas con estos vehículos por carreteras convencionales todavía no operan de forma totalmente libre ya que cuentan con un sistema de protección a forma de respaldo en el caso de que la conducción autónoma falle.

4.2.6 Nivel 5: Nivel de automatización completo para la conducción.

El nivel 5 de la clasificación del SAE es el último nivel y consiste en la automatización completa, el coche opera completamente sin la necesidad de la intervención del conductor. El vehículo puede conducirse por sí solo en todas las condiciones y superar todos los retos que se presenten.

Puede parecer que el salto del nivel 4 al nivel 5 es bastante simple, sin embargo, es el cambio de nivel más complicado y todavía no se ha llegado a implementar e incluso a veces parece que se llega a dudar si alguna vez se llegarán a implementar vehículos con este nivel totalmente robotizados en los que se puede prescindir de volante y pedales.

Imagen 1: LEVELS OF DRIVING AUTOMATION

		SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?		You are driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			You are not driving when these automated driving features are engaged – even if you are seated in "the driver's seat"		
		You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
What do these features do?		These are driver support features			These are automated driving features		
		These features are limited to providing warnings and momentary assistance	These features provide steering OR brake/acceleration support to the driver	These features provide steering AND brake/acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met	This feature can drive the vehicle under all conditions	
Example Features		<ul style="list-style-type: none"> • automatic emergency braking • blind spot warning • lane departure warning 	<ul style="list-style-type: none"> • lane centering OR • adaptive cruise control 	<ul style="list-style-type: none"> • lane centering AND • adaptive cruise control at the same time 	<ul style="list-style-type: none"> • traffic jam chauffeur 	<ul style="list-style-type: none"> • local driverless taxi • pedals/steering wheel may or may not be installed 	<ul style="list-style-type: none"> • same as level 4, but feature can drive everywhere in all conditions
For a more complete description, please download a free copy of SAE J3016: https://www.sae.org/standards/content/J3016_201806/							

Fuente: SAE

Hasta la fecha ya hay automóviles en el mercado con un nivel de automatización de la conducción del nivel 3 y ya hay en prueba automóviles con nivel de automatización de la conducción en el nivel 4, aunque todavía no han sido sacados en el mercado para la población en general, solo están en periodo de prueba.

Los retos que involucran a los vehículos con un nivel de automatización de nivel 5 son muy amplios y complejos. En la actualidad la conducción autónoma al 100% no es legal todavía.

4.3 Beneficios que traerán los coches autónomos

La realidad de los vehículos automatizados de forma completa es un futuro no tan lejano, por lo que se puede estimar que beneficios traerán para la sociedad.

- No será necesario tener los coches en propiedad, ya que los vehículos se usarán bajo demanda. Esto permitirá a los usuarios ahorrarse el dinero que gastarán en mantenimiento y propiedad, además de no tener preocupaciones por el robo del vehículo y de tener que buscar sitio para aparcar.
- Al empezar a ser un servicio bajo demanda y no tener los coches en propiedad, las ciudades se adaptarán a las necesidades de los peatones y otros vehículos como las bicicletas y los patinetes, ya que no habrá tanto espacio reservado para el aparcamiento.
- Se prevé que la mayoría de los coches autónomos sean eléctricos por lo que los niveles de contaminación se verán reducidos de forma significativa.
- Un estudio de la revista literaria The Atlantic estima que gracias a los coches autónomos el número de accidentes mortales en carretera se reducirían al 90% ya que la principal causa de los accidentes es el ser humano.

5. LA ÉTICA EN EL DISEÑO DE LA INTELIGENCIA ARTIFICIAL

La ética en el diseño de la inteligencia artificial tiene que velar por el impacto que va a tener en la vida de las personas y por las transformaciones que lleva su implantación, tanto en la economía como en la sociedad.

Es importante delimitar los principios éticos que se han tenido en cuenta en la implementación de sus modos de funcionamiento ya que, en la actualidad, las máquinas sustituyen a los humanos, no son complementos (Brynjolfsson y McAfee, 2014)

5.1 Principios éticos

La Inteligencia artificial está revolucionando el mundo que conocemos y su aplicación supone riesgos por lo que es importante que cuente con un marco ético basado en 5 principios éticos (Cortina, 2019)

- Creación de una IA libre y justa.
- Propósito de no-maleficencia, que haya seguridad y privacidad, y no se infrinja daño de forma deliberada o intencionada.
 - Orientada a la búsqueda de beneficencia, para los seres humanos como para el medioambiente, inquiriendo en un impacto positivo.
 - Principio de autonomía del ser humano: la IA no puede coaccionar o subordinar al ser humano.
 - Principio de aplicabilidad: El algoritmo debe ser transparente y los usuarios pueden entender su funcionamiento.

5.2 Beneficios

Gracias a la Inteligencia Artificial y a la evolución de ésta, los beneficios para la sociedad han sido numerosos. Entre ellos se encuentran la optimización de los procesos de la cadena valor, la optimización sobre todo en el análisis y en el entendimiento de datos.

La IA permite conocer mejor los gustos de las personas y sus inquietudes, gracias a este conocimiento los productos y servicios que se le pueden ofrecer son mucho más adecuados a sus necesidades y conseguirá satisfacer aquellas que hasta la fecha no han podido satisfacerse

5.3 Riesgos

Uno de los mayores riesgos con los que cuenta el impacto de la IA en la sociedad es el de condicionar la forma que tienen de relacionarse los humanos y las interacciones que hay entre ellos, la privacidad de los usuarios se verá afectada y es muy probable que la discriminación y los prejuicios aumenten, consiguiendo tener una mayor desigualdad e injusticia en las sociedades.

Adicionalmente se acentuarán brechas ya existentes en nuestra sociedad como la brecha entre ricos y pobres, los poderosos y los vulnerables. Habrá una transformación al completo de la economía que conocemos y muchos de los empleos existentes desaparecerán y con ello la pérdida de empleo será considerable.

Es necesario tener en cuenta el impacto que tendrá en el medioambiente, ya que el consumo de energía será mucho mayor y con ello la contaminación aumentará.

Esto que hemos mencionado anteriormente es el impacto y el alcance que tendrá la IA, pero es importante preguntarse quiénes serán los afectados y de qué manera les afectará. ¿Quiénes tendrán acceso a este tipo de Inteligencia artificial? ¿Quiénes se verán beneficiados y quienes se verán excluidos?

5.4 ¿Deberían ser las máquinas consideradas agentes morales?

La agencia moral es el entendimiento de uno mismo y de otros agente morales, de la repercusión de sus acciones (conscientes, autónomas e intencionadas) cuando estás están basadas en creencias morales, emociones, sentimientos, objetivos...

Estas acciones deben estar motivadas por razones morales, deben tener el propósito de llevar la acción a cabo de forma voluntaria y dándose cuenta de lo que están haciendo. (Monasterio, 2019)

¿Habrá que dotar de agencia moral total a las máquinas? o es una simple máquina
¿Serán capaces de discernir entre el bien y el mal? Las acciones de estas máquinas

tienen consecuencias en el mundo y consecuencias morales, aunque no tengan capacidad de sentir, de tener emociones y de ser responsables de sus acciones.

Las máquinas con Inteligencia Artificial pueden llegar a tomar decisiones más racionales que los humanos ya que no interfieren los sentimientos y emociones, pero detrás de estas decisiones morales que la máquina toma están los valores éticos de los seres humanos. En consecuencia las acciones tanto buenas como malas que realicen las máquinas son en función de nuestros criterios éticos, sin tener plena consciencia plena voluntad y una intencionalidad positiva.

Por lo tanto, según Wendell Wallach y Colin Allen sería lógico y racional dotar a las máquinas con Inteligencia Artificial de una “moralidad funcional” (Coeckelbergh, 2021) A través de esta moralidad funcional las máquinas con Inteligencia Artificial podrían ser capaces de hacer un análisis sobre la repercusión de sus acciones y decisiones. Esta moralidad estaría representando las acciones morales y características que tomarían los seres humanos.

