



Máster Universitario en Ingeniería Industrial

TRABAJO FIN DE MÁSTER

Creación de un marco ético para su implementación en
vehículos autónomos.

Autor: Javier Castillejo Huerta

Director: Juan Norberto Moriñigo

Madrid

Agosto de 2025

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título
Creación de un marco ético para su implementación en vehículos autónomos
en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el
curso académico 2024/25 es de mi autoría, original e inédito y
no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido
tomada de otros documentos está debidamente referenciada.



Fdo.: Javier Castillejo Huerta

Fecha: ..20../ ..08../ 2025

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo: Juan Norberto Moriñigo

Fecha://



Máster Universitario en Ingeniería Industrial

TRABAJO FIN DE MÁSTER

Creación de un marco ético para su implementación en
vehículos autónomos.

Autor: Javier Castillejo Huerta

Director: Juan Norberto Moriñigo

Madrid

Agosto de 2025

CREACION DE UN MARCO ETICO PARA SU IMPLEMENTACION EN VEHICULOS AUTONOMOS

Autor: Castillejo Huerta, Javier

Director: Norberto Morfiño, Juan.

Entidad Colaboradora: ICAI – Universidad Pontificia Comillas

RESUMEN DEL PROYECTO

En este proyecto se estudia el actual estado del vehículo autónomo con respecto a la ética a la que se ve expuesto. Se analiza el impacto positivo que tendrá su implementación y se crea unas pautas a la hora de crear el marco ético a nivel arquitectura de software, marco legal, y escogiendo la ética utilitarista como marco moral.

Palabras clave: Ética, Redes Neuronales, Vehículo Autónomo

1. Introducción

El vehículo autónomo se define como aquel vehículo capaz de tener una circulación completa, eficaz y segura sin implicación manual del ser humano [2]. A pesar de ser un concepto lejano se empiezan a ver en la carretera niveles menos sofisticados de automatización y es necesario comenzar a pensar en la regulación sobre estos vehículos.

Uno de los puntos más controversiales es la ética que regirá a estos vehículos, que además gracias a dilemas clásicos como el del tranvía e implementaciones novedosas como los robo taxis de tela ha hecho que se convierta en uno de los temas del momento. Este trabajo plantea estudiar dos cosas. Primero si es necesario el coche autónomo desde un punto de vista moral y segundo y eje principal; como debe ser el marco ético que rijan este tipo de vehículos.

2. Definición del proyecto

Para llevar a cabo el Proyecto se han hecho tres investigaciones principales. Primero una revisión del estado del arte para comprender cual es el estado actual de esta tecnología. Principalmente enfocada en los problemas a los que se enfrentan los fabricantes y cuáles son las herramientas novedosas que están utilizando para atacar los problemas éticos.

El segundo bloque es un modelo de implementación del vehículo autónomo, realizado con el propósito de demostrar si la introducción del vehículo autónomo es moralmente aceptable. Se darán respuesta a niveles de implementación, aumento o disminución de la seguridad vial o como evolucionarán las emisiones y eficiencia del sector del transporte terrestre gracias a un vehículo completamente autónomo.

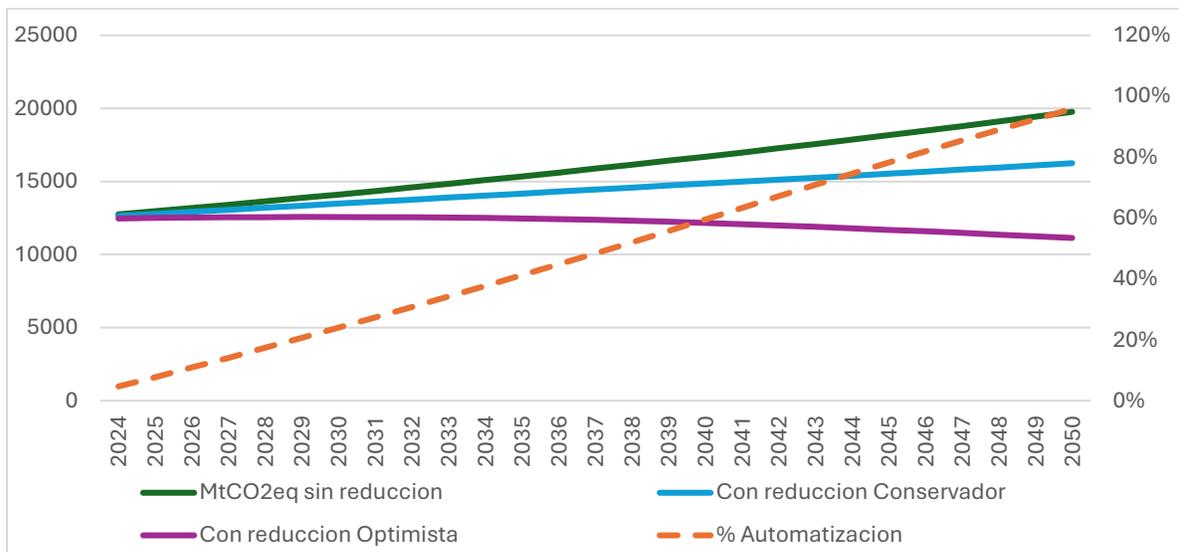
Finalmente se confeccionará un marco ético, en el que se definirán tres puntos principales que los fabricantes deberán tener en cuenta. Primero la arquitectura de software y como deberían ser los algoritmos que rijan su comportamiento. Segundo, marco ético por utilizar, analizando los tres existentes en la ética clásica y encontrando el más adecuado para el vehículo autónomo. Finalmente, que rol deberá tener el regulador en cuanto a supervisor general de estos vehículos.

3. Descripción del modelo/sistema/herramienta

El modelo de penetración del coche autónomo se basa en tres conceptos principales. El primero es el actual estado del parque de vehículos y como este evolucionara en los próximos 25%, calculando el número de vehículos con automatización, así como el nivel promedio de automatización. El segundo concepto es la seguridad que se derivara de este aumento. Siguiendo datos de otros estudios se ha computado la reducción en países en desarrollo y en países desarrollados teniendo en cuenta conceptos calculados en el primer apartado como nivel promedio para adaptar la reducción. Finalmente, un estudio de reducción de tiempos de conducción y emisiones, donde se han identificado las principales acciones que cambiaran la eficiencia de estos vehículos para intentar calcular la reducción de emisiones. Cabe recalcar que no se ha pronosticado el incremento de eficiencia de las unidades motoras, como por ejemplo la transición a vehículos híbridos o eléctricos.

Para este modelo se ha empleado Excel como herramienta principal para el calculo de los diferentes parámetros.

Gráfico de la reducción de emisiones pronosticada



Por otro lado, el marco ético también se estructura en tres pilares. Primero la arquitectura de software que se refiere al tipo de software que regirá las acciones del vehículo autónomo. Se han estudiado modelos de reglas preprogramadas, basados en reglas neuronales y modelos mixtos donde los dos anteriores son capaces de interactuar entre ellos. El pilar más distintivo es el segundo, en el cual se evalúan los tres modelos de la ética clásica, el utilitarismo, la ética del deber y la ética de la virtud. Se analizan los pros y contras, pero no desde un punto filosófico, sino desde un punto ingenieril a nivel de implementación en un sistema basado en software y más cuantitativo que cualitativo. Finalmente, todo el modelo ha de estar regulado

para asegurar una correcta implementación. En la sección de regulación diferentes conceptos como culpabilidad, diferenciación entre regiones y peritaje de los sistemas será estudiada.

4. Resultados

Los resultados son claros. Primero en el modelo de introducción se observa como las ventajas que se le atribuyen al vehículo autónomo son claramente palpables, reduciéndose los accidentes en 2050 más de un 30% y las emisiones entre el 15% y 45% siendo pesimistas.

Por otro lado, en el modelo ético se llega a la conclusión de que el software debería basarse en un sistema mixto con reglas preprogramadas como límites y con una matriz de valores ética que aporte prejuicio o penalizaciones a ciertas decisiones de la red neuronal que se comportara como el conductor. En cuanto a la ética a utilizar deberá ser una variación de la ética utilitarista propuesta por Karl Popper [1] en el que se minimiza en vez de maximizar. Finalmente, a nivel legislativo se regulan las responsabilidades de cada agente y se marcan los límites que cada agente deberá tener sobre el funcionamiento y diseño del vehículo autónomo.

5. Conclusiones

El trabajo demuestra la importancia del vehículo autónomo y como una implementación rápida y responsable debería ser una prioridad para órganos reguladores. Se aporta una visión clara sencilla y explicable al problema de la ética del vehículo autónomo, dejando una solución que pueda ser realísticamente adoptada y la población sea capaz de confiar en ella. Aun así, se proponen preguntas a estudiar en próximos trabajos, como si con una automatización completa los humanos deberíamos tener prohibido la conducción fuera de lugares de recreo.

6. Referencias

[1] Popper, K. (1945). *The Open Society and Its Enemies*

[2] MIT. (s.f.). *Moral Machine*. Obtenido de MIT: <https://www.moralmachine.net/>

CREATION OF AN ETHICAL FRAMEWORK FOR ITS IMPLEMENTATION IN AUTONOMOUS VEHICLES

Author: Castillejo Huerta, Javier

Director: Norberto Moríñigo, Juan.

Entity: ICAI – Universidad Pontificia Comillas

ABSTRACT

In this project, the current state of autonomous vehicles regarding the ethical challenges they face is studied. The positive impact of their implementation is analyzed, and guidelines are established for creating the ethical framework at the levels of software architecture and legal framework, choosing utilitarian ethics as the moral framework.

Key Words: Ethics, Neural Networks, Autonomous Vehicle

7. Introduction

An autonomous vehicle is defined as a vehicle capable of operating completely, efficiently, and safely without manual human involvement [2]. Although it is still considered a distant concept, less sophisticated levels of automation are already appearing on the roads, making it necessary to start thinking about regulations for these vehicles.

One of the most controversial points is the ethics that will govern these vehicles. Classic dilemmas, such as the trolley problem, and innovative implementations like robo-taxis, have turned this into a trending topic. This work aims to study two things: first, whether the autonomous car is necessary from a moral point of view, and second, and as the main focus, how the ethical framework governing these types of vehicles should be designed.

8. Project definition

To carry out the project, three main areas of research have been conducted. First, a review of the state of the art was performed to understand the status of this technology, focusing mainly on the challenges faced by manufacturers and the innovative tools they are using to address ethical issues.

The second section is an implementation model for autonomous vehicles, created with the purpose of demonstrating whether the introduction of autonomous vehicles is morally acceptable. This section will address questions related to levels of implementation, the potential increase or decrease in road safety, and how emissions and the efficiency of the land transport sector might evolve thanks to fully autonomous vehicles.

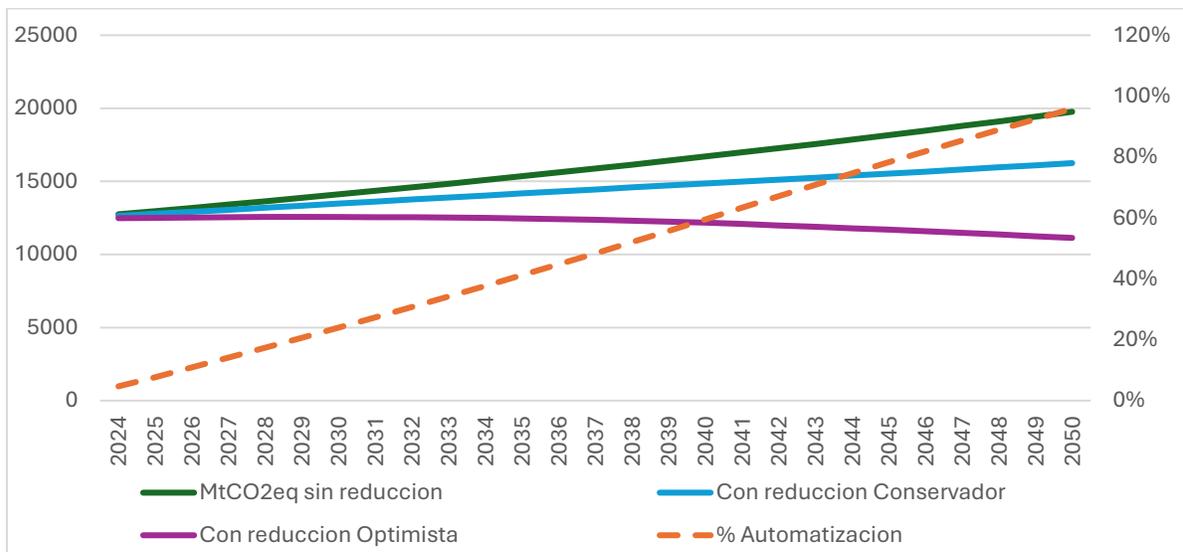
Finally, an ethical framework will be developed, defining three main points that manufacturers should consider. First, the software architecture and how the algorithms governing vehicle behavior should be designed. Second, the ethical framework to be used, analyzing the three existing approaches in classical ethics and identifying the most suitable one for autonomous vehicles. Lastly, the role of the regulator as the general supervisor of these vehicles will be defined.

9. Model/Tools description

The penetration model for autonomous vehicles is based on three main concepts. The first is the current state of the vehicle fleet and how it will evolve over the next 25 years, calculating the number of vehicles with automation as well as the average level of automation. The second concept is the safety that will result from this increase. Using data from other studies, the reduction in accidents has been estimated for both developing and developed countries, considering factors calculated in the first section, such as the average level of automation, to adapt the reduction accordingly. Finally, there is a study on the reduction of driving times and emissions, where the main actions that will change the efficiency of these vehicles have been identified in order to estimate the reduction in emissions. It is important to note that the increased efficiency of power units, such as the transition to hybrid or electric vehicles, has not been forecasted.

Excel has been used as the main tool for calculating the different parameters in this model.

Forecasted emissions reduction



On the other hand, the ethical framework is also structured around three pillars. The first is software architecture, which refers to the type of software that will govern the actions of the autonomous vehicle. Models based on pre-programmed rules, neural rule-based systems, and mixed models, where both approaches can interact, have been studied.

The most distinctive pillar is the second, in which the three models of classical ethics are evaluated: utilitarianism, deontological ethics, and virtue ethics. Their pros and cons are analyzed not from a philosophical perspective, but from an engineering standpoint, focusing on implementation in a software-based system and using a more quantitative than qualitative approach.

Finally, the entire model must be regulated to ensure correct implementation. In the regulation section, various concepts such as liability, regional differentiation, and system assessment will be studied.

10. Results

The results are clear. First, in the introduction model, it is observed that the advantages attributed to autonomous vehicles are indeed tangible, with accidents projected to decrease by over 30% by 2050 and emissions reduced by between 15% and 45%, even under pessimistic scenarios.

On the other hand, the ethical model concludes that the software should be based on a mixed system, with pre-programmed rules serving as boundaries and an ethical value matrix providing biases or penalties for certain decisions made by the neural network, which will function as the driver. Regarding the ethical approach to be used, it should be a variation of the utilitarian ethics proposed by Karl Popper [1], in which the goal is to minimize rather than maximize harm.

Finally, at the legislative level, the responsibilities of each agent are regulated, and the limits that each agent must have regarding the operation and design of the autonomous vehicle are defined.

11. Conclusions

It's demonstrated the importance of autonomous vehicles and how their rapid and responsible implementation should be a priority for regulatory bodies. It provides a clear, simple, and explainable perspective on the ethical issues surrounding autonomous vehicles, offering a solution that can be realistically adopted and trusted by the public. Nevertheless, questions remain for future research, such as whether, with complete automation, humans should be prohibited from driving outside recreational areas.

12. References

- [1] Popper, K. (1945). *The Open Society and Its Enemies*
- [2] MIT. (s.f.). Moral Machine. From MIT: <https://www.moralmachine.net/>

Índice

1.	Introducción	6
2.	Estado de la Cuestión.....	8
2.1.	Introducción	8
2.2.	Historia del coche autónomo	9
2.2.1.	Décadas de 1920-1950: Los primeros conceptos.....	9
2.2.2.	Décadas de 1980-1990: Los pioneros académicos	10
2.2.3.	Años 2000: El impulso de DARPA.....	11
2.2.4.	2010-2020: La carrera tecnológica	11
2.3.	Contexto actual del desarrollo de los vehículos autónomos	12
2.4.	Niveles de automatización de los vehículos autónomos.....	15
2.4.1.	Nivel 0: Sin automatización.....	16
2.4.2.	Nivel 1: Asistencia al conductor	17
2.4.3.	Nivel 2: Automatización parcial	18
2.4.4.	Nivel 3: Automatización condicional.....	19
2.4.5.	Nivel 4: Alta automatización	20
Nivel 5:	Automatización total	21
2.5.	Hardware en vehículos autónomos	23
2.5.1.	Sensores para la percepción del entorno	23
2.5.2.	Sistemas de posicionamiento y navegación.....	25
2.5.3.	Actuadores para el control del vehículo.....	27
2.5.4.	Unidades de procesamiento	27
2.6.	Tecnologías clave de software en vehículos autónomos.....	28
2.7.	Regulación actual de los vehículos autónomos	31

2.7.1.	Regulación a nivel global.....	31
2.7.2.	Regulación en Europa.....	32
2.7.3.	Regulación en España.....	33
2.8.	Ética en vehículos autónomos.....	33
3.	Motivación.....	38
4.	Objetivos del Proyecto.....	40
5.	Alineación con los Objetivos de Desarrollo Sostenible (ODS).....	42
6.	Modelo sobre los beneficios del coche autónomo.....	44
6.1.	Modelo de beneficios, ¿por qué?.....	44
6.2.	Análisis de mercado.....	45
6.3.	Siniestralidad Vial.....	48
6.4.	Eficiencia y emisiones.....	53
7.	Modelo Ético.....	61
7.1.	Introducción.....	61
7.2.	Arquitectura de software.....	61
7.3.	Marco ético.....	65
7.3.1.	Especificaciones de la ética utilitarista.....	67
7.4.	Marco legislativo.....	69
7.5.	Solución y cuestiones existentes:.....	70
8.	Conclusiones.....	72
8.1.	Conclusiones.....	72
8.2.	Modelo de los beneficios del coche autónomo.....	72
8.3.	Modelo ético.....	73
8.4.	Puntos abiertos.....	73

9. Bibliografía 75

Índice de imágenes

Ilustración 1: Vehículo de Waymu (The End of driving: Autonomous Cars, s.d.)	9
Ilustración 2: Vehículo de Francis Houdina, American Wonder 1925	10
Ilustración 3: Volkswagen Passat modificado para el Darpa Urban de 2007. Tangi Quemener, Getty Images	11
Ilustración 4: Robotaxi de Tesla; Tim Goessman, Bloomberg News	12
Ilustración 5: Noticia de la publicación de una nueva versión de NVIDIA DRIVE (Washabaugh, 2021)	14
Ilustración 6: Diferentes niveles de automatización (Siddahantjain, 2023)	16
Ilustración 7: Sistema del Audi A8 traffic jam (Green Car Congress , 2017).....	20
Ilustración 8 Cruise Origin (Miller, 2025)	22
Ilustración 9: Funcionamiento del Lidar (Badoni, 2021).....	24
Ilustración 10: Vista real del sistema InnovizOne (Badoni, 2021)	24
Ilustración 11: INS de un misil balístico francés, Museo del Aire y del Espacio, Paris	26
Ilustración 12: Herramienta interactiva de Tensorflow para divulgación sobre redes neuronales (TensorFlow, s.d.).....	30
Ilustración 13: MIT Moral Machine (MIT, s.d.)	34
Ilustración 14: Listado de los Objetivos de Desarrollo Sostenible (Naciones Unidas , s.d.).....	43
Ilustración 15: Gráfico de la penetración de vehículos autónomos de nivel 3 y superior. (Goldman Sachs, 2024).....	47
Ilustración 16: Representación del efecto acordeón (Etrasa, 2019).....	58

Índice de Tablas

Tabla 1: Pronostico de la evolución de vehículos en circulación	46
Tabla 2: Pronostico de la penetracion y nivel de automatizacion del parque de vehiculos global 2025-2050	48
Tabla 3: Distribución de los accidentes graves entre países desarrollados y en vías de desarrollo por tipo de vehículo	50
Tabla 4: Resultados del analisis de reduccion de accidentes	51
Tabla 5: Porcentaje de reducción de accidentes por año.....	53
Tabla 6 Tabla con las diferentes reducciones de emisiones por mejora en la conducción autónoma (Massar, 2021)	55
Tabla 7: Tabla con las reducciones de emisiones utilizadas en el estudio	57
Tabla 8: Gráfico de la reducción de emisiones pronosticada.....	58

1. Introducción

La evolución de la conducción autónoma ha transformado radicalmente el concepto de movilidad, planteando no solo desafíos técnicos sino también dilemas éticos que requieren un marco normativo claro para su implementación segura y transparente. A medida que los sistemas de inteligencia artificial asumen el control de la toma de decisiones en carretera, surge la necesidad de definir principios que guíen su comportamiento ante situaciones de riesgo, tal como lo han explorado diversas iniciativas académicas como la Moral Machine del MIT (MIT, s.d.). La variabilidad en las preferencias éticas entre sociedades hace que la estandarización de estos principios sea un reto complejo, similar al debate sobre la ética de la IA en otros ámbitos como la medicina, la seguridad o la justicia.

En este contexto, resulta interesante revisar las Tres Leyes de la Robótica formuladas por Isaac Asimov, diseñadas originalmente para garantizar la seguridad de los humanos frente a la autonomía de los robots: (1) un robot no hará daño a un ser humano ni permitirá, por inacción, que sufra daño; (2) debe obedecer las órdenes dadas por los humanos, salvo que entren en conflicto con la primera ley; y (3) debe proteger su propia existencia en la medida en que esto no entre en conflicto con las dos anteriores. Posteriormente, se añadió la Ley Cero, que prioriza la seguridad de la humanidad en su conjunto sobre la de un individuo en particular. Si bien estas reglas nacieron en un contexto de ciencia ficción, su lógica resulta aplicable a la conducción autónoma, donde la toma de decisiones debe equilibrar la protección de los ocupantes, peatones y otros vehículos. Sin embargo, a diferencia de un robot con un único objetivo programado, los vehículos autónomos operan en un entorno complejo donde la ética debe coexistir con restricciones técnicas y legales.

Para abordar estos dilemas, se han explorado enfoques filosóficos como el utilitarismo, que prioriza minimizar el daño total, o la deontología, que enfatiza el cumplimiento de normas fijas sin considerar las consecuencias individuales. Sin embargo, trasladar estos principios a un sistema automatizado presenta desafíos, especialmente por la dificultad de programar criterios morales universales en un contexto dinámico como la conducción.

Otro aspecto clave es el sesgo algorítmico, ya que la IA aprende de datos históricos que pueden contener prejuicios, afectando potencialmente la toma de decisiones del vehículo. Para mitigar este riesgo, se investiga en IA explicable (XAI), con el objetivo de hacer transparentes y auditables las decisiones del sistema.

Además, los vehículos autónomos enfrentan retos legales y normativos, especialmente en cuanto a la asignación de responsabilidad en caso de accidente. Las regulaciones aún están en desarrollo, con iniciativas como el Reglamento de IA de la UE, que busca establecer marcos de transparencia y seguridad. En este sentido, el principio de precaución sugiere la necesidad de regulaciones rigurosas antes de su implementación masiva para minimizar riesgos.

Uno de los aspectos más relevantes desde el punto de vista ingenieril es que la ética en la conducción autónoma representa un desafío fundamentalmente distinto a los problemas habitualmente abordados en las disciplinas técnicas. Mientras que en ingeniería se tiende a trabajar con problemas de naturaleza binaria: resueltos o no resueltos, funcionales o defectuosos; la toma de decisiones éticas en vehículos autónomos escapa a esta lógica, ya que no existe una única solución correcta e indiscutible. Al igual que en cuestiones globales como la erradicación del hambre o la lucha contra el narcotráfico, cualquier enfoque propuesto en este ámbito es susceptible de generar contradicciones y controversias. No obstante, el objetivo debe ser desarrollar un marco de decisión que, dentro de la complejidad intrínseca del problema, garantice el mayor grado posible de transparencia, equidad y seguridad.

Este trabajo busca desarrollar un marco ético basado en una matriz de toma de decisiones, considerando las implicaciones legales y normativas para generar confianza en esta tecnología emergente y facilitar su integración en la sociedad. La necesidad de un marco ético robusto se vuelve aún más crítica en un contexto donde la inteligencia artificial no solo debe reaccionar a situaciones imprevistas, sino también alinearse con valores humanos que no siempre son universalmente compartidos. Por ello, es fundamental analizar modelos previos de ética en IA, como los aplicados en decisiones médicas automatizadas o en sistemas de seguridad, para evaluar cómo estos enfoques pueden adaptarse a la conducción autónoma y garantizar una convivencia armónica entre máquinas y humanos en las carreteras del futuro.

2. Estado de la Cuestión

2.1. Introducción

Un coche autónomo, también denominado vehículo autónomo, self-driving car o driverless car, es un automóvil diseñado para realizar todas las tareas de conducción sin intervención humana, integrando un conjunto avanzado de tecnologías que le permiten percibir el entorno, procesar información en tiempo real y tomar decisiones de manera autónoma. Este tipo de vehículo se basa en una combinación de hardware y software, incluyendo sensores (como LIDAR, radares y cámaras), sistemas de navegación global (GPS y GNSS), actuadores para el control del vehículo (dirección, aceleración y frenado) y algoritmos de inteligencia artificial (IA) que procesan grandes volúmenes de datos para interpretar el entorno y ejecutar maniobras seguras. La Sociedad de Ingenieros Automotrices (SAE, s.d.) establece una clasificación estándar en seis niveles de autonomía, desde el nivel 0 (sin asistencia alguna) hasta el nivel 5 (autonomía total, donde el vehículo puede operar en cualquier escenario sin conductor humano). Un coche autónomo de nivel 5 no requiere volante ni pedales, ya que es capaz de gestionar situaciones complejas, como tráfico urbano denso o condiciones climáticas adversas, de manera completamente independiente. Además, los vehículos autónomos podrían incorporar sistemas de comunicación avanzados, como vehículo-a-vehículo (V2V) y vehículo-a-infraestructura (V2I), que optimizan la coordinación en carretera y mejoran la eficiencia del tráfico.



Ilustración 1: Vehículo de Waymu (The End of driving: Autonomous Cars, s.d.)

El concepto de coche autónomo va más allá de la mera automatización de la conducción: busca revolucionar la movilidad, incrementando la seguridad vial (al eliminar errores humanos, responsables de la mayoría de los accidentes), reduciendo la congestión, optimizando el consumo energético y facilitando el acceso a la movilidad para personas con limitaciones físicas. Sin embargo, su desarrollo plantea dilemas éticos, técnicos y legales que requieren un análisis profundo, especialmente en lo relativo a la toma de decisiones en situaciones críticas y la responsabilidad en caso de fallos.

2.2. Historia del coche autónomo

La idea de un vehículo capaz de conducirse solo ha fascinado a ingenieros y visionarios durante décadas, con raíces que se remontan a experimentos e innovaciones tecnológicas del siglo XX. Aunque los coches autónomos parecen un fenómeno reciente, su historia combina avances en ingeniería mecánica, electrónica, informática y, más recientemente, inteligencia artificial (Bimbrow, 2015).

2.2.1. Décadas de 1920-1950: Los primeros conceptos

Los orígenes del coche autónomo pueden rastrearse hasta experimentos rudimentarios de automatización. En 1925, el inventor estadounidense Francis Houdina presentó el American

Wonder, un coche controlado por radio que desfiló por las calles de Nueva York, aunque dependía de un vehículo cercano para su control remoto, por lo que no era autónomo en el sentido moderno. Durante los años 50, General Motors (GM) imaginó un futuro con carreteras automatizadas en su exposición Futurama, donde los coches se guiaban por cables eléctricos incrustados en la vía, una idea más cercana a la infraestructura inteligente que a la autonomía del vehículo en sí.

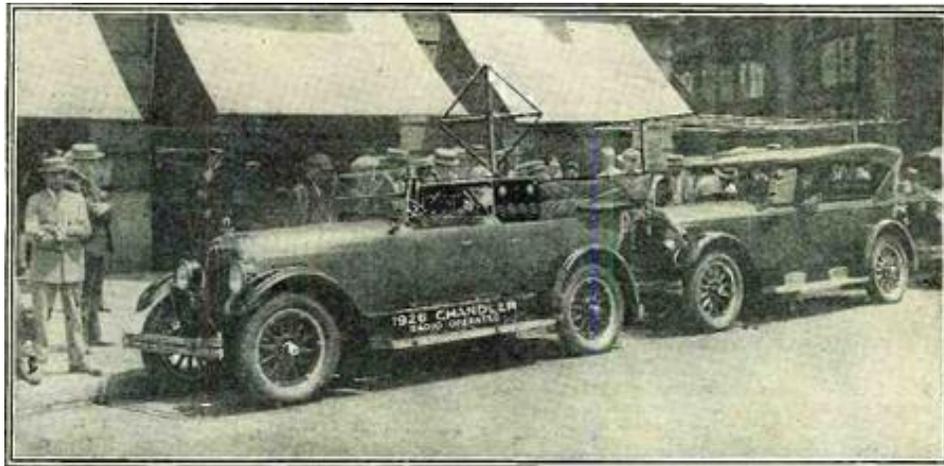


Ilustración 2: Vehículo de Francis Houdina, American Wonder 1925

2.2.2. Décadas de 1980-1990: Los pioneros académicos

El verdadero impulso hacia los coches autónomos llegó con los avances en robótica y computación. En los años 80, el proyecto Navlab de la Universidad Carnegie Mellon (EE. UU.) desarrolló los primeros prototipos de vehículos autónomos, como el Navlab 1 (1986), un todoterreno equipado con cámaras y computadoras que podía seguir caminos simples a baja velocidad. En Europa, el programa PROMETHEUS (1987-1994), financiado por la Unión Europea, marcó un hito al involucrar a fabricantes como Mercedes-Benz y BMW en el desarrollo de tecnologías de asistencia a la conducción, como el control de crucero adaptativo. En 1995, el equipo de Carnegie Mellon logró que un vehículo recorriera 4.500 km a través de EE. UU. de manera semiautónoma, un avance que demostró el potencial de los sistemas de visión por ordenador.

2.2.3. Años 2000: El impulso de DARPA

El punto de inflexión llegó con los DARPA Grand Challenges (2004-2007), competiciones organizadas por la Agencia de Proyectos de Investigación Avanzados de Defensa de EE. UU. En 2004, ningún vehículo completó el recorrido desértico propuesto, pero en 2005, el equipo de Stanford, liderado por Sebastian Thrun, ganó con Stanley, un coche que combinaba LIDAR, GPS y algoritmos de aprendizaje. En 2007, el DARPA Urban Challenge exigió a los vehículos navegar en entornos urbanos simulados, respetando señales y otros vehículos. Estos eventos catalizaron la inversión en autonomía y atrajeron a gigantes tecnológicos y automotrices al campo.



Ilustración 3: Volkswagen Passat modificado para el Darpa Urban de 2007. Tangi Quemener, Getty Images

2.2.4. 2010-2020: La carrera tecnológica

En 2009, Google (ahora Alphabet) lanzó su Self-Driving Car Project (rebautizado como Waymo), que marcó el inicio de una carrera por la autonomía total. Los prototipos de Google, basados en coches Toyota modificados, acumularon millones de kilómetros en pruebas reales, perfeccionando el uso de LIDAR y mapas digitales. Paralelamente, Tesla irrumpió en el sector con su sistema Autopilot (2014), que ofrecía autonomía parcial (nivel 2) y prometía mejoras continuas mediante actualizaciones de software. Empresas como Uber, Lyft, General Motors (Cruise) y startups como Zoox también invirtieron fuertemente, mientras que fabricantes tradicionales como Volkswagen y Toyota desarrollaron sus propias plataformas.

Durante esta década, los avances en inteligencia artificial, especialmente en deep learning, permitieron a los vehículos interpretar entornos complejos con mayor precisión. Sin embargo, incidentes como el accidente fatal de un Tesla en 2016 y el atropello de un peatón por un vehículo de Uber en 2018 pusieron en el punto de mira los retos éticos y técnicos, resaltando la necesidad de regulaciones claras.

2.3. Contexto actual del desarrollo de los vehículos autónomos

El desarrollo de los vehículos autónomos ha evolucionado significativamente desde los experimentos iniciales, alcanzando un punto en el que los avances tecnológicos están comenzando a traducirse en aplicaciones prácticas, aunque con retos importantes que aún limitan su adopción generalizada (Anderson, 2016). Este contexto actual refleja el legado de décadas de innovación, desde los primeros prototipos de los años 80 hasta las competiciones como el DARPA Grand Challenge.



Ilustración 4: Robotaxi de Tesla; Tim Goessman, Bloomberg News

En el ámbito comercial, empresas como Waymo y Cruise están liderando la implementación de robotaxis, vehículos capaces de operar sin conductor humano en entornos específicos. Waymo ha desplegado servicios de transporte autónomo en ciudades como Phoenix y San Francisco, utilizando una combinación de sensores avanzados y algoritmos para navegar en áreas urbanas mapeadas. Cruise, por su parte, ha expandido sus operaciones en San Francisco, ofreciendo servicios de transporte sin supervisión humana en condiciones controladas. En China, Baidu ha consolidado su plataforma Apollo, implementando flotas de robotaxis en ciudades como Pekín y Shanghái, donde los vehículos se adaptan a entornos dinámicos gracias a la integración de tecnologías de percepción y planificación. Estas iniciativas muestran cómo los avances históricos, como los prototipos del Navlab o los logros del DARPA, han evolucionado hacia aplicaciones reales, aunque limitadas a entornos predefinidos que aún no abarcan la complejidad total del mundo vial.

Paralelamente, los fabricantes tradicionales y las empresas tecnológicas están explorando aplicaciones en niveles intermedios de independencia. Tesla, por ejemplo, ha implementado su sistema Autopilot, que permite al vehículo realizar maniobras como cambios de carril o ajustes de velocidad en autopistas, aunque requiere supervisión humana constante. Este sistema se basa en datos recopilados de su flota global para mejorar su rendimiento, un enfoque que refleja la influencia de los avances en inteligencia artificial que comenzaron a consolidarse en la década de 2010 (Burns, 2013). Empresas como NVIDIA, con su plataforma DRIVE, están desarrollando soluciones que integran hardware y software para permitir una conducción más autónoma, mientras que Motional, una colaboración entre Hyundai y Aptiv, está probando vehículos en entornos urbanos, buscando un equilibrio entre seguridad y adaptabilidad. Estas iniciativas demuestran cómo la carrera tecnológica iniciada por proyectos como el de Google ha dado lugar a una competencia global entre diversos actores.

NVIDIA DRIVE OS 5.2.6 Linux SDK Now Available

Aug 25, 2021

+1 Like Discuss (0)

By [Katie Washabaugh](#)



Ilustración 5: Noticia de la publicación de una nueva versión de NVIDIA DRIVE (Washabaugh, 2021)

El contexto actual también incluye aplicaciones más allá de los turismos, un reflejo de la versatilidad de la tecnología autónoma. Empresas como TuSimple y Daimler están desarrollando camiones autónomos para rutas de larga distancia, aprovechando los avances en sensores y conectividad para optimizar la logística. En el sector de la entrega, Nuro ha implementado vehículos de reparto autónomos para servicios de último kilómetro, mientras que Navya ha desplegado autobuses autónomos en campus y áreas urbanas controladas, utilizando sistemas de comunicación para coordinarse con la infraestructura vial. Estas aplicaciones, que tienen sus raíces en los experimentos de automatización de las décadas pasadas, muestran cómo la tecnología está comenzando a transformar diversos sectores, aunque aún enfrenta limitaciones significativas.

A pesar de los progresos, los retos técnicos y regulatorios son una barrera importante para una autonomía más amplia. La fiabilidad en escenarios impredecibles, como obras viales o condiciones climáticas adversas, sigue siendo un desafío, ya que los sistemas actuales dependen de entornos bien mapeados y estructurados. La ciberseguridad es otra preocupación creciente: la dependencia de conectividad y actualizaciones en la nube introduce vulnerabilidades, ya que un ataque informático podría comprometer la seguridad del vehículo.

El impacto social y económico también está moldeando el panorama actual. La promesa de reducir los accidentes causados por errores humanos, que representan más del 90% de los incidentes viales, ha impulsado la inversión en esta tecnología, pero también plantea preguntas sobre el futuro del empleo, especialmente para conductores profesionales, y sobre la aceptación pública, que se ha visto afectada por incidentes como el accidente fatal de un vehículo de Uber en 2018. Estos aspectos, junto con los retos técnicos y regulatorios, serán explorados en mayor detalle en los siguientes apartados, donde se analizarán los niveles de autonomía que definen las capacidades actuales y futuras de los vehículos autónomos.

2.4. Niveles de automatización de los vehículos autónomos

Una vez explorado el estado actual del coche autónomo, es importante recalcar las diferencias entre los diferentes tipos o niveles de automatización. La clasificación de los niveles de automatización, definida por la Sociedad de Ingenieros Automotrices (SAE, s.d.) en el estándar J3016, establece un marco global para categorizar los vehículos según su grado de autonomía. Esta taxonomía, que abarca desde el nivel 0 hasta el nivel 5, se basa en la distribución de responsabilidades entre el conductor humano y el sistema de conducción automatizada (*Automated Driving System, ADS*). Cada nivel refleja un avance en la capacidad tecnológica del vehículo, desde sistemas de asistencia básica hasta la autonomía completa, con implicaciones significativas en el diseño del hardware, el software y la infraestructura requerida.