Los filósofos Luciano Floridi y J.W Sanders a diferencia de Wallach y Allen dotarían a las máquinas con IA de una “moralidad a-mental” dependiendo de su nivel de interactividad y autonomía. Floridi y Sanders defienden que cuando una Inteligencia Artificial tiene autonomía de sus programadores y diseñadores hay que dotarlo de una intención humana (Coeckelbergh, 2021)

Las máquinas con IA son capaces de tomar decisiones y de realizar acciones, pero no son conscientes de las repercusiones que tendrán, por lo tanto, aunque sus acciones tengan consecuencias éticas no pueden hacerse “**moralmente responsables**” de las acciones que han llevado a cabo.

Que las máquinas no sean moralmente responsables de las decisiones y acciones que llevan a cabo nos obliga a no poder catalogarlas como agentes morales, ya que no realizan acciones de forma voluntaria, no precisan de libre albedrío ni de emociones, no pueden reflexionar sobre las consecuencias de realizar una acción u otra.

La conducta de las máquinas está definida por un programa informático que determina lo que es capaz de realizar y qué respuesta va a dar, estas acciones son automáticas y predeterminadas por lo que no hay reflexión sobre el porqué de las acciones y no cuentan con libertad. (Monasterio, 2019)

Plantear un marco teórico para la Inteligencia artificial nos plantea cuales son los códigos éticos correctos y nos hace plantearnos si la sociedad en la que vivimos los cumple y por lo tanto si esta sociedad es justa, nos plantea cuál es el sentido de la vida y cómo la valoramos. Este código ético tiene que adaptarse a la velocidad de cambio de la sociedad que como estamos pudiendo observar cada vez es más veloz y a cada sociedad. Cada sociedad tiene un código distinto y puede que en algunos sea el mismo, pero puede que en otras sociedades sea el opuesto totalmente, por lo que hay que reconocer esta pluralidad y diversidad. Hay que elaborar un plan en el que intervengan no solo programadores, sino que también deben intervenir filósofos, juristas...(Nieto,2019) para que la adaptación sea lo más segura y justa posible.

Hasta el momento hemos observado la ética aplicada a la Inteligencia Artificial desde el punto de vista de las máquinas y cómo les afectaría. Pero no hay que olvidar que también hay que tener en cuenta quiénes tendrán acceso a estos datos y cómo los utilizarán.

6. GRANDES PROBLEMÁTICAS ÉTICAS

Para que el coche autónomo sea capaz de operar sin la intervención humana o con muy poca participación humana necesita ser capaz de percibir su entorno gracias al análisis de miles de datos y que serán procesados por algoritmos.

Las tomas de decisión en los vehículos autónomos son más complejas que la simple cumplimentación de las leyes, muchas de estas decisiones deben necesitar de un marco y un sentido ético, y esta es la dificultad a la que se enfrentan ya que no es una simple programación y es una tecnología muy innovadora de la que todavía se está investigando y no hay respuestas exactas.

La participación de la IA en el día a día de los seres humanos ha supuesto muchos avances en actividades que resultaban complicadas, llegando a hacer la vida más cómoda y segura. Pero la inminente cooperación de la IA en los coches totalmente autónomos va a tener consecuencias en las sociedades que se conocen hasta el momento y es inevitable que se encuentren nuevos retos, conflictos y surjan una serie de problemáticas éticas para las que, hasta el momento, no se tiene respuesta.

El vehículo autónomo con nivel de automatización 5 aportará más seguridad que aquellos que son conducidos por personas, además de que aportarán mucha mayor eficacia. Cuantos más coches autónomos haya mayor será esta seguridad. La seguridad de los coches autónomos no será del cien por cien ya que hay más variables en juego que pueden suponer un riesgo para la seguridad aparte de los otros coches, cómo animales, peatones, ciclistas, tiempo meteorológico...

Hay que programar al vehículo para evitar en mayor medida cualquier tipo de accidente y nos plantea varias preguntas con contenido ético ¿Deben los coches autónomos estar programados para evitar el mayor número de siniestros? ¿Qué debe priorizar el vehículo cuando la opción se encuentre entre dos atropellos? ¿Deberán decidir salvar a sus pasajeros?, ¿Deberán atropellar a personas de edad avanzada o a aquellas personas más jóvenes? Todas estas preguntas nos plantean cuáles serán los principios morales que van a estar implementados.

Estas problemáticas éticas las vamos a diferenciar en cuatro bloques distintos: La privacidad, la transparencia, la seguridad y finalmente la responsabilidad.

6.1 Privacidad

La privacidad es una de las mayores problemáticas éticas con las que cuenta el coche autónomo, el vehículo va a recolectar mucha información del usuario, tanto de manera directa como de manera indirecta.

No se conoce la finalidad de los datos recogidos ni con qué propósito u objetivo se van a utilizar. Se conocerán los horarios, rutinas, que se hace en cada momento, lugares de trabajo, domicilio... Mucha de esta información se obtendría sin que el usuario lo haya buscado o comentado. Esto supone una invasión de la privacidad y pone en riesgo la aceptación de los vehículos.

El uso ético de la Inteligencia artificial exige que la forma en la que son seleccionados, tratados y procesados los datos permita la privacidad de los usuarios, sabiendo en todo momento el usuario, qué datos son los recopilados, qué sucede con esos datos y cómo están siendo procesados. El único problema es que a medida que la Inteligencia Artificial va avanzando, cumplir este código ético cada vez es más difícil, ya que los usuarios no saben con exactitud qué está pasando con sus datos y cómo están siendo procesados.

Que el usuario no sepa qué datos y de qué manera se están tratando supone una forma de manipulación y explotación. Estos datos se utilizarán para analizar nuestros gustos, ideas... e incluso vendérselo a terceros. El peligro es que esta Inteligencia Artificial puede llegar a manipular o ser una vigilancia constante sin que nos demos cuenta.

¿Qué supone la decisión de salvar a una persona frente a otra en base a características como haber realizado buenas obras o ser buen estudiante?. Esta diferenciación supone un control por parte del vehículo (e indirectamente de aquellos que traten la información) ya que todo el mundo estaría identificado de una manera totalitaria.

La necesidad intrínseca de tener que conocer cuantos más datos e información para tomar “la mejor decisión posible” puede llegar a derivar en una sociedad totalitarista ya que habría control sobre cada uno de los ciudadanos y sus acciones.

Hasta la fecha no ha habido ningún dispositivo electrónico que no haya sido hackeado, por lo que la posibilidad de que un vehículo autónomo sea hackeado es muy alta con la única diferencia de que en este caso es mucho más peligrosa ya que pone en peligro tanto la información personal del usuario y su data, como el control del vehículo y posibles vidas. Si los datos que está analizando para tomar las decisiones han sido modificados y sesgados las decisiones que el vehículo autónomo tomará también estarán sesgadas, no serán correctas e incluso puede que incumplan los principios éticos. ¿Qué pasa si deciden tomar el control del automóvil y provocar un accidente? En ese caso ¿Quién sería el responsable? ¿Cómo se puede prevenir? ¿Cómo puede reaccionar el vehículo?

A medida que la Inteligencia Artificial va tomando más importancia en nuestras vidas también lo hace nuestra dependencia en ella y nos volvemos más vulnerables y en el caso de los coches autónomos las vulnerabilidades de la Inteligencia artificial también se convierten en la vulnerabilidad de los seres humanos.

Para paliar estas inseguridades, producidas sobre todo por la inseguridad jurídica y que los futuros usuarios cada vez vayan teniendo más confianza en la IA en 2018 la Comisión decretó una estrategia sobre IA y en el 2019 publicaron una normativa con siete instrucciones sobre cómo debería ser una IA fiable. (Comisión Europea, 2020)

6.2 Transparencia

Otra de las grandes problemáticas éticas que está muy ligada a la de la privacidad es la transparencia.