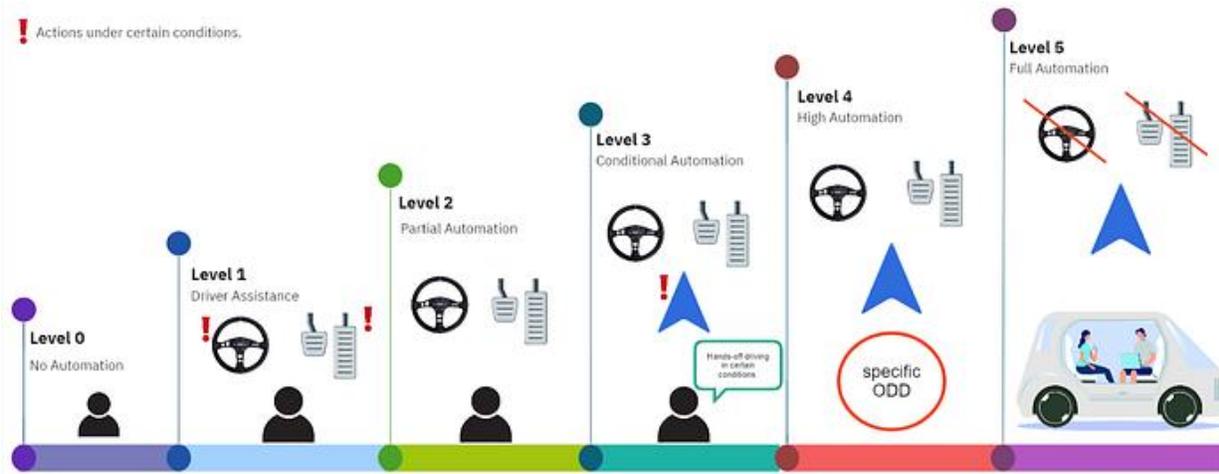


Ilustración 6: Diferentes niveles de automatización (Siddahantjain, 2023)

2.4.1. Nivel 0: Sin automatización

En el nivel 0, el vehículo carece de funciones de conducción automatizada. Todas las tareas dinámicas de conducción (*Dynamic Driving Tasks, DDT*), incluyendo la dirección, la aceleración, el frenado y la monitorización del entorno, son ejecutadas exclusivamente por el conductor humano. Los sistemas presentes, si los hay, se limitan a advertencias pasivas sin intervención en el control.

- **Tecnologías asociadas:** Incluyen sensores básicos como ultrasonidos para alertas de proximidad en maniobras de aparcamiento, cámaras traseras para asistencia visual y sistemas de advertencia simples, como detección de fatiga mediante análisis de movimientos del conductor.
- **Ejemplos:** Vehículos tradicionales sin asistencias avanzadas, como modelos antiguos o utilitarios de bajo coste.
- **Limitaciones:** La seguridad depende completamente de la capacidad y atención del conductor, lo que incrementa la probabilidad de errores humanos, principales responsables de los accidentes de tráfico según estadísticas internacionales.

- **Estado actual:** Este nivel está en retroceso en mercados desarrollados, pero persiste en vehículos económicos o en regiones con menor penetración tecnológica, donde la simplicidad mecánica prevalece sobre la integración de sistemas avanzados.

2.4.2. Nivel 1: Asistencia al conductor

El nivel 1 introduce sistemas de asistencia que automatizan una única tarea de conducción, ya sea el control longitudinal (aceleración y frenado) o el lateral (dirección), pero nunca ambas de forma simultánea. El conductor mantiene la responsabilidad total de la DDT, monitorizando el entorno y estando preparado para intervenir en todo momento.

- **Tecnologías asociadas:**
 - Control de crucero estándar, que mantiene una velocidad constante ajustada manualmente.
 - Control de crucero adaptativo (*Adaptive Cruise Control, ACC*), basado en radares o sensores láser para ajustar la velocidad según la distancia al vehículo precedente.
 - Asistente de mantenimiento de carril (*Lane Keeping Assist, LKA*), que utiliza cámaras para detectar líneas de carril y aplica correcciones suaves a la dirección.
 - Frenado automático de emergencia (*Autonomous Emergency Braking, AEB*), que activa los frenos ante obstáculos detectados por sensores si el conductor no responde.
- **Ejemplos:** Modelos como el Toyota Corolla con ACC o el Ford Focus con LKA.
- **Limitaciones:** Los sistemas son reactivos, sin capacidad predictiva para escenarios complejos. La dependencia del conductor introduce riesgos de distracción o confianza excesiva en las asistencias, lo que puede comprometer la seguridad.
- **Estado actual:** El nivel 1 está ampliamente implementado en vehículos de gama media y alta, gracias al bajo coste y la estandarización de sus tecnologías. Normativas como EuroNCAP han impulsado la adopción de sistemas como AEB en mercados europeos.

2.4.3. Nivel 2: Automatización parcial

En el nivel 2, el vehículo puede ejecutar simultáneamente el control longitudinal y lateral en condiciones específicas, asumiendo una porción limitada de la DDT. No obstante, el conductor debe supervisar constantemente el sistema y el entorno, estando preparado para retomar el control inmediato. Este nivel requiere que el conductor mantenga las manos en el volante (*hands-on*).

- **Tecnologías asociadas:**
 - Cámaras de alta resolución con algoritmos de visión por computadora para detectar carriles, señales y obstáculos.
 - Radares y, en algunos casos, LIDAR para medir distancias y velocidades con precisión.
 - Sistemas de fusión sensorial que integran datos de múltiples sensores para crear un modelo del entorno en tiempo real.
 - Actuadores electrónicos (*steer-by-wire* y *throttle-by-wire*) para ejecutar comandos precisos en dirección y aceleración.
 - Funciones combinadas como *Traffic Jam Assist*, que gestiona la conducción en atascos a baja velocidad.
- **Ejemplos:** Tesla Autopilot en modo básico, Cadillac Super Cruise, Nissan ProPilot Assist.
- **Limitaciones:** Los sistemas no pueden manejar situaciones fuera de su dominio de diseño operativo (*Operational Design Domain, ODD*), como intersecciones complejas o condiciones climáticas adversas. La supervisión humana constante puede generar fatiga de monitorización, y la confianza excesiva en el sistema ha sido vinculada a incidentes de seguridad.
- **Estado actual:** El nivel 2 es el estándar en vehículos premium y se está extendiendo a modelos de gama media. Representa un punto intermedio que combina avances tecnológicos con la necesidad de intervención humana, lo que plantea retos ergonómicos.

2.4.4. Nivel 3: Automatización condicional

El nivel 3 permite al vehículo ejecutar la DDT completa dentro de un ODD específico, como autopistas con tráfico predecible, incluyendo la monitorización del entorno y la respuesta a eventos esperados. El conductor no necesita supervisar de forma continua, pero debe estar preparado para retomar el control cuando el sistema lo solicite (*fallback*), con un tiempo de transición típico de 8 a 10 segundos.

- **Tecnologías asociadas:**
 - Mapas de alta definición (*HD Maps*) que proporcionan información detallada sobre carriles, semáforos y límites de velocidad.
 - Sensores redundantes (LIDAR, radar, cámaras) para garantizar fiabilidad en diversas condiciones.
 - Algoritmos de inteligencia artificial basados en *machine learning* para predecir trayectorias de vehículos y peatones, y gestionar transiciones seguras al conductor.
 - Interfaces hombre-máquina (*Human-Machine Interface, HMI*) con alertas visuales, auditivas y hápticas para notificar la necesidad de intervención.
- **Ejemplos:** Audi A8 Traffic Jam Pilot (limitado por regulaciones), Honda Legend con aprobación para nivel 3 en Japón.
- **Limitaciones:** La transición del control al conductor plantea riesgos si este está distraído, como durante actividades secundarias (lectura, uso de dispositivos). Definir un ODD claro y garantizar que el sistema reconozca sus límites requiere avances técnicos y regulatorios significativos (Merat, 2016).

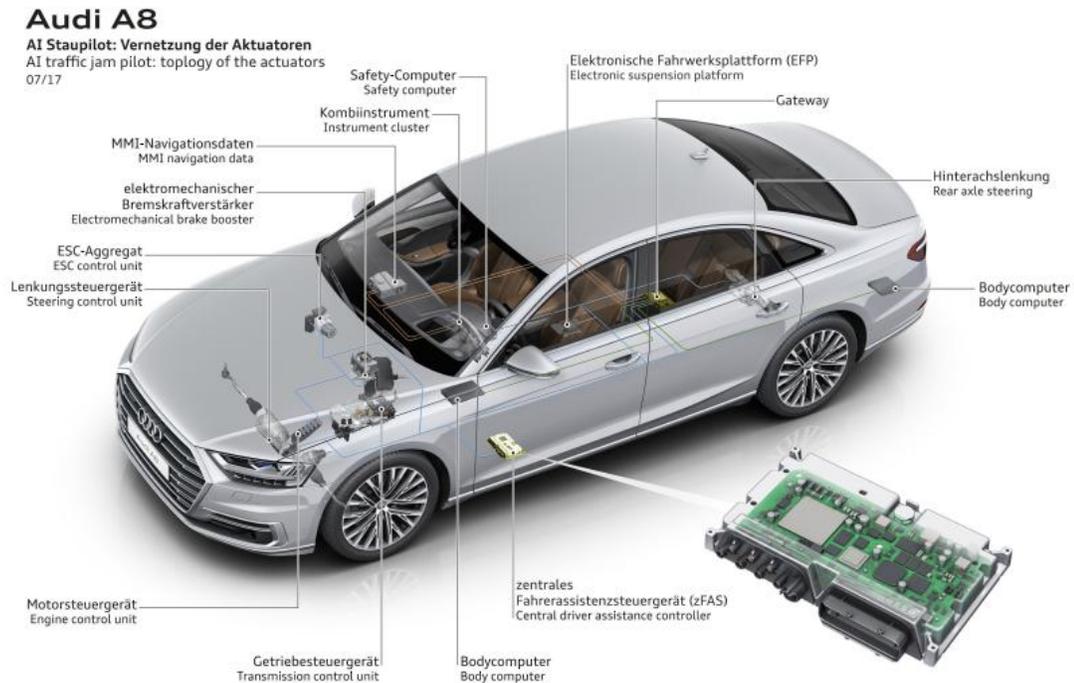


Ilustración 7: Sistema del Audi A8 traffic jam (Green Car Congress , 2017)

- **Estado actual:** El nivel 3 está en fase inicial de implementación, con aprobaciones limitadas en países como Japón y Alemania. La falta de armonización regulatoria global y los riesgos asociados a la supervisión humana restringida frenan su adopción generalizada.

2.4.5. Nivel 4: Alta automatización

En el nivel 4, el vehículo ejecuta la DDT completa dentro de un ODD definido sin intervención humana, gestionando incluso los fallos del sistema (*fallback* asumido por el ADS). Fuera del ODD, el vehículo puede alcanzar una condición de riesgo mínimo (*Minimal Risk Condition, MRC*), como detenerse de forma segura. No requiere supervisión humana en los escenarios para los que está diseñado.

- **Tecnologías asociadas:**
 - Sistemas redundantes de sensores, procesadores y actuadores para garantizar operatividad ante fallos.

- Conectividad vehículo-a-vehículo (*V2V*) y vehículo-a-infraestructura (*V2I*) mediante redes 5G para coordinar maniobras y recibir datos en tiempo real.
- Modelos de *deep learning* para anticipar comportamientos complejos en entornos dinámicos, como áreas urbanas.
- Protocolos de ciberseguridad para proteger el vehículo contra amenazas informáticas, dada su dependencia de redes externas.
- **Ejemplos:** Robotaxis de Waymo en Phoenix, autobuses autónomos de Navya en entornos controlados, vehículos de reparto de Nuro.
- **Limitaciones:** La definición y validación del ODD es compleja, especialmente en entornos urbanos con variables impredecibles. La transición a una MRC fuera del ODD depende de infraestructuras viales avanzadas, que no están disponibles en muchas regiones.
- **Estado actual:** El nivel 4 lidera las aplicaciones comerciales actuales, particularmente en servicios de movilidad autónoma y logística. Su escalabilidad está limitada por factores regulatorios, de infraestructura y de aceptación pública.

Nivel 5: Automatización total

El nivel 5 representa la autonomía completa, donde el vehículo puede ejecutar la DDT en cualquier condición de conducción y entorno que un conductor humano competente manejaría, sin restricciones de ODD (Schwartzing, 2018). No requiere controles manuales ni intervención humana, funcionando como un sistema íntegramente autónomo.

- **Tecnologías asociadas:**
 - Algoritmos de inteligencia artificial con capacidad de generalización para adaptarse a entornos desconocidos sin mapas predefinidos.
 - Sensores universales capaces de operar en condiciones extremas, como tormentas, niebla o carreteras sin señalizar.

- Conectividad V2X integral para optimizar decisiones de tráfico a nivel global mediante comunicación continua con otros agentes.
- Diseños de cabina sin controles tradicionales, orientados a espacios de trabajo o entretenimiento.
- **Ejemplos:** No existen implementaciones comerciales. Prototipos como Zoox o el Cruise Origin de General Motors apuntan a este nivel, pero operan en entornos experimentales.



Ilustración 8 Cruise Origin (Miller, 2025)

- **Limitaciones:** Alcanzar el nivel 5 requiere avances en inteligencia artificial general, robustez sensorial e infraestructura global estandarizada. La complejidad de entornos impredecibles y la necesidad de decisiones en tiempo real representan barreras significativas.
- **Estado actual:** El nivel 5 es un objetivo a largo plazo, con proyecciones que lo sitúan entre 2030 y 2050. Aunque lejano, impulsa innovaciones que benefician a niveles inferiores, redefiniendo el futuro de la movilidad.

2.5. Hardware en vehículos autónomos

El desarrollo de los vehículos autónomos se sustenta en un ecosistema tecnológico sofisticado que combina hardware y software para permitir la percepción del entorno, el procesamiento de datos y la ejecución de maniobras sin intervención humana. El hardware constituye la base física de estos sistemas, proporcionando las herramientas necesarias para recopilar información en tiempo real, interpretar el contexto vial y controlar el vehículo con precisión. Este apartado se centra en los componentes de hardware esenciales, describiendo su función, características técnicas, limitaciones y estado actual en el contexto de los coches autónomos.

2.5.1. Sensores para la percepción del entorno

La capacidad de un vehículo autónomo para comprender su entorno depende de un conjunto de sensores que operan de manera complementaria para generar una representación detallada y redundante del mundo exterior. Estos dispositivos son el equivalente a los sentidos humanos, capturando datos en todas las direcciones para detectar obstáculos, señales de tráfico, peatones y otros vehículos.

El LIDAR (*Light Detection and Ranging*) (Shan, 2020) utiliza pulsos láser para crear mapas tridimensionales de alta resolución del entorno. Al emitir millones de rayos láser por segundo y medir el tiempo de retorno, el LIDAR calcula distancias con precisiones de centímetros, lo que resulta crítico para identificar objetos pequeños o lejanos, como un peatón a 100 metros. Su rango típico alcanza los 200 metros, y su capacidad para operar en condiciones de baja luminosidad lo hace invaluable. Sin embargo, su eficacia disminuye en lluvia intensa o niebla, y su alto coste (decenas de miles de euros en modelos de alta gama) ha llevado a debates sobre su escalabilidad. Empresas como Waymo integran LIDAR como componente central, mientras que otras, como Tesla, lo descartan en favor de enfoques basados en cámaras.

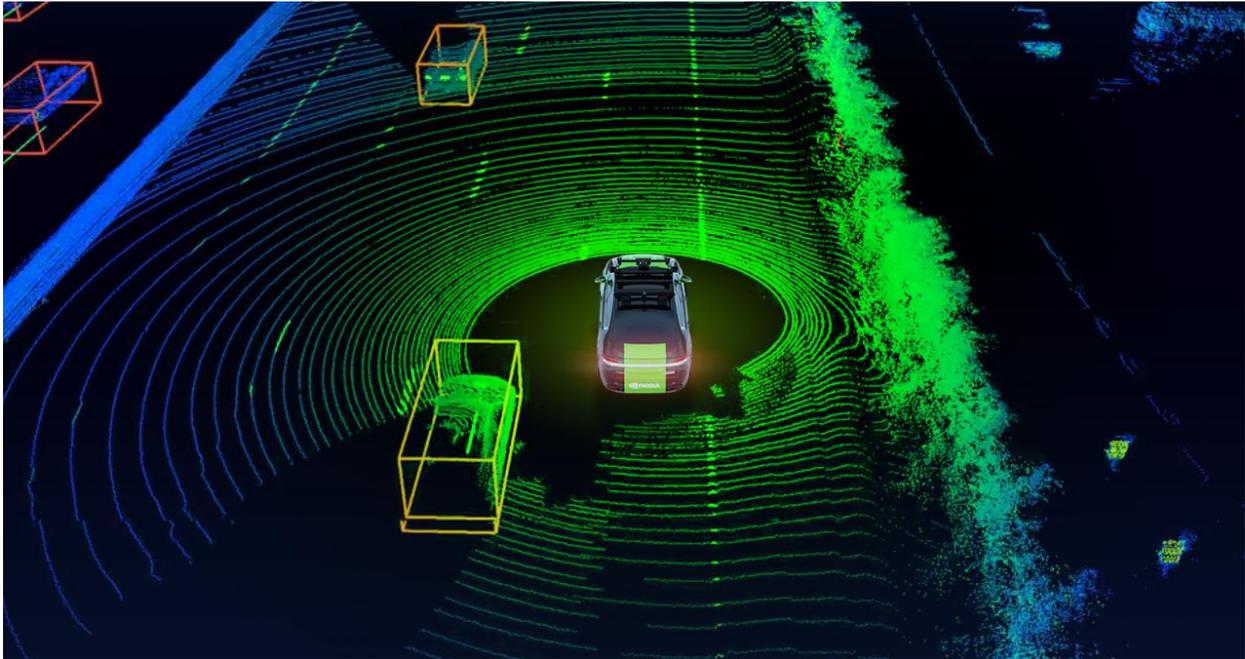


Ilustración 9: Funcionamiento del Lidar (Badoni, 2021)

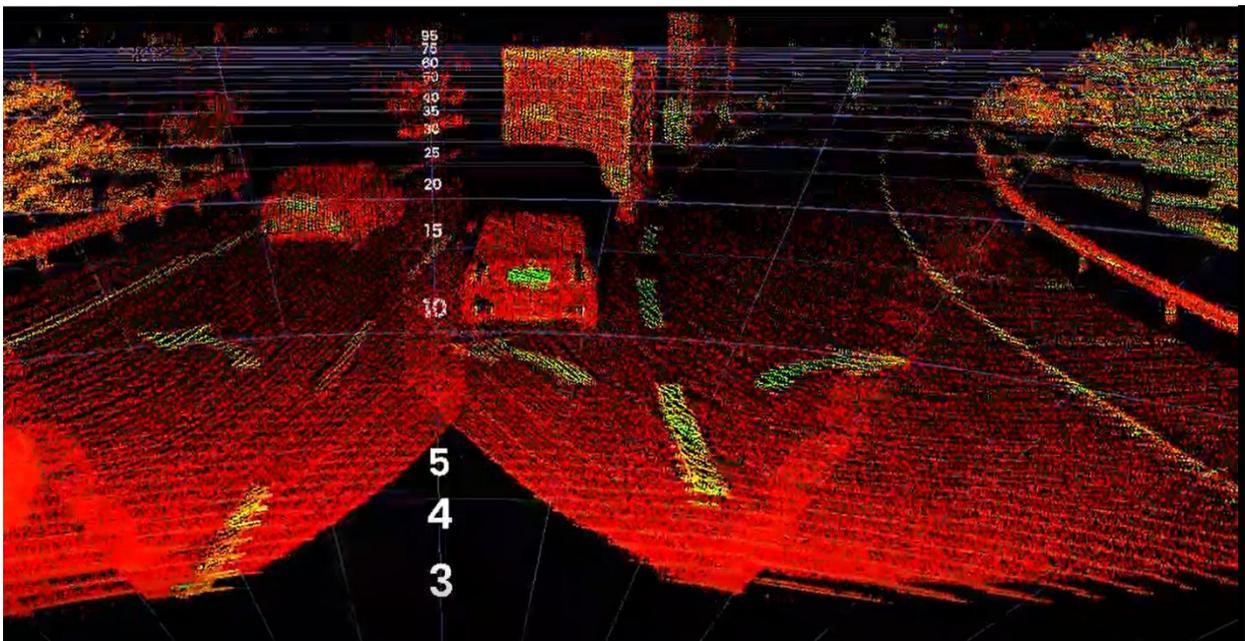


Ilustración 10: Vista real del sistema InnovizOne (Badoni, 2021)

Los radares emplean ondas de radio para detectar objetos y medir su velocidad relativa, destacando por su robustez en condiciones climáticas adversas, como tormentas o niebla densa. Operan en frecuencias milimétricas (generalmente entre 24 y 77 GHz), lo que les permite atravesar partículas

de agua o polvo. Aunque ofrecen menor resolución espacial que el LIDAR, son esenciales para funciones como el control de crucero adaptativo o la detección de vehículos en carriles adyacentes. Su coste más bajo y su madurez tecnológica los han convertido en un estándar en vehículos de nivel 2 y superiores, aunque su dependencia de algoritmos para interpretar datos puede introducir ruido en entornos urbanos complejos.