La inteligencia artificial tiene sesgos y no es perfecta, esta es programada y diseñada por personas que tienen valores propios y valores éticos distintos a cualquier persona y pueden tener sesgos discriminatorios. Pero estos sesgos no solo vienen dados por sus programadores y diseñadores, la Inteligencia Artificial procesa miles de datos y macrodatos de bases de datos, de resultados de búsqueda, etc. Dependiendo de la ubicación habrá unos sesgos u otros que son reales en la vida real y hará que las

decisiones que se tomen sean injustas. La Inteligencia artificial tiene un aprendizaje automático, este aprendizaje está basado en el aprendizaje gracias a unos patrones o reglas que se repiten y el problema es que puede que la Inteligencia artificial puede fortalecer estos sesgos.

La solución a estos sesgos sería ser cuidadosos con la selección de los datos que se van a analizar y la forma en la que los algoritmos van a tratarlos. Es muy importante que haya una transparencia en los algoritmos que se implementan ya que es necesario ver cuales son los valores y criterios éticos en los que se sustentan las decisiones automatizadas.

La mayoría de las veces estos sesgos no están buscados ni son intencionados, incluso puede que los programadores y desarrolladores no estén al tanto. Por eso es importante revisar las distintas fases de la Inteligencia Artificial para encontrar posibles sesgos y de esta manera evitar posibles consecuencias graves. Las distintas fases son la de diseño, prueba y aplicación, aunque la que más relevancia tiene es la de diseño ya que es la base de todo.

La fase de diseño, consiste en la selección de datos que los algoritmos de la Inteligencia Artificial van a analizar. Esta selección de datos puede ser no representativa o estar incompleta, por lo que las decisiones tomadas estarán sesgadas.

En la fase de diseño también es muy importante que haya diversidad entre los desarrolladores y programadores, tanto de sexo, edad y raza para que ninguno de los colectivos se vea excluido y no afecten las experiencias personales evitando así en su mayor medida los prejuicios. Aunque durante la fase de prueba se encuentren causas de correlación en los datos eso no implica que haya causalidad (Koekelberg, 2021)

Aunque seleccionemos con el mayor cuidado posible el conjunto de datos, las predicciones y decisiones que tome la Inteligencia artificial nunca van a estar libres de sesgos, ya que estos datos y el diseño y programación de los algoritmos han sido creado por humanos y son un reflejo de la sociedad, por lo que nunca estará libre de sesgos, aunque sí que se puede intentar minimizar el impacto de estos. (Digital Europe, 2019)

El uso de la IA también tiene sus vulnerabilidades y esto presenta amenazas que hace que los usuarios se sientan inseguros y con inquietudes que puedan dificultar la

implementación a largo plazo ya que la decisión racional individual no existiría y eso no es justo.

Es necesario que la transparencia sea conocida, para así poder tener la oportunidad de elegir los diferentes algoritmos que se quiere que estén involucrados. De esta manera, si no hay transparencia, habría sesgos por poder adquisitivo, antecedentes y el mercado estaría falseado.

En el caso de vender un coche, si se analizan datos como el poder adquisitivo o los gustos tecnológicos y preferencias, los precios se podrían elevar no siendo igual para todos y de la misma manera pasaría con los servicios de VTC, podrían analizar la urgencia de asumir ese servicio, al lugar al que te diriges y subir los precios en función de tus características, por lo que la decisión racional individual no estaría existiendo y eso no es justo.

Relacionando estas dos primeras problemáticas se nos plantean preguntas de concepción política, ¿Qué tipo de democracia queremos? ya que todos los datos de los usuarios estarán disponibles e incluso no se sabrá qué instituciones tienen acceso a nuestros datos y con qué fines se están utilizando. Si toda la información de un individuo, el estado o el sistema político tuviera un control total.

Esta revolución tecnológica no está sirviendo a la democracia, ya que cada vez los usuarios serían menos libres y están más condicionados. Por eso es importante que para que suponga un crecimiento, todos los colectivos, entre ellos los gobiernos, las empresas, los individuos tengan que colaborar, implementando los cambios y reformas necesarias dando el tiempo necesario para adaptarse. (Aurik, 2017)

6.3 Seguridad

Se podría decir que uno de los mayores retos que llevan asociados los vehículos autónomos son las decisiones que toman ante situaciones de peligro para preservar la seguridad.

A diferencia de los seres humanos, como hemos explicado anteriormente, estas acciones no se pueden considerar fruto de una reacción ya que la decisión ha sido premeditada y analizada.

Desde los diferentes estudios vemos que nos podemos enfrentar a cinco tipos de dilemas morales que conllevan una toma de decisión:

6.3.1 Primer dilema moral: Niño pequeño Vs persona mayor

Supongamos que un coche autónomo se encuentra en la situación de atropellar a una persona mayor o a un niño pequeño. En cualquier decisión que tome habrá un fallecimiento.

Uno de los pensamientos con más peso podría ser salvar al niño pequeño ya que tiene toda la vida por delante y le queda mucho por vivir y experimentar a diferencia de la persona mayor que ya ha vivido y ha tenido muchas más experiencias que el niño pequeño.

Pero cualquiera de las dos decisiones es éticamente incorrecto, ya que según uno de los códigos éticos del Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) se estaría haciendo una discriminación de edad. Este código no permite “hacer distinciones de factores como la raza, la religión, el sexo, la discapacidad, la edad o su país de origen”. (IEEE, s.f). Según el imperativo categórico Kantiano explicita que todas las personas merecen ser tratadas como fines en sí mismos y nunca como medios. Aunque la vida de la persona mayor es igual de válida que la vida del niño pequeño, salvar a este último sería la decisión que parece más correcta.

¿Cuál es la decisión éticamente correcta? Por eso es importante dotar de un sentido ético a los coches autónomos. La situación comentada con anterioridad, al estar en igualdad de condiciones, habría una discriminación al elegir colisionar contra el niño o con la persona mayor, pero si las condiciones son distintas sí que se podría hacer distinciones. Estas distinciones se hacen en función de una justicia. Las elecciones no pueden tener una justicia ciega y que se elija una opción al azar. En el dilema del niño pequeño y la persona mayor, una de estas decisiones se puede basar en tomar aquella decisión que favorezca a la persona en mayor desventaja, pero es necesario tener en cuenta que si la ley empieza a regirse por aquellas reglas en las que las probabilidades de sobrevivir son menores las personas cada vez empezarán a incumplir más las leyes, como veremos en otros dilemas.

La reflexión desde el punto de vista de la tradición deontológica, cada ser humano es digno en sí mismo y por lo tanto no es de recibo que a un autómata se le implemente una decisión de si es mejor quitar la vida a un niño o a una persona mayor, ya que los dos son sujetos de dignidad.

6.3.2 Segundo dilema moral: Una persona Vs varias personas

Teniendo en cuenta que el vehículo se encuentra en la situación de colisionar contra una persona o contra varias personas, la acción puede parecer bastante intuitiva ya que cuantas más personas se salven mejor. Desde el punto de vista utilitarista se estaría maximizando el beneficio de la mayoría.

El dilema se complica cuando la decisión se encuentra entre colisionar contra una persona pero es un familiar o un conocido frente a varias personas desconocidas, porque ahí intervienen los sentimientos de los humanos, que aunque se salven menos vidas preferirán salvar al familiar. Desde el punto de vista deontológico no se estaría considerando a cada persona como digno en sí mismo, ya que no todos serían considerados iguales y desde el punto de vista utilitarista no se estaría maximizando el beneficio de la mayoría. Estas decisiones están basadas en el sentido de justicia que se le dote y se le programe a la Inteligencia Artificial.

Rawls defiende que la justicia tiene que ver con la equidad y no con el igualitarismo. La justicia no tiene que ver con el beneficio de la mayoría social, porque como dijo Kant todos los seres humanos deben ser tratados como fines en sí mismos y nunca como medios. La equidad tiene que ver con la proporcionalidad racional, por lo tanto, la decisión que tome el vehículo tendría que tener en cuenta más perspectivas para discernir, ya que una decisión será tanto más justa cuantos más criterios racionales se implementen.

En este dilema también se podría incluir la colisión contra otro vehículo que sea capaz de soportar mejor el impacto, en el caso de que un coche autónomo tenga que colisionar contra algún otro automóvil que esté a su alrededor y tenga que elegir uno de ellos, para evitar el mayor número de siniestros colisionará contra aquel que sea más seguro y pueda soportar mejor el golpe antes que aquel vehículo que sea más pequeño e inseguro, ya que el daño sería menos grave.