Las cámaras proporcionan imágenes de alta definición para tareas de visión por computadora, como la identificación de señales de tráfico, líneas de carril y comportamientos de peatones (Chen L. S.). Equipadas con sensores CMOS y lentes de gran angular, pueden capturar detalles en un rango dinámico amplio, adaptándose a cambios de iluminación. Su bajo coste y versatilidad las hacen omnipresentes, pero su rendimiento se ve afectado en condiciones de poca luz o en escenarios con objetos parcialmente ocultos. Además, requieren un procesamiento computacional intensivo para traducir imágenes bidimensionales en modelos tridimensionales del entorno, lo que impone demandas significativas al hardware de cálculo.

Por último, los sensores ultrasónicos complementan el sistema en distancias cortas, especialmente en maniobras a baja velocidad, como el aparcamiento automático. Utilizan ondas sonoras de alta frecuencia para medir distancias de hasta 5 metros, siendo económicos y fiables en entornos controlados. Sin embargo, su rango limitado y su sensibilidad a superficies irregulares los relegan a funciones secundarias.

La integración de estos sensores requiere un diseño redundante para garantizar la fiabilidad. Por ejemplo, un vehículo autónomo de nivel 4 puede combinar múltiples LIDAR, radares y cámaras para cubrir puntos ciegos y mitigar fallos individuales. Esta redundancia incrementa la complejidad y el coste, pero es indispensable para alcanzar los estándares de seguridad exigidos en entornos reales.

2.5.2. Sistemas de posicionamiento y navegación

La navegación precisa es un pilar fundamental para los vehículos autónomos, que deben conocer su ubicación exacta en todo momento para planificar rutas y ejecutar maniobras. El hardware de posicionamiento incluye sistemas globales de navegación por satélite (*Global Navigation Satellite*

Systems, GNSS), como GPS, GLONASS o Galileo, que ofrecen una precisión de varios metros en condiciones óptimas. Sin embargo, en entornos urbanos con edificios altos o túneles, la señal satelital puede degradarse, lo que obliga a complementar el GNSS con otras tecnologías.

Los sistemas inerciales de navegación (INS), basados en acelerómetros y giroscopios MEMS (*Micro-Electro-Mechanical Systems*), miden los movimientos del vehículo para estimar su posición relativa cuando las señales satelitales no están disponibles (Titterton, 2004). Estos sistemas son precisos a corto plazo, pero acumulan errores con el tiempo, lo que requiere una integración constante con datos de sensores externos.

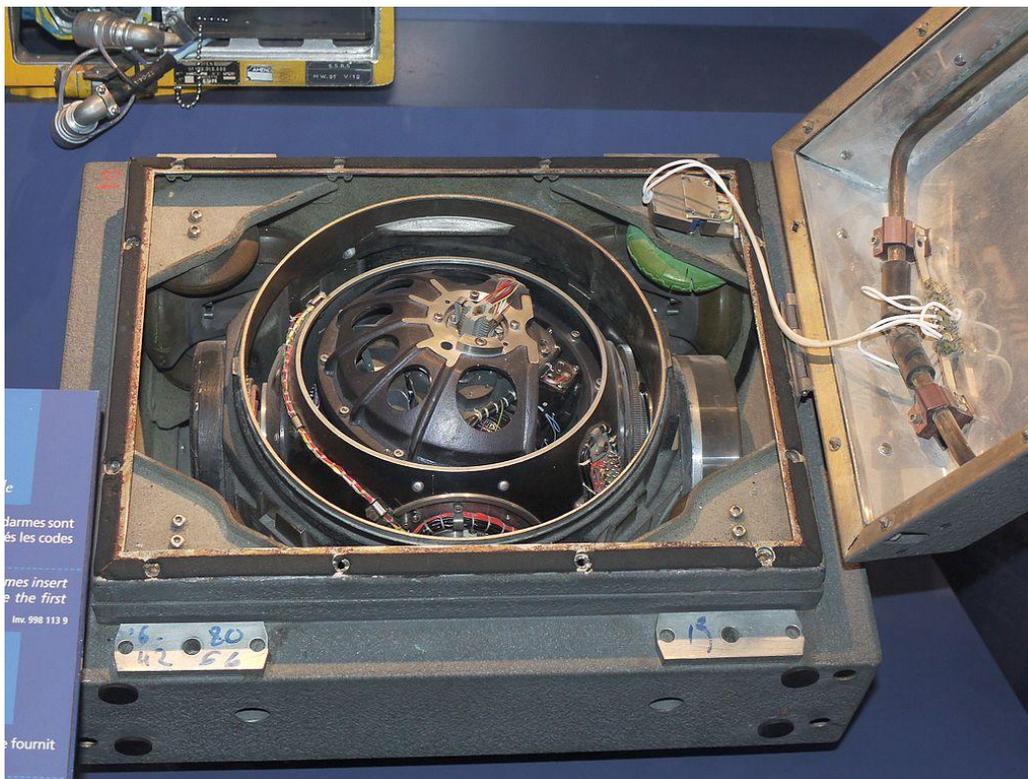


Ilustración 11: INS de un misil balístico francés, Museo del Aire y del Espacio, Paris

Además, las unidades de odometría visual utilizan cámaras para rastrear puntos de referencia en el entorno, complementando los datos de posicionamiento en escenarios donde el GNSS falla. Este enfoque, combinado con mapas de alta definición (*HD Maps*), permite al vehículo ubicarse con una precisión de centímetros, lo que es crucial para maniobras en entornos urbanos densos. Los mapas digitales, almacenados localmente o en la nube, contienen información detallada sobre

carriles, semáforos y límites de velocidad, pero su mantenimiento exige una infraestructura robusta para actualizaciones en tiempo real.

2.5.3. Actuadores para el control del vehículo

Una vez que el vehículo ha percibido su entorno y determinado una acción, los actuadores convierten las decisiones del sistema en movimientos físicos. Estos dispositivos reemplazan los controles mecánicos tradicionales por sistemas electrónicos (*by-wire*), que ofrecen mayor precisión y rapidez de respuesta.

Los sistemas de dirección (*steer-by-wire*) (Li, 2021) utilizan motores eléctricos para girar las ruedas, eliminando la conexión mecánica entre el volante y el eje. Uno de los vehículos más polémicos en introducir esta tecnología es el Tesla Cybertruck, que es el primer vehículo con una dirección completamente digital sin una conexión mecánica con las ruedas. Esta tecnología permite ajustes dinámicos en función de las condiciones de la carretera y reduce el peso del vehículo, aunque requiere redundancia para evitar fallos críticos. De manera similar, los sistemas de aceleración y frenado (*throttle-by-wire* y *brake-by-wire*) emplean actuadores electrónicos para modular la potencia del motor y la presión de los frenos, integrándose con algoritmos que optimizan la eficiencia energética y la seguridad.

La transición hacia sistemas *by-wire* plantea desafíos en términos de ciberseguridad, ya que un fallo o un ataque informático podría comprometer el control del vehículo. Por ello, los actuadores suelen incorporar múltiples capas de redundancia y sistemas de diagnóstico en tiempo real para detectar anomalías.

2.5.4. Unidades de procesamiento

El núcleo del hardware de un vehículo autónomo es su unidad de procesamiento, encargada de integrar los datos de los sensores, ejecutar algoritmos de inteligencia artificial y generar comandos para los actuadores. Estas unidades, conocidas como *Electronic Control Units (ECUs)* o plataformas de cómputo centralizadas, deben manejar volúmenes masivos de datos en milisegundos para garantizar decisiones en tiempo real (Chen T. L., 2021).

Los procesadores modernos, como los chips NVIDIA DRIVE, Intel Mobileye o los ASIC diseñados por Tesla, combinan CPUs, GPUs y aceleradores específicos para tareas de *machine learning*. Por ejemplo, el procesamiento de imágenes de cámaras requiere redes neuronales profundas que ejecutan billones de operaciones por segundo, mientras que la fusión sensorial demanda una sincronización precisa de datos heterogéneos. Estas plataformas suelen consumir cientos de vatios, lo que impone restricciones en el diseño térmico y energético del vehículo.

La tendencia hacia arquitecturas centralizadas, frente a las ECUs distribuidas de los vehículos tradicionales, permite una integración más eficiente, pero aumenta la dependencia de un único punto de fallo. Para mitigar este riesgo, los sistemas de nivel 4 y superiores incorporan procesadores redundantes y mecanismos de *failover* que garantizan la continuidad operativa.

2.6. Tecnologías clave de software en vehículos autónomos

El software de los vehículos autónomos es el pilar que permite transformar los datos del hardware en decisiones operativas, integrando percepción, planificación y control para lograr una conducción segura y eficiente. En el estado del arte actual, los enfoques de desarrollo del software se dividen principalmente en modelos basados en reglas preprogramadas y modelos de aprendizaje automático, cada uno con características, ventajas y limitaciones que influyen tanto en su rendimiento técnico como en las implicaciones éticas de su implementación (Koopman, 2017). Además, diversas soluciones y plataformas están consolidando estos enfoques, ofreciendo herramientas que combinan ambos paradigmas para avanzar hacia niveles superiores de autonomía.

Los modelos de reglas preprogramadas se fundamentan en un conjunto de instrucciones explícitas definidas por ingenieros, diseñadas para interpretar los datos de los sensores y aplicar reglas específicas en función del entorno. Por ejemplo, el sistema puede estar programado para detenerse ante un semáforo en rojo, mantener una distancia de seguridad de 2 metros respecto al vehículo precedente o reducir la velocidad en curvas pronunciadas. Este enfoque destaca por su predictibilidad, ya que las decisiones son deterministas y fácilmente auditables, lo que facilita su alineación con normativas viales y su uso en entornos estructurados, como autopistas o áreas con señalización clara. Sin embargo, su principal limitación radica en su falta de adaptabilidad frente

a situaciones nuevas o ambiguas, como un peatón que cruza de forma inesperada o una señalización vial poco visible, donde la rigidez de las reglas puede derivar en respuestas ineficientes o incluso inseguras. Desde el punto de vista ético, este método permite incorporar principios morales claros, como priorizar siempre la seguridad de los peatones, pero su incapacidad para manejar dilemas complejos, como elegir entre dos males inevitables en un accidente, plantea desafíos significativos.

En contraste, los modelos de aprendizaje automático se basan en el uso de grandes volúmenes de datos para entrenar algoritmos, principalmente redes neuronales profundas, que permiten al vehículo aprender patrones y tomar decisiones en tiempo real. Este enfoque utiliza datos de sensores, como imágenes de cámaras o mapas de LIDAR, para entrenar modelos que pueden predecir comportamientos, como la trayectoria de un ciclista, o adaptarse a entornos dinámicos, como una intersección urbana con tráfico denso. La capacidad de estos sistemas para manejar situaciones complejas y nuevas los hace ideales para niveles superiores de autonomía, ya que pueden aprender de la experiencia y mejorar su rendimiento con el tiempo. Sin embargo, su dependencia de la calidad y cantidad de los datos de entrenamiento introduce limitaciones importantes (Amodei, 2016): un conjunto de datos sesgado o incompleto puede llevar a decisiones erróneas, como no reconocer a un peatón en un entorno poco representado en los datos. Además, estos modelos operan como una "caja negra", lo que dificulta entender cómo se toman las decisiones, un aspecto crítico desde la perspectiva ética, ya que complica la asignación de responsabilidades en caso de fallos y plantea preguntas sobre la transparencia y la equidad de las decisiones.

Un aspecto clave del aprendizaje automático es el uso de redes neuronales, que forman el núcleo de muchos sistemas de conducción autónoma. Estas redes están compuestas por capas de nodos interconectados que procesan datos de entrada, como imágenes o señales de sensores, para generar salidas, como comandos de dirección o frenado. Durante el entrenamiento, las redes ajustan sus parámetros internos analizando millones de ejemplos etiquetados, como imágenes de carreteras con indicaciones de los objetos presentes, para minimizar errores y aprender patrones complejos. Por ejemplo, una red neuronal convolucional puede identificar un semáforo en una imagen al detectar primero bordes y colores, y luego correlacionarlos con patrones aprendidos, mientras que

una red recurrente puede predecir la trayectoria de un vehículo basándose en su movimiento previo. Aunque estas tecnologías han avanzado significativamente, su opacidad y su dependencia de datos masivos generan preocupaciones éticas, especialmente en lo que respecta a sesgos en los datos y la dificultad de garantizar decisiones justas en situaciones críticas.

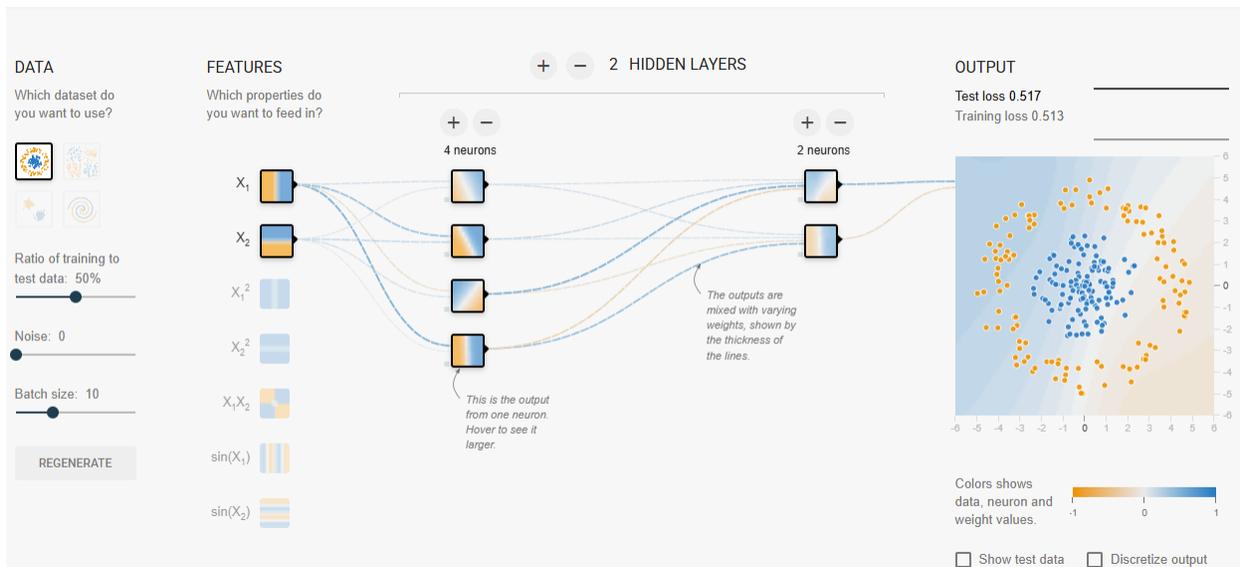


Ilustración 12: Herramienta interactiva de TensorFlow para divulgación sobre redes neuronales (TensorFlow, s.d.)

También se están desarrollando soluciones que combinan ambos enfoques, buscando equilibrar la predictibilidad de las reglas preprogramadas con la adaptabilidad del aprendizaje automático. Este enfoque híbrido utiliza reglas predefinidas para garantizar un comportamiento seguro y conforme a la ley en situaciones críticas, como detenerse ante un cruce peatonal, mientras que el aprendizaje automático se emplea para optimizar decisiones en escenarios menos predecibles, como la navegación en tráfico denso. Plataformas como NVIDIA DRIVE y Motional han adoptado este método, integrando módulos de reglas con algoritmos de aprendizaje para ofrecer flexibilidad y seguridad en niveles intermedios de autonomía (L3/L4). Por ejemplo, NVIDIA DRIVE combina un núcleo de reglas para cumplir con normativas viales con redes neuronales que aprenden a anticipar comportamientos de otros conductores, logrando un equilibrio entre transparencia y adaptabilidad.

Además, existen plataformas abiertas que están impulsando el desarrollo del software para vehículos autónomos, ofreciendo frameworks personalizables que combinan ambos paradigmas.

Apollo, desarrollado por Baidu, es un ejemplo destacado, proporcionando un entorno de código abierto que integra reglas preprogramadas y aprendizaje automático para permitir a los desarrolladores crear sistemas adaptados a entornos específicos, como áreas urbanas en China. De manera similar, Autoware ofrece un framework flexible que combina módulos de percepción, planificación y control, permitiendo a los investigadores implementar soluciones que equilibren reglas y aprendizaje según las necesidades del proyecto. Estas plataformas no solo facilitan la innovación, sino que también abren la puerta a la colaboración global, un aspecto crucial para estandarizar y escalar las tecnologías de software en el sector.

2.7. Regulación actual de los vehículos autónomos

La regulación de los vehículos autónomos es un aspecto crucial para su desarrollo y despliegue, ya que establece el marco legal que garantiza su seguridad, define responsabilidades y aborda los dilemas éticos que serán centrales en este trabajo. A medida que la tecnología ha avanzado desde los experimentos históricos hasta las aplicaciones actuales descritas en los apartados previos, los gobiernos han trabajado para adaptar sus normativas, aunque con ritmos y enfoques dispares.