6.3.3 Tercer dilema moral: Animales

En el caso de la aparición de animales en la vía, las opciones que tiene un vehículo son las de: atropellar al animal o desviarse del carril.

En el caso de que el animal sea pequeño como un conejo las probabilidades de accidente son menores que en el caso de que sea un animal grande con una gran envergadura. Si la colisión es contra un animal grande las consecuencias pueden ser graves y si el vehículo es capaz de analizar el entorno y ve que hay una posibilidad de desviarse sin causar accidentes esa sería la decisión más correcta, pero en el caso de que la vía esté muy transitada sería mejor tener una colisión frontal intentando reducir la velocidad lo máximo posible para que no el vehículo no de vueltas de campana y haya el menor número de afectados posible.

La reflexión utilitarista busca la maximización del beneficio de la mayoría y que el principio de actuación trate a las personas en sí mismas como fines y no como medios. Pero a diferencia de los seres humanos, los animales, al no tener capacidad de razonar, sí que pueden ser tratados en sí mismos como medios por lo que la opción sería colisionar contra el animal.

Pero este razonamiento no es correcto ya que los animales también son dignos, tienen sentimientos y pueden sentir dolor, por lo que hay que considerar a los animales de manera igualitaria.

6.3.4 Cuarto dilema moral: Personas que cumplen la ley vs personas que no cumplen la ley

En el caso de que un peatón esté cruzando la calle con el semáforo en rojo y otro este cruzando la calle con el semáforo en verde y el coche autónomo deba colisionar contra alguno de ellos elegirá colisionar contra aquella persona que está cruzando en rojo ya que está incumpliendo la ley. El peatón que está cruzando en verde y está haciendo las cosas de forma legal no tiene por qué sufrir las consecuencias, en este caso que los dos peatones están en igualdad de condiciones la

solución parece simple pero este dilema cada vez se va haciendo más complicado cuando ya no empiezan a estar en igualdad de condiciones.

En el caso de que sean tres personas cruzando en rojo y solo una de ellas cruzando en verde, contra quién es preferible colisionar ¿Contra aquellas tres que están cruzando en rojo? Si se colisiona contra las tres personas estás haciendo daño a un número mayor de personas, pero si se colisiona contra aquella que está cruzando en verde no sería justo.

De la misma manera pasa con el caso de dos motoristas, uno de ellos lleva casco y las protecciones necesarias para en caso de sufrir un accidente tener los menores daños posibles, y otro de ellos no lleva casco y va sin protecciones. El que está incumpliendo la ley es aquel que no lleva casco, pero si se colisiona contra él por no estar cumpliendo las leyes sus probabilidades de sobrevivir se ven reducidas ¿Sería justo colisionar contra aquel motorista que cumple la ley porque sus probabilidades de sobrevivir son mayores?

Ante todos estos casos hay otras opciones a parte de una posible colisión, todos los coches autónomos podrían ser programados para que el conductor pueda coger los mandos y que sea él el que tome la decisión. Uno de los retos y dudas que se plantean ante esta posible solución es si daría tiempo a que, en caso de accidente, el conductor pueda retomar el control sobre el vehículo, y si en el caso de que la tomé, su respuesta fuese más rápida y eficiente que la que tomaría el coche autónomo.

Como hemos podido analizar son muchas las variables las que hay que tener en consideración a la hora de tomar una decisión u otra y los humanos al ser seres vivos que tienen emociones, sienten y padecen pueden sufrir bloqueos o dudar mucho a la hora de tomar decisiones ya que hay muchas variables en juego, por eso los coches autónomos tomarían la mejor posible decisión que tuviese menos impacto en un tiempo mucho menor ya que no intervienen las emociones ni el estado de ánimo.

El dilema ético que se nos plantea es que los coches autónomos deberán tener la misma moral que un humano, aunque la tecnología permita salvar más vidas, lo éticamente correcto es que haga lo mismo que haga un humano para que las máquinas no lleguen a superar a los humanos.

Desde los principios éticos una máquina no debería de superar la capacidad de una persona porque te metes en el ámbito de yo decido, yo no decido, aunque sea capaz de salvar más vida. La pregunta sería ¿Esto es verdaderamente ético?, todas las vidas humanas son iguales en dignidad, pero si tenemos la oportunidad de salvar vidas y de hacer de este mundo mejor ¿Por qué no programar a los coches autónomos para evitar el mayor número de siniestros en vez de programarlos para que actúen como lo haría un humano?

Estas decisiones morales nos plantean preguntas como, ¿Será necesario la construcción de coches autónomos con Inteligencia artificial que aprendan de las acciones de sus conductores habituales? De esta manera “el vehículo tomaría acciones morales” basadas en las mismas que tomaría su conductor y la moralidad del ser humano se estaría representando. Se podría cuestionar si los seres humanos toman decisiones morales acertadas y si esta sería una decisión acertada ya que la moralidad de las máquinas con Inteligencia Artificial sería superior al ser estas máquinas superiores a los seres humanos.

En el caso de que la reacción sea producida por parte del ser humano y haya un siniestro o un accidente es muy probable que no se considere delito porque no ha habido negligencia, mala intención o previsión de realizarlo. Pero en el caso de que esta reacción haya sido producida por un coche autónomo no se podría decir lo mismo, ya que la decisión ha sido deliberada, no ha actuado por instintos.

Después de analizar estas problemáticas, llegamos a la conclusión de que los vehículos autómatas siempre serán máquinas y se les pueden implementar algoritmos por los cuales sus acciones serán socialmente más positivas como por ejemplo ser más sostenible desde el punto de vista medioambiental, pero hay que ser conscientes de que hay límites intolerables. Las personas no son sujetos de utilidad ni de cálculo utilitario por lo que un algoritmo deberá llevar implementado aquellas directrices que ayuden a vivir a las personas, no aquellas a las que debe suprimir.

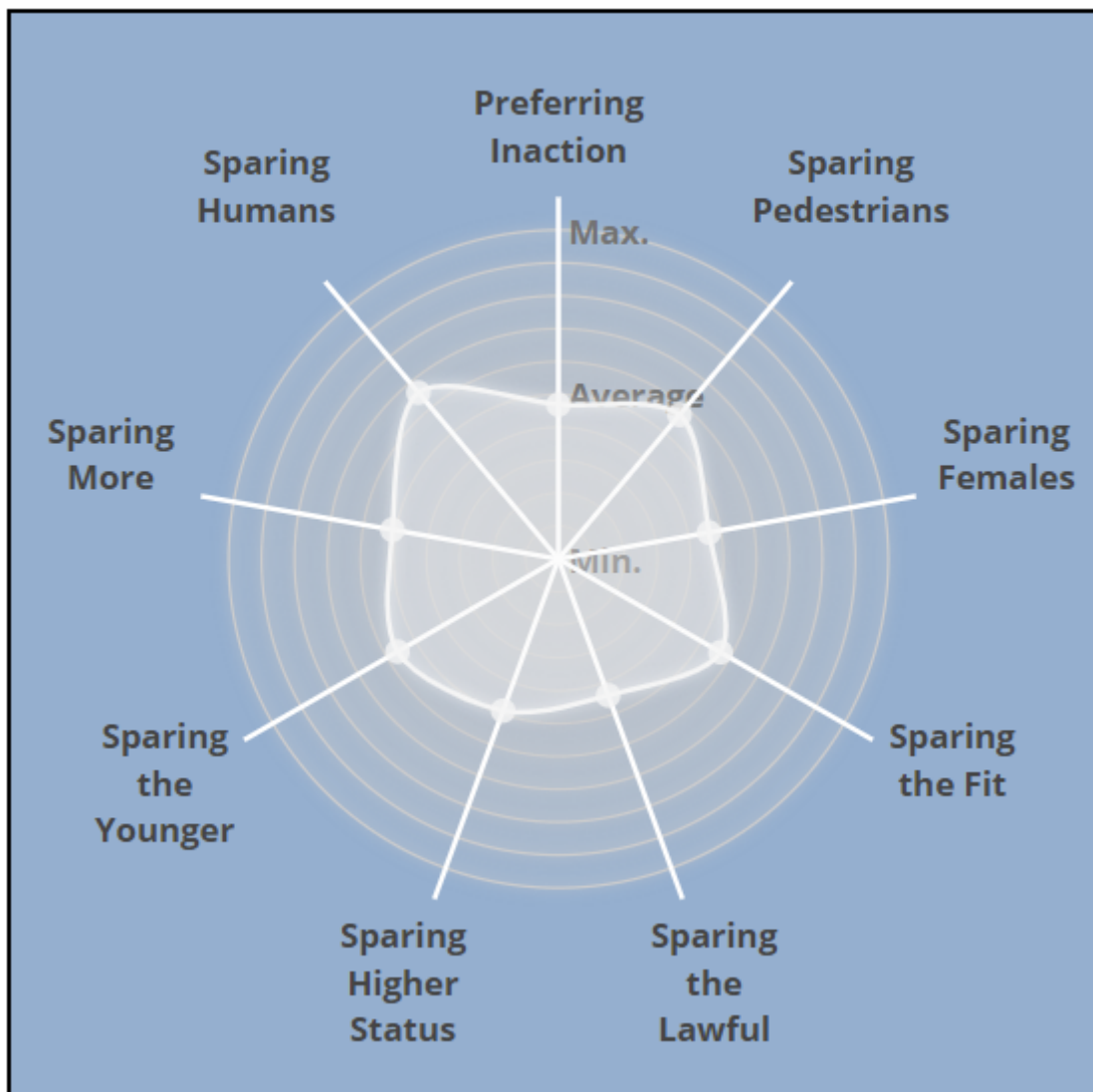
6.3.6 Moral Machine

El MIT (Massachusetts Institute of Technology) creó una plataforma online en la que todas las personas con acceso a Internet pueden jugar al juego en el que deben elegir “el mal menor de dos males”. Plantea dilemas éticos y morales para observar que preferencias tienen los seres humanos con respecto a la edad, especie, cumplimiento de la ley, sexo, poder adquisitivo...

Entre estos dilemas éticos el jugador tendrá que elegir entre atropellar a niños, personas mayores, animales, ladrones, personas de mayor poder adquisitivo como empresarios, médicos o vagabundos y ladrones. Además, también tendrán la opción de desviar el vehículo y matar a los pasajeros del coche en vez de a los seres que estén cruzando la calzada. Tendrán la opción de elegir de atropellar a personas que están cruzando la calzada de manera correcta o de manera incorrecta y de elegir entre atropellar a dos pasajeros o a cinco peatones.

Todo el estudio se realiza desde la perspectiva del vehículo, ya que se está intentando recopilar las perspectivas humanas sobre las decisiones morales que deberían tomar los coches autónomos. Han participado ya más de 4 millones de personas en este estudio y hasta la fecha es el atlas moral más desarrollado.

Imagen 2: Moral Machine Results



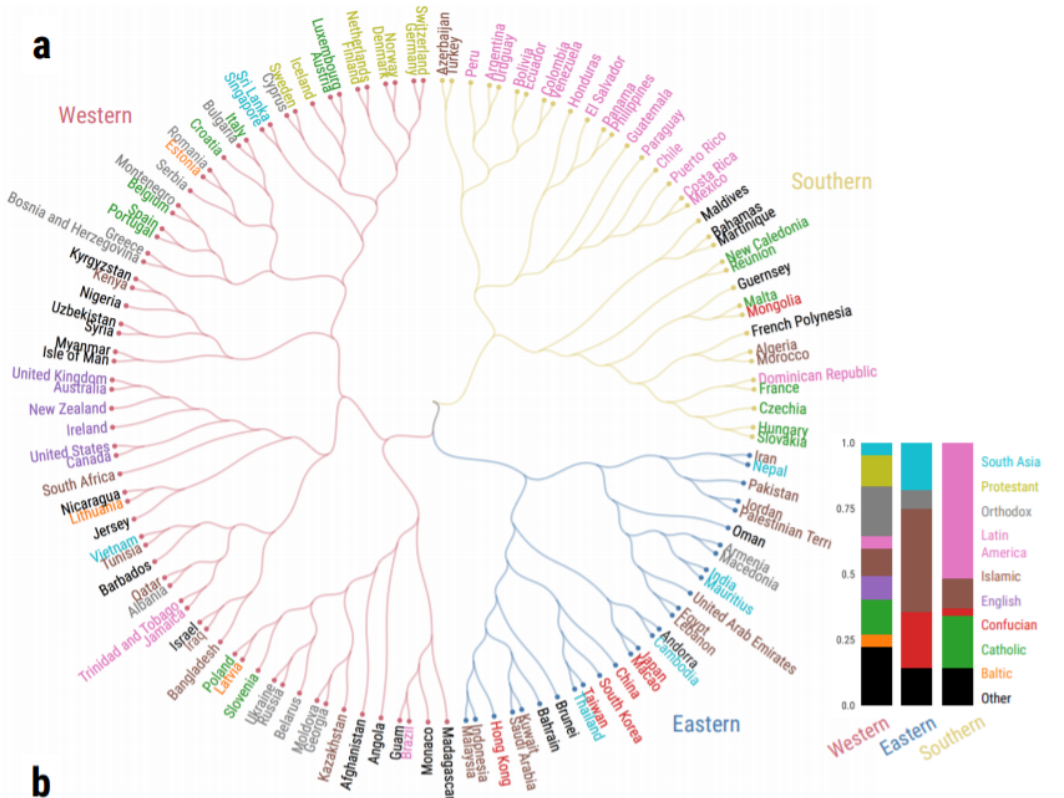
Fuente: Massachusetts Institute of Technology

En esta imagen se representa la media de las respuestas de todo el mundo, y hay decisiones más aceptadas de forma universal, estas son las de salvar a las personas antes que a los animales, salvar al mayor número de personas posibles y finalmente salvar preferentemente a la población más joven antes que a la población de edad más avanzada.

También se observa que de forma más universal hay una preferencia mayor por salvar a aquellas personas que tienen mayor poder adquisitivo a aquellas que son delincuentes o que no tienen hogar.

Aunque estas decisiones más aceptadas de forma universal no se pueden tomar como referencia para crear un marco teórico ético ya que hay rasgos discriminatorios y además entre distintas culturas estos parámetros cambian como podremos ver a continuación:

Imagen 3: The Moral Machine Experiment.



Fuente: University of Exeter

Estos gráficos creados por Awad, E; Dsouza, S; Kim, R; et al en 2019 son un análisis de los resultados del experimento Moral Machine. Este análisis dividía las respuestas en tres distintos grupos culturales. Estos grupos culturales surgen por la proximidad geográfica y por tener una cultura parecida, lo que llevó a que las respuestas sobre la ética que deben tener las máquinas fuesen muy parecidas, pero entre los tres grupos culturales hay mayores diferencias entre las respuestas, esto plantea problemas.

El primer grupo cultural es el Occidental (Western), en este grupo cultural se incluyen los siguientes países: Norteamérica y países pertenecientes a Europa que su orientación religiosa es católica, protestante y cristiana ortodoxa.

Tanto el grupo cultural Occidental como el oriental tienen una mayor preferencia por salvar a los humanos que el grupo cultural Meridional, en este se observa que hay una mayor preferencia por salvar a los animales.

El segundo grupo cultural es el Oriental (Eastern), en este grupo cultural se incluyen países del lejano oriente tanto confucionistas como islámicos: Japón, Taiwán, Arabia Saudí, Indonesia y Pakistán entre otros.

En el grupo cultural Oriental se observa una mayor preferencia en salvar a las personas de edad avanzada frente a los más jóvenes y de la misma manera ocurre con las personas de mayor estatus, tienen menor preferencia a salvarles frente al resto de segmentos de la sociedad.

El tercer grupo cultural es el Meridional (Southern), en este grupo cultural se incluyen países de América del Sur, Centroamérica y algunos países con influencia francesa.