2.7.1. Regulación a nivel global

A nivel global, la regulación de los vehículos autónomos está marcada por esfuerzos de armonización, pero también por una notable fragmentación. La Convención de Viena sobre la Circulación Vial de 1968, que rige en muchos países, originalmente exigía la presencia de un conductor humano, lo que limitaba el uso de vehículos autónomos. Sin embargo, una enmienda adoptada por el Foro Mundial para la Armonización de Reglamentos de Vehículos de la ONU (UNECE) en 2022 introdujo el artículo 34bis, que permite considerar que el requisito del conductor se cumple si el vehículo utiliza un sistema de conducción automatizada (ADS) que cumpla con las normativas técnicas nacionales e internacionales. Este cambio ha facilitado el despliegue de vehículos autónomos en países signatarios, como Alemania y Francia, aunque su implementación varía según las leyes locales. Por otro lado, países como Japón, que no es signatario de la Convención de Viena, han avanzado rápidamente: desde 2020, Japón permite vehículos de nivel 3 en vías públicas y ha establecido procesos de certificación para garantizar su seguridad. En China,

más de 20 ciudades, incluidas Pekín y Shanghái, han autorizado pruebas de nivel 4, con un enfoque estratégico para liderar el mercado global, apoyado por una regulación que fomenta la innovación tecnológica. Sin embargo, la falta de un marco global unificado sigue siendo un desafío, ya que las diferencias en normativas dificultan la interoperabilidad transfronteriza, un aspecto crítico para la movilidad autónoma y los dilemas éticos relacionados con la responsabilidad en accidentes internacionales.

2.7.2. Regulación en Europa

En Europa, la regulación de los vehículos autónomos ha avanzado hacia una mayor armonización, aunque persisten diferencias entre los estados miembros. La Unión Europea ha priorizado la seguridad vial y la innovación tecnológica, como se refleja en la Regulación General de Seguridad (GSR) de 2019, que entró en vigor en julio de 2022 y establece requisitos de tipo para vehículos automatizados y conectados. Esta normativa introdujo sistemas obligatorios de asistencia al conductor, como el frenado de emergencia autónomo, y sentó las bases legales para aprobar vehículos de nivel 3 y 4, como robotaxis y autobuses urbanos, con requisitos estrictos de ciberseguridad, registro de datos y monitoreo de seguridad. Además, la UE ha impulsado proyectos de conectividad, como los corredores transfronterizos 5G, que permiten probar la seguridad vial y la interoperabilidad de los vehículos autónomos en rutas como la de Porto-Vigo, entre España y Portugal. A nivel nacional, países como Alemania han liderado el camino: desde 2021, Alemania permite vehículos de nivel 4 en áreas definidas, eliminando la necesidad de un conductor de seguridad, y ha establecido un marco de responsabilidad que asigna la culpa al fabricante en caso de fallos del sistema. Francia, por su parte, ha implementado un marco regulatorio desde 2021 que permite pruebas y despliegue de vehículos autónomos en servicios de movilidad, con más de 50 proyectos de prueba desde 2014. Sin embargo, la falta de una certificación unificada a nivel europeo, que se espera para 2027, sigue complicando el despliegue transfronterizo, un desafío que impacta directamente en los dilemas éticos relacionados con la ética cultural y la personalización ética, ya que las normativas nacionales reflejan valores y prioridades diferentes.

2.7.3. Regulación en España

En España, la regulación de los vehículos autónomos se encuentra en una fase de desarrollo, con un enfoque inicial en pruebas y una progresiva adaptación al marco europeo. Desde 2015, la Dirección General de Tráfico (DGT) ha permitido pruebas de vehículos autónomos de niveles 3 a 5 en vías públicas, siempre que se obtenga una autorización previa del gobierno, según la Instrucción 15/V-113. Esta normativa exige que los solicitantes de pruebas cuenten con un seguro obligatorio que cubra posibles incidentes, aunque no existen requisitos específicos de seguro para vehículos autónomos más allá de las leyes generales de circulación. En 2018, la DGT firmó un acuerdo con Mobileye para preparar la infraestructura vial y las políticas regulatorias, con el objetivo de reducir accidentes y facilitar la integración de la autonomía. Además, España participa en proyectos europeos, como el corredor 5G Porto-Vigo previamente mencionado, que prueba la conectividad y la seguridad de los vehículos autónomos en un contexto transfronterizo. A nivel legislativo, el Congreso de los Diputados aprobó en 2017 una Proposición No de Ley para promover el desarrollo de los vehículos autónomos, y se están elaborando modificaciones a las leyes de seguros para crear un marco legal más completo, aunque aún no se han implementado normativas específicas para el uso comercial de vehículos de nivel 3 o superior. Esta situación refleja un enfoque cauteloso, que prioriza la seguridad y la alineación con las normativas europeas.

2.8. Ética en vehículos autónomos

La ética en los vehículos autónomos constituye el eje central de este trabajo, ya que la introducción de sistemas autónomos en la conducción plantea dilemas morales que desafían los paradigmas tradicionales de responsabilidad, equidad y seguridad vial. Este apartado del estado del arte se centra en los marcos éticos, enfoques filosóficos, dilemas clásicos y debates actuales que configuran el panorama ético de los coches autónomos, sentando las bases para un análisis profundo de cómo estas cuestiones impactan en su diseño, implementación y aceptación social. La programación de decisiones éticas, ya explorada en el apartado de software, se convierte aquí en el núcleo del análisis, destacando cómo las elecciones técnicas reflejan valores morales que deben ser cuidadosamente definidos para garantizar un comportamiento justo y responsable.

What should the self-driving car do?

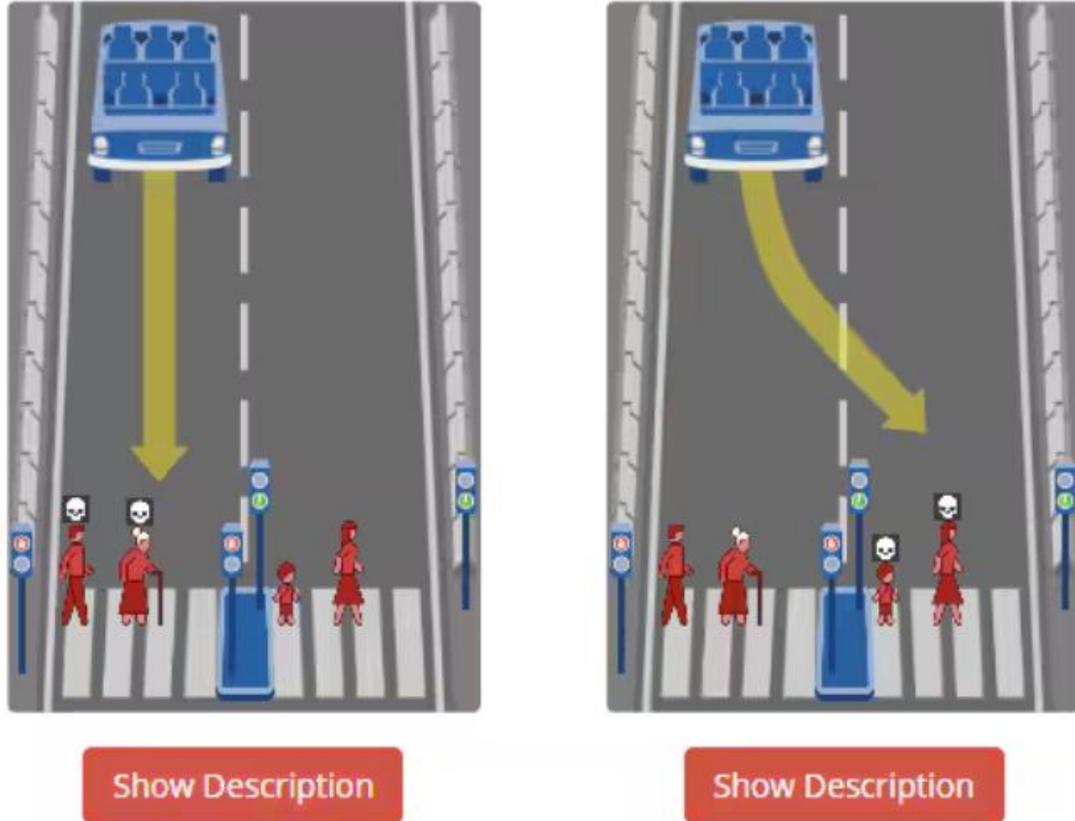


Ilustración 13: MIT Moral Machine (MIT, s.d.)

Un marco clásico para abordar la ética de los sistemas autónomos son las tres leyes de la robótica propuestas por Isaac Asimov en 1942, que, aunque ficticias, ofrecen un punto de partida para reflexionar sobre las prioridades éticas de un vehículo autónomo. La primera ley establece que un robot no debe dañar a un ser humano ni, por inacción, permitir que sufra daño, lo que implica que un coche autónomo debe minimizar el riesgo para todos los involucrados, ya sean ocupantes, peatones u otros conductores. Sin embargo, en un accidente inevitable, ¿cómo se define "minimizar el daño"? La segunda ley exige que un robot obedezca las órdenes humanas, salvo que entren en conflicto con la primera ley, lo que plantea preguntas sobre el grado de control que los usuarios deben tener sobre el vehículo: ¿puede un pasajero ordenar al sistema que priorice su seguridad sobre la de otros, y en qué medida es ético permitirlo? La tercera ley indica que un robot debe proteger su propia existencia, siempre que no contradiga las leyes anteriores, lo que podría

traducirse en evitar daños al vehículo, pero solo si no compromete la seguridad humana. Estas leyes, aunque teóricamente útiles, resultan insuficientes frente a la complejidad de los dilemas reales, donde las decisiones no son binarias y los valores en juego; como la vida, la justicia o la autonomía, entran en conflicto.

Los enfoques filosóficos proporcionan herramientas más matizadas para abordar estos dilemas, cada uno con implicaciones directas para la programación ética de los vehículos autónomos. El utilitarismo, que busca maximizar el bienestar general, sugiere que un vehículo debería tomar decisiones que minimicen el daño total, incluso si eso implica sacrificar a un individuo para salvar a un grupo mayor. Por ejemplo, en un escenario donde el coche debe elegir entre atropellar a un peatón o desviarse y arriesgar la vida de sus ocupantes, un enfoque utilitarista podría programarse para calcular el "coste" de cada acción y optar por la que resulte en menos víctimas. Sin embargo, esta lógica genera controversias, ya que priorizar el número de vidas salvadas podría percibirse como injusto, especialmente si introduce variables como la edad o el estatus social de los involucrados, lo que lleva a debates sobre discriminación. La deontología, en cambio, se basa en reglas universales, como no causar daño intencionadamente, lo que podría traducirse en programar el vehículo para evitar acciones que resulten en daño directo, independientemente de las consecuencias globales. Este enfoque, aunque más predecible, puede ser problemático en situaciones donde todas las opciones implican algún tipo de daño, como en un accidente inevitable. Finalmente, la ética de la virtud propone que el vehículo emule las decisiones de un conductor humano ideal, incorporando valores como la prudencia o la empatía, lo que conecta con los modelos de aprendizaje automático mencionados en el apartado de software: las redes neuronales podrían entrenarse para imitar comportamientos éticos, pero con el riesgo de reproducir sesgos presentes en los datos, como priorizar a ciertos grupos sobre otros debido a una representación desigual en el entrenamiento.

Los dilemas éticos clásicos son fundamentales para entender las tensiones que enfrenta la programación de vehículos autónomos, ya que ilustran los desafíos prácticos de traducir principios morales en algoritmos. El problema del tranvía, adaptado al contexto automotriz, es un ejemplo paradigmático: si un vehículo autónomo se enfrenta a una colisión inevitable, ¿debería atropellar a un grupo de peatones o desviarse y arriesgar la vida de un solo peatón o de sus ocupantes? Este

dilema pone de manifiesto la dificultad de asignar valor a las vidas humanas, un problema que se agrava si se introducen variables como la edad, el género o la condición social de los involucrados, lo que lleva al riesgo de discriminación en las decisiones. Por ejemplo, un algoritmo que priorice a los ocupantes sobre los peatones podría reflejar sesgos culturales o comerciales, mientras que uno que priorice a los peatones podría ser percibido como injusto por los propietarios del vehículo. Este riesgo de discriminación está directamente relacionado con los modelos de aprendizaje automático, donde los sesgos en los datos de entrenamiento, como una sobrerrepresentación de entornos urbanos occidentales, podrían llevar a decisiones que desfavorezcan a grupos menos representados, como peatones en áreas rurales o en países con infraestructuras diferentes a los que el modelo ha sido expuesto en el entrenamiento. Otro dilema crucial es la asignación de responsabilidad en caso de accidente: si un vehículo autónomo causa un daño, ¿quién es responsable, el fabricante, el programador o el propietario? La opacidad de los sistemas de "caja negra", como las redes neuronales, complica esta cuestión, ya que dificulta determinar cómo se tomó una decisión, lo que resalta la necesidad de desarrollar software más transparente y explicable, un desafío técnico y ético que será central en el desarrollo de este trabajo.

Los debates éticos actuales amplían estas cuestiones, explorando las implicaciones sociales, culturales y prácticas de la autonomía en la conducción. La personalización ética es un tema candente: ¿deberían los vehículos permitir a los usuarios definir sus propias preferencias éticas, como priorizar la seguridad de los ocupantes sobre la de los peatones, o deben seguir un estándar universal? Permitir personalización podría aumentar la aceptación pública, pero también generar desigualdades, ya que los usuarios con más recursos podrían configurar sistemas que los favorezcan, lo que plantea preguntas sobre justicia distributiva. La ética cultural añade otra capa de complejidad, ya que los valores morales varían entre regiones; por ejemplo, en algunas culturas podría ser más aceptable priorizar al colectivo sobre el individuo, mientras que en otras se enfatiza la protección de los ocupantes, lo que requiere que los fabricantes decidan si programar sistemas globales o adaptarlos a contextos locales. Este debate conecta con la necesidad de datos diversos en los modelos de aprendizaje automático, ya que un sistema entrenado con datos sesgados podría tomar decisiones que no respeten las normas culturales de ciertas regiones, generando rechazo o incluso riesgos de seguridad. Finalmente, el debate sobre autonomía versus control humano

cuestiona el equilibrio entre la independencia del vehículo y la intervención humana: un exceso de autonomía podría reducir la capacidad de los usuarios para intervenir en situaciones críticas, mientras que un control humano excesivo podría anular los beneficios de la automatización, como la reducción de errores humanos, un aspecto que será explorado en profundidad en los siguientes apartados del trabajo.

3. Motivación

El desarrollo de los vehículos autónomos representa una de las innovaciones tecnológicas más significativas del siglo XXI, con un potencial transformador que ha captado la atención de la comunidad académica e industrial, especialmente en el ámbito de la Ingeniería Industrial. La motivación de este trabajo surge de la necesidad de analizar dos aspectos fundamentales: el impacto que estos vehículos pueden tener en la sociedad y la urgencia de definir cómo deben comportarse desde una perspectiva ética, garantizando transparencia y generando confianza para su aceptación social, un enfoque que se alinea con los retos identificados en el estado del arte.

El primer aspecto que impulsa este estudio es la necesidad de comprender el impacto de los vehículos autónomos, que abarca beneficios sociales, económicos y ambientales, como se ha señalado en el apartado de contexto actual del estado del arte. Uno de los argumentos más sólidos a favor de esta tecnología es su capacidad para reducir los accidentes viales, dado que más del 90% de estos son causados por errores humanos, como distracciones o conducción imprudente. Al eliminar este factor, los coches autónomos podrían salvar vidas y disminuir los costes asociados a los accidentes, que afectan tanto a nivel humano como económico. Además, su capacidad para optimizar rutas y coordinarse puede reducir los atascos, lo que no solo acorta los tiempos de viaje, sino que también disminuye las emisiones de CO₂, contribuyendo a la sostenibilidad ambiental, un objetivo clave en el contexto actual de cambio climático. Otro impacto relevante es la inclusión que esta tecnología puede ofrecer a personas con discapacidad, como aquellas con limitaciones visuales o de movilidad, quienes podrían ganar independencia al acceder a un medio de transporte autónomo, mejorando su calidad de vida de manera significativa. En el primer apartado de análisis del proyecto se valorarán los diferentes impactos para dar motivos de peso a la necesidad del desarrollo del coche autónomo como concepto más allá de la movilidad.

El segundo aspecto que motiva este trabajo es la necesidad de definir cómo deben comportarse los coches autónomos para garantizar que sus decisiones sean éticamente responsables y socialmente aceptadas. La programación del software ya sea mediante reglas preprogramadas o aprendizaje automático, implica tomar decisiones en situaciones críticas que reflejan valores éticos. Estas decisiones no solo requieren un marco ético claro, sino también transparencia para que la sociedad

entienda cómo se toman y pueda confiar en el sistema. La opacidad de tecnologías como las redes neuronales, que dificulta explicar las decisiones del vehículo, es un obstáculo significativo para generar confianza, especialmente tras incidentes como el accidente fatal de Uber en 2018, donde la falta de claridad sobre el comportamiento del sistema generó controversias. Este trabajo se motiva en la urgencia de explorar cómo integrar principios éticos en el diseño de los vehículos autónomos, asegurando que sean justos, equitativos y adaptados a contextos culturales. El objetivo es contribuir al desarrollo de una tecnología que no solo sea avanzada, sino también responsable, alineándose con los valores humanos y fomentando una aceptación social que será clave para su éxito.

4. Objetivos del Proyecto

El propósito de este trabajo es abordar de manera integral las implicaciones éticas asociadas a los vehículos autónomos, examinando los desafíos que emergen de su integración en la sociedad y proponiendo estrategias para garantizar un diseño responsable que alinee la tecnología con los valores humanos. Para ello, se han definido los siguientes objetivos, que estructuran el análisis y las propuestas del estudio, abarcando dimensiones éticas, tecnológicas, sociales, legales y prácticas, con el fin de contribuir al desarrollo de una movilidad autónoma que sea segura, equitativa y socialmente aceptada.

El primer objetivo es analizar los dilemas éticos de la conducción autónoma, profundizando en las situaciones donde los vehículos deben tomar decisiones moralmente complejas. Esto incluye explorar casos en los que el sistema debe elegir entre diferentes opciones con consecuencias humanas, como priorizar entre la seguridad de los ocupantes y la de los peatones en un accidente inevitable, o decidir cómo actuar frente a grupos de personas con características diversas. El análisis busca identificar los principios éticos que deberían guiar estas decisiones, considerando tensiones entre valores fundamentales como la seguridad, la justicia y la equidad, y evaluando cómo estas decisiones pueden impactar en la percepción pública de la tecnología.

El segundo objetivo se centra en evaluar tecnologías actuales y futuras que soportan los vehículos autónomos, con especial atención a su influencia en la toma de decisiones éticas. Esto implica examinar los sistemas de aprendizaje automático y las reglas preprogramadas, analizando cómo cada enfoque incorpora o podría incorporar principios éticos en su funcionamiento. Por ejemplo, se estudiará cómo los algoritmos de aprendizaje automático pueden entrenarse para reflejar valores éticos sin introducir sesgos, y cómo las reglas preprogramadas pueden garantizar consistencia en situaciones críticas. Además, se evaluarán las limitaciones tecnológicas, como la opacidad de algunos sistemas, y su impacto en la confiabilidad y la transparencia, buscando formas de mejorar la capacidad de estas tecnologías para tomar decisiones éticamente responsables.

El tercer objetivo es determinar el impacto social y legal de los modelos de conducción autónoma, evaluando cómo estos sistemas afectan a la sociedad y al marco legal. Desde el punto

de vista social, se analizarán aspectos como la aceptación pública, explorando las percepciones y preocupaciones de los usuarios frente a la autonomía, y cómo esta tecnología puede promover la inclusión al facilitar la movilidad de grupos vulnerables, como personas con discapacidad. También se estudiarán las posibles desigualdades que podrían surgir, como el acceso desigual a la tecnología entre diferentes sectores de la población. Desde el punto de vista legal, se examinarán las implicaciones relacionadas con la responsabilidad en caso de incidentes, analizando cómo las normativas actuales definen quién es responsable: fabricante, programador o usuario; y qué ajustes legales son necesarios para garantizar un marco justo y claro que fomente la confianza en la tecnología.

El cuarto objetivo consiste en proponer recomendaciones para fabricantes y reguladores, ofreciendo criterios prácticos y aplicables que permitan integrar principios éticos en el diseño y la regulación de los vehículos autónomos. Estas recomendaciones incluirán estrategias para mejorar la transparencia en los algoritmos, asegurando que las decisiones del vehículo sean comprensibles para los usuarios y las autoridades. También se abordará la adaptabilidad de los sistemas a contextos culturales diversos, proponiendo formas de diseñar vehículos que respeten las diferencias en valores y normas entre regiones. Asimismo, se sugerirán directrices para los reguladores, como la creación de normativas que promuevan un desarrollo equitativo y responsable, fomentando estándares internacionales que equilibren la innovación con la protección de los derechos y la seguridad de los ciudadanos.

5. Alineación con los Objetivos de Desarrollo Sostenible (ODS)

Este trabajo aborda los desafíos éticos, tecnológicos y normativos que surgen en la toma de decisiones de los vehículos autónomos, proponiendo un análisis que no solo busca avanzar en el desarrollo tecnológico, sino también garantizar que su implementación sea responsable y alineada con principios éticos y sociales. En este contexto, el proyecto se alinea con varios Objetivos de Desarrollo Sostenible (ODS) establecidos por la ONU, contribuyendo al progreso en áreas clave relacionadas con la innovación, la sostenibilidad y la justicia. A continuación, se detalla cómo este estudio impacta en tres ODS específicos, destacando su relevancia para el desarrollo de una movilidad autónoma ética y equitativa (Naciones Unidas, s.d.).

En primer lugar, el trabajo contribuye al ODS 9: Industria, Innovación e Infraestructura, que promueve el desarrollo de infraestructuras resilientes, la industrialización sostenible y la innovación tecnológica. Los vehículos autónomos representan un avance significativo en el ámbito de la movilidad, con el potencial de optimizar las infraestructuras de transporte mediante la coordinación eficiente de flujos de tráfico y la reducción de errores humanos. Sin embargo, para que esta innovación sea sostenible y beneficiosa, es crucial que los sistemas autónomos operen bajo principios éticos claros que garanticen su seguridad y fiabilidad. Este estudio analiza cómo integrar estos principios en el diseño y la operación de los vehículos autónomos, asegurando que las decisiones tecnológicas, como las tomadas por algoritmos de aprendizaje automático, no solo sean eficientes, sino también responsables. Al hacerlo, se fomenta un desarrollo tecnológico que no solo impulsa la industria del transporte, sino que también respeta los estándares de sostenibilidad y seguridad, contribuyendo a una innovación que beneficie a largo plazo a la sociedad.

En segundo lugar, el proyecto se alinea con el ODS 11: Ciudades y Comunidades Sostenibles, que busca hacer las ciudades más inclusivas, seguras, resilientes y sostenibles. Los vehículos autónomos tienen el potencial de transformar los entornos urbanos al reducir los accidentes de tráfico, mejorar la fluidez vehicular y optimizar los sistemas de transporte, lo que puede disminuir

la congestión y las emisiones contaminantes. No obstante, para que estas mejoras sean efectivas y socialmente aceptadas, es necesario que su integración en las ciudades se realice de manera transparente y equitativa. Este trabajo propone un marco de referencia que permita a los vehículos autónomos operar en entornos urbanos respetando las necesidades de todos los ciudadanos, incluyendo grupos vulnerables como personas con discapacidad, quienes podrían beneficiarse de una mayor inclusión a través de esta tecnología. Al abordar los dilemas éticos y tecnológicos asociados, el proyecto busca asegurar que los beneficios de la movilidad autónoma se distribuyan de manera justa, promoviendo comunidades urbanas más seguras y sostenibles.

En tercer lugar, el estudio se vincula con el ODS 16: Paz, Justicia e Instituciones Sólidas, que aboga por sociedades pacíficas, acceso a la justicia y la construcción de instituciones responsables y transparentes. La toma de decisiones automatizada en los vehículos autónomos plantea dilemas legales y morales significativos, como la asignación de responsabilidad en caso de accidentes o la necesidad de garantizar que las decisiones del sistema no perpetúen desigualdades. Estas cuestiones requieren regulaciones claras que equilibren la innovación tecnológica con la protección de los derechos de los ciudadanos. Este trabajo tiene como objetivo sentar las bases para definir principios normativos que promuevan la transparencia en los sistemas de inteligencia artificial aplicados a la movilidad, asegurando que las decisiones de los vehículos autónomos sean comprensibles y auditables. Al hacerlo, se contribuye a la creación de instituciones más sólidas y responsables, que puedan regular esta tecnología de manera justa, fomentando la confianza pública y apoyando un marco legal que refleje valores éticos fundamentales.



Ilustración 14: Listado de los Objetivos de Desarrollo Sostenible (Naciones Unidas , s.d.)

6. Modelo sobre los beneficios del coche autónomo

6.1. Modelo de beneficios, ¿por qué?

Tras revisar la bibliografía y el estado del arte sobre los vehículos autónomos, este trabajo inicia su primer capítulo de investigación y modelado, un componente clave del proyecto. La construcción de un marco ético para evaluar esta tecnología requiere analizar las consecuencias de su implementación y sus implicaciones sociales, económicas y ambientales. Uno de los principales desafíos al investigar los coches autónomos es la polarización de los datos disponibles, con posturas que van desde un entusiasmo desmedido por sus beneficios hasta críticas severas sobre sus riesgos. Dado que el transporte terrestre es un pilar fundamental de la sociedad moderna, las opiniones de los relatores suelen estar marcadas por una notable parcialidad.

Para abordar este problema, este estudio contrasta múltiples fuentes con el objetivo de ofrecer una perspectiva imparcial, rigurosa y equilibrada ante la pregunta central: ¿es realmente necesario el vehículo autónomo? Asimismo, se exploran sus posibles beneficios, como la reducción de accidentes y emisiones, frente a los contrargumentos, como los riesgos tecnológicos o el impacto en el empleo. Finalmente, se evalúa si es prioritario acelerar la transición hacia modelos de movilidad autónoma.

Con este fin, la investigación se estructura en tres bloques principales:

- **Mercado del vehículo autónomo:** Un análisis del estado actual de la automatización en la industria automotriz y un pronóstico de su evolución en los próximos 25 años, considerando tendencias tecnológicas y económicas.
- **Incidencias y accidentes:** Una evaluación de cómo los vehículos autónomos podrían influir en la frecuencia y gravedad de los accidentes de tráfico, comparando su desempeño con la conducción humana.
- **Emisiones y sostenibilidad:** Un cálculo de las implicaciones ambientales de una movilidad autónoma, considerando una mayor eficiencia en el uso de recursos frente al posible aumento en la demanda de transporte.

A través de este enfoque, se busca desarrollar un marco integral que permita comprender las oportunidades y desafíos de los vehículos autónomos, contribuyendo a un debate informado sobre su rol en el futuro de la movilidad

6.2. Análisis de mercado.

Es esencial conocer las cifras actuales de vehículos con diferentes niveles de automatización y proyectar su evolución, tanto en términos cuantitativos (mayor número de vehículos con asistencia a la conducción) como cualitativos (aumento del nivel promedio de automatización del parque automotor global).

En 2025, el parque automotor global se estima en ~1,644 millones de vehículos, susceptibles de ser reemplazados por modelos autónomos o parcialmente autónomos, con ventas anuales de ~80 millones de unidades (S&P Global, 2024). Se asume un crecimiento interanual constante del 1.7% para el total de vehículos y las ventas, basado en la media de crecimiento postpandemia (2022-2024). Esta tasa constante se adopta como simplificación, aunque se reconoce que factores como la electrificación o la propia automatización podrían alterar estas cifras. Para centrar la investigación en sus objetivos principales, se aplica la misma tasa de crecimiento a ambos indicadores de 2025 a 2050.

Con estas cifras, se obtiene la siguiente curva de crecimiento para el número total de vehículos y las unidades vendidas:

Tabla 1: Pronostico de la evolución de vehículos en circulación

Año	Ventas totales de vehículos	Coches en circulación
2024	80.00	1700.00
2025	81.60	1728.90
2026	83.23	1758.29
2027	84.90	1788.18
2028	86.59	1818.58
2029	88.33	1849.50
2030	90.09	1880.94
2035	99.47	2046.35
2040	109.82	2226.30
2045	121.25	2422.08
2050	133.87	2635.08

Una vez llegadas a las cifras de vehículos totales, el siguiente paso es calcular tanto cual es la presencia actual de automatización, la calidad de esta y como el coche autónomo se desarrollará hasta 2050.

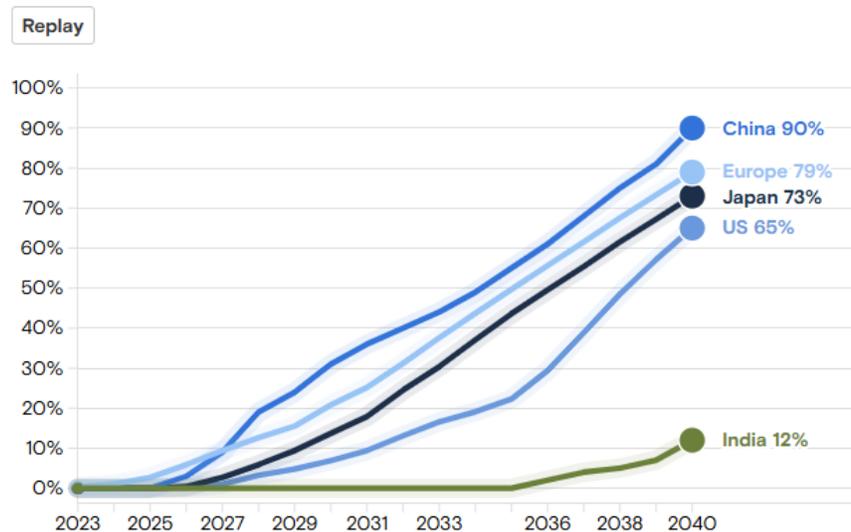
Actualmente, de los 1700 millones de vehículos que circulan en el mundo, se estiman que en torno a 54 millones de unidades tienen una automatización L2 o superior (ACEA, 2025), principalmente concentrados en los niveles bajos ya que niveles más avanzados (L4-L5) solo aparecen en taxis robóticos y pruebas de concepto (Waymo). A pesar de la relativa baja cifra, un 3.3% del total de vehículos, mejoras en la tecnología, regulaciones más favorables y precios más asequibles, se espera que en 2025 el 60% de los coches nuevos vendidos (~48 millones de los 80 millones totales) tengan al menos L2. Los sistemas más avanzados (L3-L4), como el Autopilot de Tesla o los modelos premium de Mercedes y BMW, ya representan un 6% de las ventas (~4.8 millones). Estos vehículos empiezan a mostrar ventajas claras, como mayor seguridad y eficiencia, acercándose a lo que podrían lograr los coches totalmente autónomos.

Para proyectar la evolución del panorama automotor hasta 2050, se asume, en línea con la mayoría de los estudios, que la penetración de vehículos con automatización alcanzará el 100% en la década de 2040, y que los niveles más avanzados de automatización (L4-L5) representarán entre el 50-60% del parque automotor en 2050 (Goldman Sachs, 2024), con un nivel promedio de automatización en las ventas globales sobre el nivel 3.

Con estas hipótesis, se ha elaborado la siguiente tabla que detalla las ventas de vehículos autónomos, el nivel promedio de automatización y el porcentaje que estos representan sobre el total del parque automotor global.

Autonomous vehicle penetration is forecast to grow in major economies

Key regions level 3/4/5 penetration rate



Source: Company data, Goldman Sachs Research

Goldman Sachs

Ilustración 15: Grafico de la penetracion de vehiculos autonomos de nivel 3 y superior. (Goldman Sachs, 2024)

Tabla 2: Pronostico de la penetracion y nivel de automatizacion del parque de vehiculos global 2025-2050

Año	Ventas anuales de semiautónomos	Ventas totales de vehículos	Coches en circulación	Coches nivel L1+	% Automatización	Nivel Promedio de Automatización
2024	53	80	1700	80	5%	1.70%
2025	58	82	1729	133	8%	1.74%
2026	61	83	1758	192	11%	1.78%
2027	64	85	1788	253	14%	1.82%
2028	66	87	1819	316	17%	1.86%
2029	69	88	1849	382	21%	1.90%
2030	72	90	1881	451	24%	1.94%
2035	89	99	2046	843	41%	2.16%
2040	109	110	2226	1325	60%	2.41%
2045	121	121	2422	1896	78%	2.69%
2050	134	134	2635	2527	96%	3.00%

Una vez elaborado el pronóstico de penetración y adopción del vehículo autónomo a escala global, procedemos a examinar el primer aspecto crítico relacionado con esta tecnología: la reducción de la siniestralidad vial.

6.3. Siniestralidad Vial

En este punto existe cierta controversia. Aunque a nivel mundial se asume que un vehículo completamente autónomo conseguirá reducir tanto la frecuencia de los accidentes de carretera como la gravedad de sus consecuencias, las encuestas revelan que la población, pese a reconocer las posibles mejoras, mantiene reticencias hacia la conducción autónoma. Esto se traduce en que

menos del 30 % de los participantes encuestados afirman que se sentirían seguros viajando en un coche autónomo. Por esta razón, resulta imprescindible realizar un análisis riguroso que transforme la hipótesis conceptual en una realidad respaldada por datos cuantitativos.

Para desarrollar dicho análisis se ha tomado como punto de partida el número de siniestros con víctimas en carretera que la OMS publica anualmente. En este estudio se ha optado por excluir del cálculo los accidentes menores con implicaciones únicamente materiales. No obstante, a partir del estudio principal se podrán extraer directrices para calcular, por ejemplo, el impacto económico que la reducción de accidentes podría generar.

Del Global Status Report on Road Safety 2023, la publicación más reciente de la OMS (Abdulgafoor M. Bachani, s.d.), se extrae que a nivel global en 2021 se produjeron 1,19 millones de fallecimientos en carretera, cifra a la que habría que sumar entre 20 y 50 millones de accidentes con lesiones graves. Como puede observarse, se trata de cifras inmensas que sitúan a los traumatismos viales como la octava causa de mortalidad a escala mundial (Wang, 2025). Si consideramos un promedio de 37 millones de accidentes con consecuencias severas para la salud, cabe destacar que el 92 % de estos se concentra en países en vías de desarrollo, donde tanto las deficiencias en infraestructura como las limitaciones en la normativa de circulación y el menor cumplimiento de dichas normas generan tasas desproporcionadamente elevadas en comparación con el "apenas" 8 % registrado en países desarrollados.

Con estas cifras y atendiendo a la proporción de los diferentes métodos de transporte y como de probables son de ser afectados por un accidente se obtienen las siguientes tablas

Tabla 3: Distribución de los accidentes graves entre países desarrollados y en vías de desarrollo por tipo de vehículo

Victimas Total		Fallecimientos	Lesiones<	Leiones>
Peaton	23%	273700	4600000	11500000
Motociclista	21%	249900	4200000	10500000
Ciclistas	6%	71400	1200000	3000000
Vehiculos	40%	476000	8000000	20000000
Otros	10%	119000	2000000	5000000
Total	1	1190000	20000000	50000000
Países desarrollados 8%		Fallecimientos	Lesiones<	Leiones>
Peaton	0.23	21896	368000	920000
Motociclista	0.21	19992	336000	840000
Ciclistas	0.06	5712	96000	240000
Vehiculos	0.4	38080	640000	1600000
Otros	0.1	9520	160000	400000
Total	1	7616	128000	320000
Países en desarrollo 92%		Fallecimientos	Lesiones<	Leiones>
Peaton	23%	251804	4232000	10580000
Motociclista	21%	229908	3864000	9660000
Ciclistas	6%	65688	1104000	2760000
Vehiculos	40%	437920	7360000	18400000
Otros	10%	109480	1840000	4600000
Total	1	1007216	16928000	42320000

Una vez que se saben la cantidad de accidentes actual, se procede a realizar una serie de asunciones sobre cómo el vehículo autónomo impactaría en la reducción de accidentes. En primer lugar, es necesario determinar qué tipos de siniestros serían efectivamente evitables mediante la conducción autónoma. Si bien es cierto que se podría argumentar que un coche autónomo circulando por una infraestructura inteligente debería ser capaz de prevenir accidentes más allá de los causados por error humano —por ejemplo, anticiparse al desgaste de los frenos de un camión en una pendiente pronunciada y conseguir que el propio vehículo se detenga antes de que ocurra una catástrofe—, en este análisis se han considerado únicamente los accidentes derivados del error humano como aquellos evitables por un vehículo autónomo. Este porcentaje, según la mayoría de las publicaciones especializadas, ronda el 90 % del total de accidentes, al cual se le ha aplicado una

reducción adicional del 10 % para contemplar los posibles errores que el propio sistema autónomo podría generar.

Sobre este 80 % restante, existen diversas publicaciones que analizan en qué medida los coches completamente autónomos lograrían reducir la siniestralidad. Entre los estudios más relevantes se encuentra el de la NHTSA, que proyecta una reducción del 94 %; WAYMO, que pronostica un 85% (WAYMO, 2025); y RAND Corporation, que por su parte anticipa un 80 %. Como criterio metodológico se ha optado por calcular el promedio de estas tres publicaciones, obteniendo así una tasa de reducción del 86,33 % de los errores humanos. Adicionalmente, y aplicando un enfoque conservador, se ha decidido penalizar esta reducción en un 20 % para contemplar aquellos momentos en los que el conductor pudiera decidir retomar el control del vehículo, ya que, aunque en un nivel de automatización L5 teóricamente no deberían existir estos controles manuales, resulta difícil concebir un vehículo completamente autónomo sin ningún tipo de intervención humana posible.