En el grupo Meridional a diferencia del grupo Oriental hay mayor preferencia por salvar a las personas más jóvenes en vez de a las personas más mayores, de salvar más a las mujeres, a las personas que hagan deporte y también hay una mayor preferencia por salvar a las personas con mayor status que a aquellas que no tienen tanto status.

Entre todos los grupos culturales se encuentran algunas similitudes, todos ellos tienen preferencias por salvar a los peatones antes que a los propios pasajeros del coche (aunque esta preferencia es muy débil) y una preferencia mucho más fuerte es aquella en la que prefieren salvar a las personas que cumplen la ley antes que a aquellos que no lo hacen.

Este estudio del MIT trata de recopilar las perspectivas humanas sobre las decisiones que debería tomar un vehículo autónomo en caso de que se produzca un dilema moral para crear un marco teórico, pero estas decisiones como hemos podido comprobar son muy dispares entre distintos países y se utiliza a las personas como sujetos de utilidad, no se considera a cada ser humano digno en sí mismo, ya que separa a las personas en mejores y peores, buenas o malas, triunfadoras o perdedoras.

Esta clasificación está dividiendo a las personas en seres más dignos o menos dignos y cuanto más digno seas mayor probabilidad de vivir, por lo que se está deshumanizando. Una máquina inteligente no puede tener el poder de decisión de elegir quién es más digno ya que todos los seres vivos son sujetos de dignidad por igual y no se puede discriminar.

6.4 RESPONSABILIDAD

La IA en los coches autónomos cada vez está tomando más importancia en nuestras vidas, la toma de decisiones y acciones que hace por los seres humanos va en aumento ya que cada vez se le está otorgando mayor autonomía y por esa misma razón nos planteamos si debiéramos de otorgarle responsabilidad moral. La responsabilidad moral es la toma de decisiones de manera consciente y asumiendo la repercusión que tendrán los actos.

Cuando los coches autónomos en un nivel de automatización superior al nivel 3 empiezan a tomar decisiones que antes tomaban los humanos y hay un accidente, no sabes a quién atribuir a la responsabilidad, ¿Al dueño del vehículo, a la empresa que ha creado al vehículo, al diseñador de la Inteligencia artificial, al que ha sufrido el accidente, el que ha regulado la ley?

Los seres humanos a la hora de tomar decisiones actuamos bajo una responsabilidad moral ya que sabemos que nuestros actos tienen consecuencias y sabemos discernir entre el bien y el mal. Es el ser humano el que decide delegar la capacidad de actuación al coche autónomo (en el caso de tener un coche con un nivel de automatización superior al nivel 3). Pero podríamos decir que las decisiones que los coches autónomos toman por las Inteligencias artificiales ¿son verdaderamente

decisiones? Cuando el ser humano delega las decisiones a un coche autónomo no hay que olvidarse de que es una responsabilidad moral.

Una de las condiciones de la responsabilidad moral es la condición de control, si tienes la capacidad de decir y de tener efectos sobre el entorno que te rodea, debes ser responsable de tus actos, por lo que en este caso los coches autónomos sí que deberían de tener responsabilidad moral. Pero también debes ser consciente de lo que estás haciendo y sus posibles consecuencias (Koekelberg, 2021), a raíz de esta segunda condición nos damos cuenta de que los coches autónomos no tienen conciencia y por lo tanto no son conscientes de lo que hacen ni de las repercusiones que puede llegar a tener sus acciones.

Para que exista responsabilidad moral debe existir plena conciencia, plena y libre voluntad y una intencionalidad positiva, que como ha sido explicado anteriormente los vehículos autónomos al no cumplir estas características carecen de responsabilidad moral.

Como hemos mencionado, los coches autónomos con IA no se les puede considerar agentes morales ya que carecen de conciencia para actuar libremente y pudiendo reflexionar sobre sus acciones. La responsabilidad moral va ligada a un agente moral y como los coches autónomos no son agentes morales no pueden hacerse responsables de sus acciones, por lo que son los seres humanos los responsables de estas acciones.

Pero en el caso de los coches autónomos esta afirmación se complica ya que gracias a la IA que llevan implementada, el tiempo de reacción que tienen los seres humanos ante un imprevisto es muy pequeño. Por lo que ¿Cómo podrían ser los seres humanos responsables de las acciones que tomen los coches autónomos si no tienen tiempo para intervenir en la decisión final? La responsabilidad reside en el algoritmo que ha sido programado.

Como hemos dicho anteriormente, los responsables morales de la IA son los seres humanos, pero en el caso de que se produzca un accidente y se vea envuelto un coche autónomo no sabríamos decir quien de todos los seres humanos involucrados es el responsable, ¿El dueño del coche, el fabricante, el programador, otro coche...? No sabríamos decir quién es el responsable y posiblemente no haya un solo responsable.

Otro de los aspectos que dificultan la asignación de la responsabilidad de la IA es el aprendizaje automático que tiene está, en el momento en el que la IA empieza a aprender de forma automática, la asignación de la responsabilidad se complica ya que se pierde la transparencia de como se ha llegado a tomar una decisión.

Que la explicación de cómo las máquinas llegan a tomar decisiones supone otros problemas y dilemas y nos hace plantearnos hasta qué punto queremos depender de las máquinas. Los primeros diseñadores de la IA saben cuál es el fin y que quieren conseguir, pero a medida que más gente se va viendo involucrada como otros programadores, diseñadores o incluso el consumidor final, el entendimiento de la toma de decisiones que tiene la IA se va reduciendo hasta poder llegar a ser nulo.

6.5 REFLEXIÓN ÉTICA

Tras analizar las problemáticas mencionadas en los apartados anteriores nos damos cuenta de que las implicaciones que tienen los vehículos autónomos no solo se quedan en los vehículos sino que también afectan a la sociedad y la manera que tenemos de relacionarnos dentro de ella.

Al dotarle de un sentido ético a las máquinas inteligentes como los coches autónomos tenemos que preguntarnos ¿Por qué tipo de ser humano estamos apostando? ¿Qué tipo de ser humano queremos ser y qué papel queremos jugar en la sociedad? Estamos apostando por un ser humano que va a dejar la toma de decisiones en manos de una máquina, desapareciendo por completo del centro de decisiones o seguimos apostando por la centralidad en el ser humano. No hay que olvidar que lo tecnológicamente posible puede que no sea lo humanamente deseable.

Buena parte de estas problemáticas analizadas van a afectar al modelo de sociedad, al modelo de persona, al modelo de sostenibilidad deseable, al modelo económico. respondiendo a una nueva forma de vida con una cultura distinta.

Si apostamos por apartar al ser humano de la toma de decisiones hay que ser conscientes de que el modelo de sociedad que conocemos cambiará, llegando a ser más cerrado, exclusivo y menos plural, ya que las tomas de decisión harán que los humanos dejen de ser sujetos de dignidad y serán sujetos de utilidad. La igualdad se perdería y las

brechas que hoy en día nos separan lo harán incluso más, porque se crearán estándares de perfección y las personas empezarán a valer más o menos en función de sus logros, pertenencias y capacidad.

Los seres humanos no serán dignos en sí mismos. Un vagabundo que vive en la calle, sin ingresos ni estudios, con algún problema físico o psicológico será menos válido que un ejecutivo con unos ingresos por encima de la media, felizmente casado y en perfecto estado de salud.

La reflexión ética necesita de plena consciencia, libertad e intencionalidad positiva, las tecnologías deben estar al servicio del ser humano, para que el ser humano mejore sus prestaciones, y servicios en la sociedad pero no para quitarle la capacidad de elegir. De esta nueva revolución queremos que el ser humano siga siendo el centro y no quede relevado, no hay que perder la centralidad en el ser humano.

Las personas no son sujetos de utilidad, son sujetos de dignidad por lo que no hay personas mejores o peores, o personas más dignas o menos dignas ni tan siquiera en función de la calidad moral ni de los aciertos que han tenido en la vida, por esa razón debería estar prohibido que un algoritmo decida quién debe morir y quien debe sobrevivir, pero eso no quita que haya algoritmos implementados que si favorezcan la vida humana.