Una vez aplicados estos números y se ha tenido en cuenta un último factor, y es el nivel promedio de automatización del parque de vehículos en cada año. Para ello se han cogido los cálculos del apartado anterior y se ha aplicado una reducción del efecto sobre los accidentes atendiendo al nivel de automatización de forma lineal (L1-> 0%, L5-> 100%)

Se han llegado a los siguientes resultados:

Tabla 4: Resultados del análisis de reducción de accidentes

Año	Coches	% Automatización	Nivel Promedio	Sin automatización		Con Automatización Reducido		Accidentes Evitados			
				Accidentes X 1000 Vehículos		Accidentes X 1000 Vehículos		Accidentes X1000 Vehículos		En Millones de accidentes	
				Desarrollados	En vías	Desarrollados	En vías	Desarrollados	En vías	Desarrollados	En vías
2024	1700.00	5%	1.70	0.3406	30.0306	0.3381	29.8083	0.0074	0.6539	0.00	0.23
2025	1728.90	8%	1.74	0.3406	30.0306	0.3364	29.6587	0.0121	1.0701	0.00	0.39
2026	1758.29	11%	1.78	0.3406	30.0306	0.3345	29.4927	0.0172	1.5143	0.00	0.57
2027	1788.18	14%	1.82	0.3406	30.0306	0.3325	29.3181	0.0223	1.9625	0.01	0.76
2028	1818.58	17%	1.86	0.3406	30.0306	0.3304	29.1344	0.0274	2.4151	0.01	0.98
2029	1849.50	21%	1.90	0.3406	30.0306	0.3283	28.9412	0.0326	2.8723	0.01	1.21
2030	1880.94	24%	1.94	0.3406	30.0306	0.3260	28.7381	0.0378	3.3342	0.01	1.46
2035	2046.35	41%	2.16	0.3406	30.0306	0.3125	27.5555	0.0649	5.7239	0.02	3.04
2040	2226.30	60%	2.41	0.3406	30.0306	0.2954	26.0417	0.0938	8.2695	0.04	5.33
2045	2422.08	78%	2.69	0.3406	30.0306	0.2742	24.1790	0.1233	10.8751	0.06	8.50
2050	2635.08	96%	3.00	0.3406	30.0306	0.2499	22.0337	0.1511	13.3234	0.10	12.64

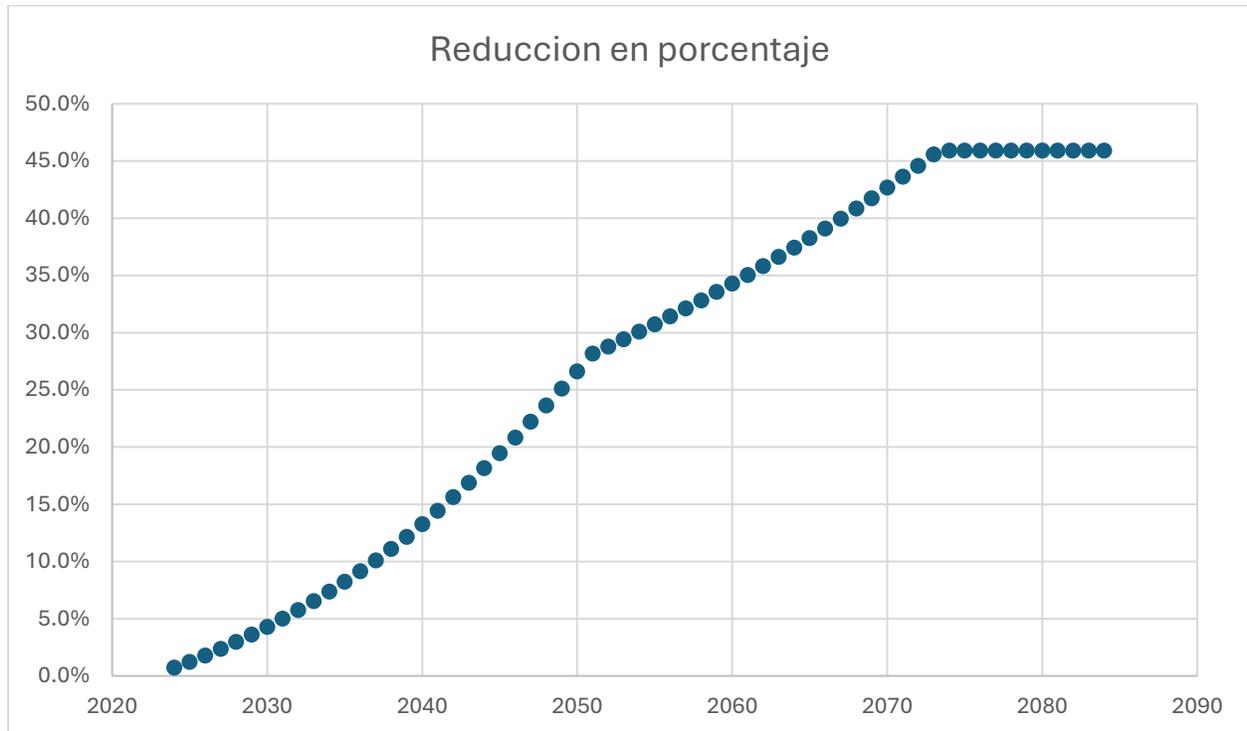
Los resultados obtenidos evidencian el impacto progresivo que la introducción de vehículos autónomos tendría sobre la reducción de la siniestralidad vial a escala global. Como se aprecia en los datos presentados, la disminución de accidentes sigue una trayectoria ascendente que se correlaciona directamente con el grado de penetración y el nivel de automatización del parque automovilístico, a pesar de que debido al aumento de la cantidad de vehículos y de trayectos, en cifras totales los accidentes aumentan.

Durante las fases iniciales de implementación, cuando predominan los vehículos con niveles de automatización L1 y L2, la reducción de accidentes resulta marginal, lo que refleja las limitaciones inherentes de estos sistemas de asistencia al conductor. Sin embargo, conforme avanza la adopción de tecnologías de mayor nivel, especialmente L4 y L5, se observa una aceleración significativa en la reducción de la siniestralidad.

Los resultados demuestran que, una vez alcanzada una penetración considerable de vehículos completamente autónomos, la reducción anual de accidentes con víctimas podría superar cifras extraordinarias, traducándose en cientos de miles de vidas salvadas y millones de lesiones evitadas.

Es importante destacar que, aunque los países desarrollados muestran una reducción absoluta menor en términos numéricos, consecuencia de su menor incidencia inicial por unas mejores condiciones de las vías, reglamentación y cumplimiento de las normas, el impacto relativo resulta igualmente significativo. Por el contrario, en los países en vías de desarrollo, donde se concentra el 92% de los fallecimientos actuales, el potencial de reducción absoluta de víctimas alcanza dimensiones que podrían transformar radicalmente las estadísticas globales de seguridad vial.

Tabla 5: Porcentaje de reducción de accidentes por año



Es interesante ampliar el estudio a lo que sería en años posteriores. Ahora se diferencian tres zonas claras, la primera de 2025 a 2050, en la cual con la introducción de más vehículos automatizados la reducción a nivel de porcentaje es prácticamente exponencial al estarse incrementando tanto porcentaje como nivel medio del vehículo autónomo. De 2050 a 2075 se la curva pierde pendiente, al tratarse el aumento de reducción únicamente dado por un mayor nivel de automatización promedio. Finalmente, a partir de 2075 se observa el nivel final de reducción derivado de los cálculos y previsiones anteriormente expuestas.

6.4. Eficiencia y emisiones

Otro de los puntos que normalmente se comentan cuando se hablan de las ventajas del coche autónomo es la mayor eficiencia a la hora de conducir y la reducción de emisiones asociada a este tipo de vehículos. Si bien es verdad que gran parte de la población asocia el coche autónomo a un coche eléctrico, en gran parte por culpa de fabricantes como Tesla que crean la percepción de que

ambos conceptos son simbióticos. En cambio, se ha analizado como una conducción realmente más eficiente, con menos parones, atascos o despistes por focalizar la atención en el navegador o algún punto del paisaje puede tener sobre las emisiones asociadas al vehículo. Al final incluir mejoras en la unidad motor del coche diluiría las verdaderas ventajas de esta conducción.

Como punto de partida se han obtenido las emisiones en toneladas equivalentes de CO₂ a nivel global de la OMS. Las toneladas equivalentes es un concepto que unifica los diferentes gases de efecto invernadero emitidos durante procesos químicos o combustión a nivel global y los estandariza utilizando como base el CO₂. A nivel global en 2023 se emitieron 57.4BtCO₂eq (United Nations, 2024) siendo el transporte uno de los principales contribuyentes a nivel global de estas emisiones, estimándose su aportación entre el 20 al 25% de las emisiones totales. Esto equivale a unas 12BtCO₂eq, que promediadas por vehículo para poder hacer una reducción en función del nivel de automatización y porcentaje de vehículos automatizados da unas emisiones aproximadamente de 4.5tCO₂eq/Año/vehículo, equivalente a lo que emiten unos 1900L de gasolina al año.

Ahora, existen dos visiones en cuanto a la reducción se refiere. Por un lado, hay ciertos sectores que argumentan que un coche completamente autónomo aumentaría el número de trayectos realizados. Esto se debería a que por ejemplo un camión completamente autónomo no tendría en principio límites horarios en su operativa, y al no contar con conductor podría convertirse en un medio aún más usado reemplazando a trenes de mercancías en ciertas zonas. Además, facilitaría la movilidad de la población con problemas motrices, y haría al vehículo privado una alternativa más accesible, desplazando tráfico actualmente cubierto por transporte público al vehículo.

No obstante, la mayoría de los estudios coinciden en que la implantación de esta tecnología traería consigo mejoras significativas en la eficiencia de conducción. Estas mejoras no solo reducirían el consumo de combustibles y las emisiones, sino que también optimizarían los tiempos de trayecto y disminuirían los costes asociados al transporte. La mayor parte de estudios citan las siguientes mejoras como los posibles mayores reductores de emisiones (Massar, 2021):

Tabla 6 Tabla con las diferentes reducciones de emisiones por mejora en la conducción autónoma (Massar, 2021)

Causa de Reducción de GEI	Resultados (%)
Perfil del conductor y flujo de tráfico	
- Calma en horas pico	0–10%
- Calma en horas no pico	0–5%
- Mejora en flujo	10–21%
- Mejora en flujo no pico	5–11%
Eco-conducción	
- Tráfico en autopista congestionada	10–20%
- Flujo libre	~0%
- Tráfico congestionado en ciudad	5–10%
- Bajo diferentes niveles	8–23%
Señalización de tráfico eficiente	
- Optimización de semáforos	1.8–2%
Comunicación V2i/i2v	2–6%
Evitar colisiones	
- Reducción en ciudad	0–0.95%

Causa de Reducción de GEI	Resultados (%)
- Reducción máxima	0–1.9%
Platooning (agrupamiento)	
- Durante horas pico	0–12.5%
- Durante horas no pico	12.5–25%
- Máximo en horas no pico	22.5–27.5%
Reducción por redimensionamiento	~45–50%
Búsqueda de estacionamiento	2–11%
Aumento de viajes compartidos	~12%
Viajes más rápidos	
- Durante horas pico	0–10%
- Durante horas no pico	10–40%

Como se observa en la tabla anterior que representa una agrupación de estudios realizado en 2021, se han estudiado prácticamente todas las casuísticas en las que una conducción más inteligente sería efectiva reduciendo emisiones GHG. Todas ellas son interesantes, pero es cierto que se necesitaría de un porcentaje alto de automatización, así como una comunicación eficiente entre vehículos, infraestructura y conductores para que sus efectos sean realmente notables. Un ejemplo de ello es el platooning, es decir el agrupamiento de vehículos que causa retenciones por aumento del tráfico. Se asume en este caso que cada vehículo es capaz de comunicarse con el resto para

optimizar no solo su ruta, si no el conjunto de rutas que potencialmente podrían causar más tráfico en una zona determinada.

Es por ello por lo que se ha decidido no utilizar la mayoría de los inputs, para no sobredimensionar el alcance de las mejoras y se han seleccionado los siguientes puntos como principales reductores de emisiones de gases de efecto invernadero. Además, para tener en cuenta tanto escenarios más optimistas como conservadores se han realizados dos curvas de reducción:

Tabla 7: Tabla con las reducciones de emisiones utilizadas en el estudio

Reducción de emisiones	Conservador	Optimista
Por atascos: https://www.mdpi.com/2071-1050/14/11/6910	15%	30%
Rutas optimas: https://pmc.ncbi.nlm.nih.gov/articles/PMC8197118/	5%	10%
Eco Driving	15%	35%
Menos búsqueda de aparcamiento	3%	5%
Aumento del uso del vehículo	30%	30%
Total reducción por vehículo autónomo	13%	49%

Como se observa la mayor parte de la reducción viene dada por el llamado Eco driving, en el que el vehículo optimiza la aceleración, el manejo y las frenadas que se hacen en un trayecto para consumir lo menos posible, así como una notable reducción de atascos por dos factores principales. Primero sería la reducción de accidentes comentada anteriormente. El primero es la reducción de accidentes, ya discutida previamente. El segundo, y no menos relevante, es la implementación de una conducción inteligente que mitiga significativamente el conocido efecto acordeón o fantasma. Este fenómeno ocurre cuando pequeñas irregularidades en la conducción, como frenadas bruscas o aceleraciones repentinas, desencadenan un efecto mariposa que se propaga entre los conductores, creando atascos aparentemente "de la nada". La conducción autónoma tiene el potencial de reducir este efecto al mantener un flujo de tráfico más uniforme y eficiente.

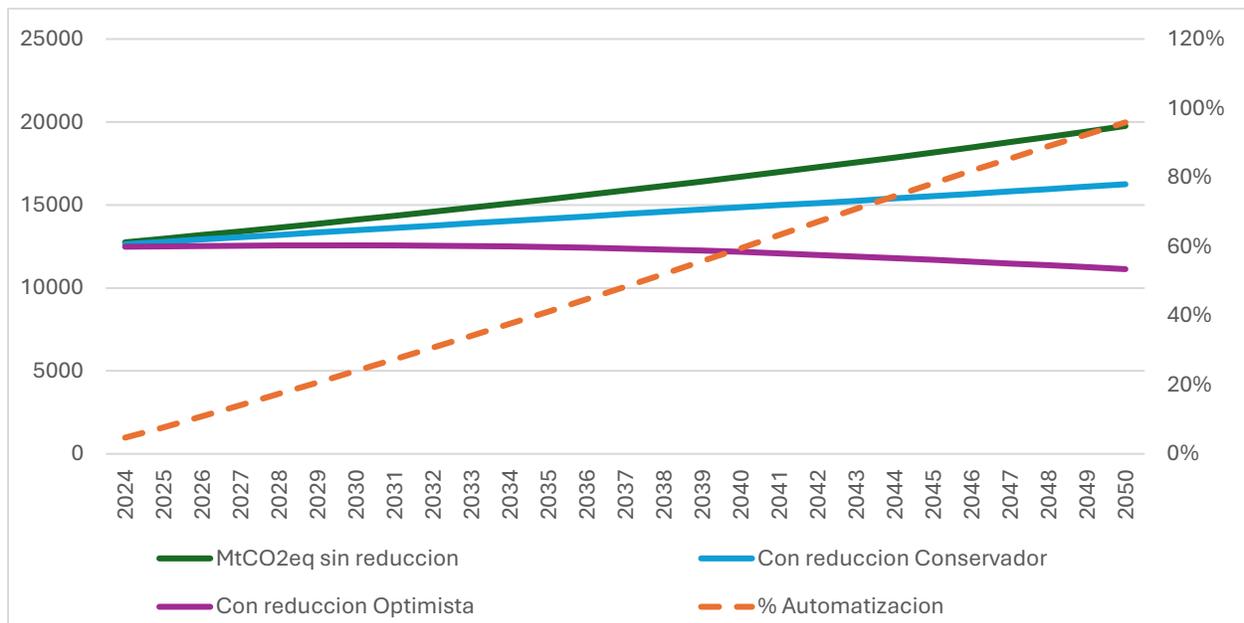


Ilustración 16: Representación del efecto acordeón (Etrasa, 2019)

Finalmente, para no despreciar el aumento de desplazamientos que el vehículo autónomo proporcionara se ha decidido optar por aplicar una penalización del 30% sobre el total de emisiones en ambos casos.

Con estos números se llegan a las siguientes curvas de reducción de emisiones:

Tabla 8: Gráfico de la reducción de emisiones pronosticada



Como se observa la mejora es notable en ambos casos, tanto en el conservador como como en el más optimista. En el más optimista es muy interesante observar como a pesar de que las emisiones aumentarían en los primeros años debido a un mayor número de vehículos en circulación, con aproximadamente una penetración de vehículos autónomos sobre el 40% las emisiones totales del transporte caerían por debajo de niveles actuales, situándose en 2050 con una automatización completa, aunque de nivel 3 promedio, sobre las 10GtCO₂eq. Este número, comparado con la proyección realizada de emisiones totales si no se avanza con la automatización, representa un corte sobre el total de emisiones de casi un 50%, o lo que es igual, un ahorro de 30.000 TWh (MITECO, 2016) equivalentes de gasoil o las emisiones asociadas a la combustión de una producción equivalente a 190 días actuales (a pesar de tratarse de un porcentaje alto es correcto ya que se asume que la producción continuara aumentando para abastecer la demanda)

Con los resultados obtenidos de ambos estudios, tanto el de impacto sobre la seguridad en carretera, como el impacto medioambiental del vehículo autónomo es evidente. En ambos casos una gran implantación de este tipo de movilidad resulta en mejoras que la sociedad ha de buscar. Es importante este análisis, ya que, a pesar de desarrollar un marco ético posteriormente para esta medida, es muy importante el análisis de si realmente es una medida correcta. Además, conocer el impacto que puede tener sobre la sociedad hará que la población cuente con un mayor grado de confianza a la hora de la implantación. Como comentado en el estado del arte, el principal problema al que se enfrenta la conducción autónoma es la reticencia de la sociedad a adoptar este progreso. Un buen estudio, que sea comprensible y al que la población tenga acceso es necesario para que interioricen las ventajas del vehículo autónomo.

Finalmente, además es un buen indicador de porque el coche autónomo en si es una decisión ética. La adopción de este implicara mejoras sobre tres de los pilares en los que se fundamenta la ética. El primero es el cuidado y respeto al prójimo. Como resaltan los resultados el vehículo autónomo es por diseño más seguro que un conductor humano, por lo que al adoptarlo se le otorga valor a la vida, tanto propia, como de la sociedad que rodea. El siguiente es la libertad, y es que la introducción de esta forma de transporte conlleva una mayor accesibilidad a desplazamientos a gran parte de la población que actualmente no pueden. Y por último concepto la responsabilidad. El coche autónomo es responsable con la sociedad y el medioambiente, siendo una de las

soluciones, junto a la movilidad eléctrica, que más respetan el entorno en el que están, reduciendo gastos innecesarios, aumentando la seguridad y devolviendo tiempo a otras cosas más importantes.