Los coches autónomos traerán numerosos beneficios a la sociedad, como reducir la tasa de accidentes ya que podrá actuar de una manera más rápida y eficaz para prevenir un accidente que un ser humano al poder interactuar con el entorno, serán grandes propulsores de la sostenibilidad, ya que entre ellos se reducirá: el tráfico, los gases de efecto invernadero, se podrán dar nuevas utilidades a los espacios reservados para aparcar ya que el modelo de uso de los vehículos cambiará. Pero todas estas mejoras están enfocadas al ser humano, él sigue siendo el centro por ello no se puede programar un vehículo con un algoritmo que decida quién debe vivir y quién no.

Para que el futuro de los coches autónomos sea exitoso, este debe de tener la confianza de la sociedad. Para ello es importante que haya comprensión de su modelaje, funcionamiento, trata de datos... Tienen que sentirse seguros y para ello es importante que haya entendimiento del funcionamiento.

La disposición de la gente a comprar vehículos autónomos, también dependerá del marco ético que estos tengan, cuanto más se parezca al código ético que ellos tengan más disposición a comprar. Aunque las preferencias éticas de una sociedad no tengan que ser el determinante para crear una política ética no se puede dejar de lado ya que es un factor importante.

Además, para que tengan éxito también será muy importante dotar de infraestructuras adecuadas, mejoras en las estructuras y calles que puede llegar a ser una gran inversión que no todos los países podrían soportar o no se podría implementar en todas las zonas de un mismo país.

Que los coches autónomos lleguen a ser una realidad en nuestro día a día requiere de estudios de nuevos escenarios y de investigaciones, ya que todavía se encuentra en la fase de embrión, es necesario un gran progreso. Una de las conclusiones a las que llegamos es que quizás no sea necesario que haya un código ético universal pero sí que sería necesario que todos los usuarios sepan explicar el por qué de las acciones que se han llevado a cabo, que sepan explicar el razonamiento de por qué se ha tomado esa decisión y no otra y para ello es muy necesaria la transparencia.

Aunque los vehículos autónomos hayan sido diseñados para reducir los accidentes y ser más efectivos que los humanos, esto requiere de que la mayoría sean autónomos también y puedan comunicarse entre ellos. Es importante aclarar que aunque puedan reducir la tasa de mortalidad no hay que olvidar que son máquinas y pueden fallar de forma ocasional.

Quedan muchas dudas por resolver y muchas de ellas no tienen una postura pública sobre cómo van a ser abordadas, todas estas cuestiones ralentizan la implementación de los vehículos autónomos en la sociedad y la autorización de marcos legales ya que no hay regulación para que puedan circular.

7. CONCLUSIONES

Por último indicar que los objetivos planteados en este trabajo se ven cumplidos:

- Se ha comprendido que el desarrollo de la inteligencia artificial, ya presentes en nuestras vidas, ha supuesto una nueva revolución, cambiando la forma que se tenía de entender el funcionamiento de la sociedad. Este cambio tecnológico en el sector de la automoción ha cambiado la manera que teníamos de ver el transporte y de la influencia que tendremos sobre él.
- Se han analizado las problemáticas éticas que supone la integración de la inteligencia artificial en un coche autónomo y se ha realizado una reflexión ética sobre la repercusión que tendrán en la sociedad, viendo en qué aspectos supondrá una mejora y en cuales supondrá un agravamiento. Puede traer muy buenas oportunidades si se adapta de manera correcta, pero puede suponer nuevos problemas graves si no se implementa apropiadamente.

Con la investigación de este trabajo se quiere dar a entender que la automatización ya ha llegado aunque no seamos conscientes del todo. En esta cuarta revolución los cambios se producen a gran velocidad, por ello es importante actuar con rapidez en la regularización de estos nuevos vehículos. Mediante el análisis expuesto en este documento se pretende ayudar a entender la automatización y todo lo que conlleva detrás. Es una industria muy incipiente por lo que el trabajo no se ha abordado con mucho detalle ya que no cuenta con restricciones reglamentarias y todavía no han sido autorizados, pero puede seguirse profundizando a medida que vayan siendo más conocidas y más aceptadas, ya que esta revolución es inminente y traerá beneficios para todos los colectivos de la sociedad.

8. BIBLIOGRAFÍA:

1. REAL ACADEMIA ESPAÑOLA: *Diccionario de la lengua española*, 23.^a ed., [versión 23.4 en línea] <https://dle.rae.es/inteligencia#2DxmhCT> recuperado el 1 de marzo de 2021
2. Andreas, M., & Kaplan, A. (2019, 17 julio). *A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence*. Berkely Haas.
https://journals.sagepub.com/doi/full/10.1177/0008125619864925?casa_token=T1Wr0uWye50AAAAA%3ArPiNapsXKBaE7AWjiz9umWA8zN04wDN238nRzHiCPUrS683cAzWfYmJvW9j8GpLajcwVCE0ozjg
3. Capgemini. (2019). *Accelerating automotive's AI transformation: How driving AI enterprise-wide can turbo-charge organizational value*. Capgemini. Recuperado el 20 de Julio de 2020 en <https://www.capgemini.com/es-es/wp-content/uploads/sites/16/2019/03/Ai-in-automotive-research-report.pdf>
4. López de Mántaras Badia, R. (2015, diciembre). *Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial*. Novática, No 234. <https://digital.csic.es/bitstream/10261/136978/1/NOV234%282015%2997-101.pdf>
5. Ortiz, A (2013). *Coche autónomo y debates éticos*. Xataka. Recuperado el 09 de septiembre de 2020 de <https://www.xataka.com/automovil/coche-autonomo-ydebates-eticos>
6. García, M. S. (2019). *Ética e inteligencia artificial*. Cátedra CaixaBank de Responsabilidad Social Corporativa. Recuperado el 20 de Julio de 2020 de <https://www.iese.edu/wp-content/uploads/2019/12/ST-0522.pdf>

7. Koekelberg, M. (2021). *Ética de la inteligencia artificial*. Madrid, España, Cátedra.
8. Huawei. (s. f.). *Inteligencia Artificial Fuerte vs. Inteligencia artificial débil* [Fotografía]. https://img.blogs.es/iahuawei/wp-content/uploads/2018/10/IA_infografia.png
9. Brynjolfsson, E. & McAfee, A (2014). *The Second Machine Age*. Nueva York, W.W. Norton.
10. Modelo determinista. (2020, 15 noviembre). En *Wikipedia, la enciclopedia libre*. https://es.wikipedia.org/wiki/Modelo_determinista
11. Oliver, N. (2020). *Inteligencia artificial, naturalmente*. ONTSI. En <https://www.ontsi.red.es/sites/ontsi/files/2020-06/InteligenciaArtificialNuriaOliver.pdf>
12. Valero-Matas, J. A., & De La Barrera, A. (2019, octubre). *The Autonomous Car: A better future?* Ediciones Universidad Valladolid. Recuperado el 14 de Julio de 2020 <http://revistas.uva.es/index.php/sociotecnologia/article/view/4225/3280>
<https://doi.org/10.24197/st.1.2020.136-158>
13. *Código ético del IEEE*. (s. f.). IEEE. Recuperado 1 de febrero de 2021, de http://www.ieec.uned.es/investigacion/ieee_dieec/co_etico_ieee.htm
14. Maurer, M., Lenz, B., Gerdes, J. C., & Winner, H. (2015). *Autonomous Driving*. Springer Open. <https://doi.org/10.1007/978-3-662-48847-8>

15. UNESCO. (2020, 2 octubre). *Inteligencia artificial: ejemplos de dilemas éticos*. <https://es.unesco.org/artificial-intelligence/ethics/cases#autonomouscar>
16. Nieto, N. H. (2019, 23 diciembre). *Claves 2030: Los dilemas éticos de la inteligencia artificial*. KPMG Tendencias. <https://www.tendencias.kpmg.es/claves-decada-2020-2030/robots-etica-inteligencia-artificial/>
17. Castañeda, S., & Peñalosa, E. (2016). *Fenomenología de Agencia Académica*. Universidad Autónoma Metropolitana, 153. http://devcua.cua.uam.mx/pdfs/revistas_electronicas/libros-electronicos/2017/agencia_academica/agencia_electronico.pdf#page=153
18. Digital Europe. (2019, 13 noviembre). *DIGITALEUROPE's Recommendations on Artificial Intelligence Policy*. <https://www.digitaleurope.org/resources/digitaleuropes-recommendations-on-artificial-intelligence-policy/>
19. Cortina, A. (2019). *Ética de la inteligencia artificial*. Ministerio de Ciencia, Innovación y Universidades. Valencia. En <https://www.racmyp.es/docs/anales/a96-24.pdf>
[acceso 4 de noviembre de 2020]
20. Seoane, A.M. (2018). Computational Thinking between Philosophy and STEM. Programming Decision Making applied to the Behaviour of “Moral Machines” in Ethical Values Classroom. IEEE- RITA. En <https://ieeexplore.ieee.org/document/8305482>
[acceso el 14 de Julio de 2020]
21. Warrendale, P. A. (2018, 12 diciembre). *SAE International Releases Updated Visual Chart for Its “Levels of Driving Automation” Standard for Self-Driving Vehicles*. SAE INTERNATIONAL.

<https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%E2%80%9Clevels-of-driving-automation%E2%80%9D-standard-for-self-driving-vehicles>

22. SAE International. (2018, junio). *Taxonomy and Definitions for Terms Related to Driving Automation System for On-Road Motor Vehicles*. SAE International. https://doi.org/10.4271/J3016_201806
23. Galván, P. (s. f.). *Los 6 Niveles de Autonomía*. SG Buzz. Recuperado 19 de enero de 2021, de <https://sg.com.mx/revista/55/los-6-niveles-autonom>
24. Li, Y., Díaz, M., Morantes, S., & Dorati, Y. (2018, junio). *Vista de Vehículos autónomos: Innovación en la logística urbana* (Vol. 4, N°1). Universidad Tecnológica de Panamá. Lic Gestión de la Producción Industrial. <https://doi.org/10.33412/rev-ric.v4.1.1864>
25. Awad, E; Dsouza, S; Kim, R; et al. (2019, octubre). *The Moral Machine Experiment*. Nature. University of Exeter. <https://doi.org/10.1038/s41586-018-0637-6>
26. Awad, E. (2017, junio). *Moral machines: perception of moral judgment made by machines*. Massachusetts Institute of Technology. <http://hdl.handle.net/1721.1/112532>
27. López de Mántaras, R. (2019). *El futuro de la IA: hacia inteligencias artificiales realmente inteligentes*. OpenMind. <https://www.bbvaopenmind.com/articulos/el-futuro-de-la-ia-hacia-inteligencias-artificiales-realmente-inteligentes/>
28. Comisión Europea. (2020, febrero). *LIBRO BLANCO sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza* (COM

(2020) 65 final). ES.

https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf

29. LaFrance, A. (2016, 23 junio). *How Many Lives Will Driverless Cars Save?* The Atlantic.

<https://www.theatlantic.com/technology/archive/2015/09/self-driving-cars-could-save-300000-lives-per-decade-in-america/407956/>

30. Rouhiainen, L. (2018). *INTELIGENCIA ARTIFICIAL 101 COSAS QUE DEBES SABER HOY SOBRE NUESTRO FUTURO*. Planeta.

https://static0planetadelibroscom.cdnstatics.com/libros_contenido_extra/40/39308_Inteligencia_artificial.pdf

31. Caballero, J. F. (2006). La Teoría de la Justicia de John Rawls. *IBERO FORUM, II* (Año 1), 1–22.

https://d1wqtxts1xzle7.cloudfront.net/38371720/Rawls_comentado.pdf?1438632634=&response-content-disposition=inline%3B+filename%3DLa_Teoria_de_la_Justicia_de_John_Rawls.pdf&Expires=1618325583&Signature=RmY~I1kiukiUbzRNhEVwRU~jN9C6gVVYeN3f2h0OVqVTfGubwhsqruE1v~bZC4keTazpSKIVtaBSjBrcGcB1Q2MSVjv1AChVzreqCwRYjUH1EZJZlv6359pegbcYqPQqyhWeyKjh~aaB76I1SODaYDnY54a2X0zeMsQ3050NwKNqJ~UmPgKWQ3txiQCOT9wK5jRt1vI0K-VO9tKE9aZpB7xcEArn2sNWUYgzuANWyJyBmORxJYiqVD57hQWPqnO25eA8gGiQWfhzdZ25lqpSC1I~TFvQFMqVuepJTs0rddyx35gupi5tbhh1uu80hL37D~U~2i7EWmDT4vx6AsCfLJO&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

32. Perasso, V. (2016, 12 octubre). *Qué es la cuarta revolución industrial (y por qué debería preocuparnos)*. BBC Mundo.

https://docs.ufpr.br/~jrgarcia/macroeconomia_ecologica/macroeconomia_ecologica/Qu%C3%A9%20es%20la%20cuarta%20revoluci%C3%B3n%20industrial.pdf

33. Cárdenas, R., & Padilla, V. (2019, diciembre). *VEHÍCULOS AUTÓNOMOS: TEMAS VIALES – DILEMAS VITALES.LA TECNOLOGÍA TOMA EL VOLANTE* (Vol. N°8). UNIFE. <https://doi.org/10.33539/peryla.2019.n8.1956>
34. <http://riaa.uaem.mx/xmlui/bitstream/handle/20.500.12055/734/FIBVHC07T.pdf?sequence=1&isAllowed=y> DIRECTRICES PARA RESOLVER DILEMAS ÉTICOS EN ADS
35. Aurik, J. (2017, 30 enero). *La Cuarta Revolución Industrial tendrá un efecto disruptivo sobre el empleo, ¿pero cómo?* World Economic Forum. <https://es.weforum.org/agenda/2017/01/la-cuarta-revolucion-industrial-tendra-un-efecto-disruptivo-sobre-el-empleo-pero-como/>
36. <https://publications.iadb.org/es/publications/spanish/document/Revista-Integraci%C3%B3n--Comercio-A%C3%B1o-21-No-42-Agosto-2017-Robot-luci%C3%B3n-el-futuro-del-trabajo-en-la-integraci%C3%B3n-40-de-Am%C3%A9rica-Latina.pdf>
37. Carbajal Suárez, Y. (Ed.). (2010, enero). *Sector automotriz: reestructuración tecnológica y reconfiguración del mercado mundial* (Año 2 Núm. 1). Paradigma económico. <https://dialnet.unirioja.es/descarga/articulo/5961679.pdf>
38. Evans, D. (2011, abril). *Internet de las cosas Cómo la próxima evolución de Internet lo cambia todo*. Cisco Internet Business Solutions Group (IBSG). <http://audentia-gestion.fr/cisco/IoT/internet-of-things-iot-ibsg.pdf>

39. Monasterio Astobiza, Á. (2019). Ética para máquinas: Similitudes y diferencias entre la moral artificial y la moral humana*. *DILEMATA*, N°30, 129–147. <https://dialnet.unirioja.es/servlet/articulo?codigo=7416201>
40. Cortina Orts, A. (2019). *ÉTICA DE LA INTELIGENCIA ARTIFICIAL*. Ministerio de Ciencia, Innovación y Universidades. https://www.boe.es/biblioteca_juridica/anuarios_derecho/abrir_pdf.php?id=ANU-M-2019-10037900394

Imágenes:

1. SAE. (s. f.). *LEVELS OF DRIVING AUTOMATION* [Gráfico]. <https://www.sae.org/binaries/content/gallery/cm/articles/news/2018/12/j3016-levels-of-driving-automation-12-10.jpg>
2. Massachusetts Institute of Technology. (s. f.). *Moral Machine Results*. Recuperado 1 de abril de 2021, de <http://moralmachineresults.scalablecoop.org/>
3. Awad, E; Dsouza, S; Kim, R; et al. (2019, octubre). *The Moral Machine Experiment*. Nature. University of Exeter. <https://doi.org/10.1038/s41586-018-0637-6>