7. Modelo Ético

7.1. Introducción

Una vez descubierto porque es necesario y ético la introducción del vehículo toca revisar como debe de ser esta implantación desde un punto de vista ético. Como desarrollado en el estado del arte, tanto a nivel de hardware, software y ética los avances por conseguir un modelo y “cerebro” para el vehículo perfecto se están llevando a cabo. En este modelo sin embargo se van a desarrollar los aspectos en los que la ética tenga una mayor importancia, que serían tres: la arquitectura de software a implementar para poder conseguir los resultados más éticos a la vez que comprensibles; el modelo ético detrás del software, que reglas se han de implantar o cuales son los puntos que pueden o deben variar entre fabricantes; y finalmente el apartado regulatorio, sobre quien deberían recaer las responsabilidades y que debería ser limitado.

7.2. Arquitectura de software.

El primer paso de la implementación es conocer que soluciones se están tomando ahora y cuál puede ser la más adecuada para el futuro. Como comentado en el estado del arte actualmente existen 3 tipos de arquitecturas sobre las que se sustenta el código de todo vehículo autónomo.

La primera son las reglas preprogramadas. Un sistema de reglas preprogramadas determina el comportamiento del vehículo con sentencias «si X, entonces Y», tomando datos de los sensores (velocidad, distancia, ángulos, aceleraciones) y actuando sobre acelerador, freno y dirección. El control de crucero y el asistente de mantenimiento de carril son ejemplos típicos. Las ventajas en este esquema son fáciles de identificar.

- Baja demanda computacional: el algoritmo es determinista, utiliza operaciones aritméticas simples y funciona en microcontroladores de baja potencia sin requerir refrigeración especial. Además, al no requerir de un entrenamiento previo, el desarrollo se simplifica.
- Trazabilidad completa: cada decisión se vincula a una línea de código, de modo que la secuencia exacta puede reproducirse y validarse por pruebas de caja blanca.

- Integración normativa directa: límites de velocidad o distancias mínimas se traducen a umbrales numéricos; el regulador puede revisar el código y confirmar su conformidad sin interpretar modelos estadísticos.
- Familiaridad para el usuario: el conductor ya conoce estos asistentes, lo que facilita la aceptación y reduce la curva de aprendizaje.

Sin embargo, también aparecen una serie de inconvenientes con difícil solución entre los que se encuentran:

- Rigidez ante la variabilidad: el número de estados posibles crece exponencialmente y la introducción de casos raros exige reglas adicionales que complican el sistema.
- Conflictos lógicos: una regla de velocidad máxima puede contradecir la necesidad de acelerar para evitar colisiones, obligando a definir jerarquías y excepciones.
- Falta de aprendizaje: los umbrales no se ajustan con la experiencia; cualquier mejora implica reprogramar y volver a ensayar todo el conjunto de pruebas.
- Actualizaciones costosas: cada cambio normativo obliga a reescribir, certificar y desplegar nuevo software en toda la flota, con impacto logístico y tiempo de inactividad.

Es por estos factores que la utilización única de reglas no es la mejor solución al software que debe gobernar el coche autónomo. Ante este problema aparece el segundo tipo de arquitectura, las redes neuronales y sistemas de aprendizaje automático.

Las redes neuronales son sistemas de procesamiento inspirados en el funcionamiento del cerebro humano. Estas redes están compuestas por capas de nodos (neuronas), donde cada nodo realiza cálculos matemáticos y transmite sus resultados a la siguiente capa. A través de un proceso de entrenamiento, las redes ajustan los pesos de estas conexiones para aprender patrones a partir de grandes volúmenes de datos. En el contexto de los vehículos autónomos, las redes neuronales se utilizan para tareas como la detección de objetos, la clasificación de señales de tráfico, la predicción de trayectorias y la toma de decisiones complejas en tiempo real. Esta capacidad para generalizar y aprender patrones complejos les permite abordar escenarios complejos, caóticos y

dinámicos, que como discutido previamente es uno de los principales problemas de las reglas preprogramadas. Entre las principales ventajas se encuentran:

- **Capacidad de adaptación:** las redes neuronales pueden aprender de datos históricos y ajustarse a patrones complejos, lo que les permite manejar situaciones imprevistas o no programadas, como cambios en las condiciones climáticas o comportamientos erráticos de otros conductores.
- **Escalabilidad:** a medida que se recopilan más datos, el sistema puede ser reentrenado para mejorar su rendimiento sin necesidad de rediseñar la arquitectura subyacente.
- **Robustez en entornos dinámicos:** su capacidad para procesar múltiples entradas simultáneamente las hace ideales para la toma de decisiones en tiempo real en entornos complejos como el tráfico urbano.
- **Reducción de la intervención humana:** el entrenamiento automatizado minimiza la necesidad de programar manualmente cada escenario posible, lo que acelera el desarrollo.

Pero no es una solución perfecta, y es que al menos en el apartado de la ética, estas redes neuronales se encuentran con inconvenientes importantes a tratar:

- **Falta de transparencia y trazabilidad:** Las redes neuronales son esencialmente sistemas de "caja negra", lo que significa que, aunque se pueda observar su entrada y salida, el proceso intermedio que lleva a una decisión no es fácilmente interpretable. Se ha comentado previamente, pero hace que esencialmente la asignación de responsabilidades se complique, así como la confianza que la sociedad tiene en estos sistemas. Si bien es verdad que los LLM como ChatGPT están acercando esta tecnología y haciéndola accesible, aún existe gran desconfianza ante las decisiones autónomas de estos sistemas.
- **Sesgos en los datos de entrenamiento:** Las redes neuronales aprenden a partir de los datos con los que se entrenan, lo que significa que cualquier sesgo presente en esos datos se reflejará en su comportamiento. Por ejemplo, si los datos de entrenamiento no incluyen suficientes escenarios con peatones de ciertas características (edad, complexión, vestimenta), el sistema podría fallar al reconocerlos. Esto plantea un problema ético, ya

que podría discriminar a ciertos grupos sociales o demográficos, afectando su seguridad. También surge gracias a esto el debate de si debieran tener diferentes sets de entrenamiento atendiendo a la localización.

Ante este dilema entre reglas rígidas, comprensibles y auditables, o redes neuronales adaptativas, eficientes y robustas, aparece una tercera arquitectura como unión simbiótica de ambas filosofías de software.

Esta arquitectura se basa en dos capas de código. La primera es la capa normativa de reglas preprogramadas. Estas fijan un perímetro operacional al vehículo, son como las barandillas que guían el camino. El regulador sería capaz de leer y auditar, y deberían ser las que fijen los límites de seguridad del vehículo. No deberían en principio poder ser traspasadas por norma general, pero en situaciones donde hubiese contradicciones se tiene que fijar aun así un orden de prioridad entre ellas. El ejemplo claro es que evitar colisiones como regla preprogramada tiene que contar con un peso mayor que por ejemplo respetar perfectamente el límite de velocidad o un ceda el paso. La segunda capa es la capa adaptativa, sustentada en redes neuronales, opera dentro de ese perímetro para interpretar cámaras, radares y Lidar, anticipar la trayectoria de los demás usuarios y seleccionar la maniobra más suave y eficiente.

Las ventajas principales son las técnicas, y es que es una unión simbiótica en la que a nivel de arquitectura de software ambas son soluciones complementarias. El resultado en este caso no es una suma, sino el producto de ambas filosofías. La red neuronal aporta la flexibilidad, pero gracias a las reglas no se pierde la trazabilidad necesaria para el regulador. Como solución técnica es la mejor disponible actualmente, pero éticamente siguen apareciendo los siguientes inconvenientes:

- Persistencia de la “caja negra”: aunque la regla final sea visible, la génesis de la recomendación neuronal sigue siendo opaca y complica la atribución moral en caso de daño.
- Sesgos residuales: los filtros reglados limitan, pero no eliminan, la posibilidad de que un modelo reproduzca patrones discriminatorios presentes en los datos.

- Dilemas de prioridad: la jerarquía entre la capa normativa y la capa neuronal debe decidir de antemano qué vidas o bienes proteger en conflictos extremos, un terreno aún sin consenso social pleno.

En definitiva, aunque no existe una solución única y definitiva, los problemas expuestos pueden mitigarse con el siguiente elemento clave de la lógica del vehículo autónomo: el marco ético. Este marco determina cómo se comportará la arquitectura de software cuando se enfrente a dilemas morales en carretera. Diseñado de forma rigurosa, permitirá gestionar sesgos, resolver conflictos de prioridad y, al mismo tiempo, reducir el carácter de “caja negra” que hoy lastra la confianza en estos sistemas.

7.3. Marco ético

Si bien la arquitectura y estructura del software es importante para un coche eléctrico a nivel de ética lo realmente importante es el marco ético en el que el vehículo trabaje. Cuando nos referimos al marco ético, se habla de las diferentes formas de entender que es lo moral y justo. Lo primero es saber que no existe, o es imposible asumir una verdad universal que rijan absolutamente todas las situaciones. Diferentes formas de pensar sobre cómo comportarse, o que está bien o mal, dan lugar a soluciones a problemas completamente diferentes. Si bien es verdad que nunca se conseguirá un sistema perfecto que cualifique para todas las formas de pensar, todas las situaciones y culturas, en la filosofía clásica se han desarrollado tres estructuras principales sobre las cuales obtener respuestas éticas. Como desarrollado en el estado del arte, estas tres corrientes son la ética de la virtud o aristotélica, el Deontologismo o ética del deber, y Utilitarismo. Estas tres ofrecen criterios diferentes a la hora de decidir que es ético o que deja de serlo.

La ética aristotélica se fundamenta sobre la idea de que lo importante no es la acción aislada, sino el carácter de quien actúa. Se basa en cultivar virtudes como la prudencia, la justicia o la valentía, actuando en la mayor medida en el denominado término medio para evitar exceso o defecto en las decisiones. Si bien es verdad que a nivel social puede ser una buena solución, para la conducción autónoma es un criterio insuficiente. No se puede cuantificar el “carácter” o intención de un vehículo autónomo, y dejar a un código de aprendizaje automático a que decida cuáles son sus intenciones parece una idea sacada de ciencia ficción. Sería como interpretar que el coche tiene

conciencia y la utiliza de la mejor manera, por lo que no es una solución adecuada de cara al marco ético que los vehículos autónomos deberían portar.

La segunda estructura, la ética del deber, se basa en la rectitud moral ante el cumplimiento de una serie de deberes o reglas, sin aportar valor ético al resultado. Esta ética, en la cual su principal exponente es Immanuel Kant, exige actuar conforme a unos principios o máximas que puedan llegar a convertirse en ley universal, siempre utilizando al humano como un fin en sí mismo y no como un medio. Si es cierto que, de cara a la arquitectura de software, el deontologismo es una ética que se adapta lo suficientemente bien como para ser utilizada. El problema viene sobre como programar estas reglas. Como comentado en la arquitectura las reglas preprogramadas están para cumplirse en el 99% de las casuísticas usuales, donde las reglas de conducción sean siempre la mejor solución. Pero el problema aparece cuando rompiendo levemente alguna regla, como saltarse una línea continua, minimizaría las consecuencias de un accidente. Es por ello que puede ser una solución para el coche autónomo, pero implementar un sistema basado completamente en el deontologismo no soluciona ninguno de los problemas prevalentes en la arquitectura.

Finalmente, el utilitarismo tiene como idea central para decidir sobre si una acción es correcta moralmente, si las consecuencias de esta son las mejores posibles. Este enfoque lo formalizaron Jeremy Bentham y John Stuart Mill con el criterio de “Suma de felicidad”. En este enfoque se le asigna a cada consecuencia un valor “objetivo”, haciendo que se deba actuar en torno a las acciones que conlleven un mayor valor. Este enfoque a priori es compatible con el coche autónomo, ya que aporta una visión cuantitativa y no cualitativa a las acciones.

Para implementar un sistema utilitarista sería necesario crear una matriz o tabla con valores asignados a cada posible resultado de la conducción. Este es un enfoque que ya existe en las redes neuronales, ya que actualmente se pueden ajustar cada una de las capas y nodos que las componen para que las varianzas y covarianzas de diferentes entradas tengan un valor mayor o menor, haciendo que ciertas acciones se vean favorecidas con respecto a otras. Es un criterio que soluciona una gran cantidad de los problemas de la arquitectura:

- Criterio de caja negra: es impracticable incorporar un sistema de aprendizaje autónomo y pretender eliminar al completo la famosa caja negra. Pero es verdad que la matriz de

consecuencias y valores otorga una visión clara, al menos, de los criterios que la maquina tiene que seguir. Esta matriz podría ser publica y ser regulada por el gobierno o cuerpo regulador correspondiente. Es sumamente útil, ya que atendiendo a cada vehículo la matriz es adaptable, al igual que por entorno o país.

- Los sesgos residuales podrían ser regulados con una mayor certeza. La matriz, al ser sumamente ajustable sería capaz de crear un entorno igualitario, en el que el valor de por ejemplo atropellar a un peatón mayor o menor tenga el mismo valor, o el que la sociedad en su conjunto dicte por democracia como justo.
- Dilema de prioridad. Este es el dilema con una solución más elegante. Uno de los nodos de la matriz de pesos podría ser el incumplir alguna de las reglas preprogramadas. El sistema luego podría optimizar la situación de forma que si alguna de las reglas es sobrepasada esa solución cuente con una penalización importante, pero que no sea completamente inquebrantable, dando lugar a que en situaciones extremas estas reglas pueden ser sobrescritas.

Es por estos motivos que el utilitarismo se presenta como la mejor estructura ética frente a un problema en el que es una maquina quien tiene que seguir la manera de actuar más ética. Además, es una visión respaldada por ciertos estudios. Por ejemplo, en ciertos servicios de urgencias, se han entrenado redes neuronales capaces de hacer triaje. Estas funcionan intentando maximizar el denominado “vidas-año-ganadas”, priorizando a los pacientes que podrían beneficiarse de una mayor mejoría.

7.3.1. Especificaciones de la ética utilitarista

Una vez estudiado la estructura ética más favorecedora de cara a un sistema basado en software habría que adentrarse en la estructura que va a ser implementada. Al contrario del ejemplo del hospital en el que siempre la situación del paciente es mejorable, la situación más favorable en un coche autónomo sería una conducción normal, sin ningún incidente extraordinario. Es por ello por lo que el estándar a conseguir es el mismo punto que en el de partida, por lo que toda decisión límite que tome la máquina, o toda consecuencia, fuera de una conducción estándar debería estar

penalizada. Esta manera de entender el problema de la implementación se da de bruces con la ética utilitarista clásica de maximizar el bien común, es por ello por lo que la ética utilitarista correcta debería ser el utilitarismo negativo. Esta corriente, formulada por Karl Popper en su obra principal “The Open society and its Enemies” (Popper, 1945), cambia el enfoque de la ética utilitarista clásica al abordar el problema desde el punto de vista de que lo importante es reducir el sufrimiento de la sociedad. Básicamente indica que para considerar una acción como correcta tiene que seguir un “Principio Rector” que evite o mitigue el sufrimiento, aunque no aumente los placeres.

El enfoque utilitarista negativo se adapta con facilidad al dilema operativo del vehículo autónomo. En la práctica, la matriz de decisión se construye asignando un coste de sufrimiento a cada combinación de incidente y víctima potencial: colisión frontal, lateral o por alcance; tipo de vehículo implicado (motocicleta, turismo, vehículo pesado); presencia de peatones u obstáculos fijos; e incluso infracciones de normas de tráfico necesarias para evitar daños mayores. Cada casilla recibe un valor ponderado que refleja la severidad esperada de las lesiones y el número de afectados.

Cuando se desencadena una situación crítica, la red neuronal estima, para cada maniobra factible, la probabilidad de encajar en cada casilla de la matriz. Multiplicando probabilidades por costes obtiene un “sufrimiento esperado” para cada opción y selecciona la que minimiza dicho indicador. El esquema admite refinamientos: si un alcance trasero a una motocicleta genera lesiones mucho más graves que a un turismo, el peso asignado a esa casilla se incrementa; de igual modo, la protección de usuarios vulnerables puede integrarse elevando sus coeficientes de daño.

Este diseño presenta dos ventajas claras. Primero, convierte el principio moral abstracto (minimizar el daño) en una función de coste explícita, legible y auditable. Segundo, permite ajustes transparentes a medida que se disponga de mejores datos clínicos o estadísticos. La dificultad reside en consensuar públicamente los pesos asignados a cada escenario y en garantizar que el modelo probabilístico sea lo bastante fiable para que la elección mínima de sufrimiento no se base en estimaciones sesgadas o inciertas.

Estas grandes decisiones no deben de ser tomadas a la ligera. De hecho, decisiones de un calado como las expuestas no deberían estar supeditadas a los intereses de empresas privadas, que

lógicamente buscaran poner sus productos en una posición de superioridad frente a la competencia. Esto podría dar lugar a una situación similar a la que ocurre en teoría de juegos con el dilema del prisionero. Esto haría que por querer acabar en una situación ventajosa desde un plano individual, se llegue a una situación poco óptima a nivel global. Es por ello que estas decisiones deben ser tomadas en consenso por la sociedad, haciendo y dotando de importancia al tercer pilar del marco ético, la regulación que lo acompaña.

7.4. Marco legislativo

Todo lo anterior, la arquitectura del software y el marco ético solo cobra sentido si existe una ley que lo ate todo y pueda cumplirse en la práctica. Sin un reglamento claro, cada empresa seguiría su propio criterio, el ciudadano no sabría a qué atenerse y los jueces no tendrían herramientas para decidir en caso de accidente. Por eso el tercer pilar es un marco normativo capaz de transformar los principios morales en reglas uniformes, verificables y, sobre todo, comprensibles para cualquiera.

El primer escollo que debe salvar la ley es la dispersión de criterios éticos entre marcas y países. Hoy un coche puede valorar de un modo el riesgo para un peatón y otro coche de la acera contraria, fabricado en otro continente, puede valorarlo de forma distinta. Para evitar esta lotería, la normativa tendría que obligar a trabajar con una misma matriz de impacto y esa matriz debería definirse de manera pública y colegiada. Lo lógico es que la apruebe un organismo supranacional, como ya ocurre con los límites de emisiones o con las pruebas EuroNCAP, de modo que un turismo, una motocicleta o un camión usen en España, Alemania o Italia la misma escala de valores. Si un fabricante quiere ser aún más estricto y proteger por encima de la norma, podrá hacerlo, pero nunca por debajo.

Ahora bien, homogeneizar la matriz no basta: hay que garantizar que el coche la está utilizando de verdad. El regulador no necesita todo el código fuente ya que sería inviable y pondría en riesgo secretos industriales, pero sí tiene que acceder a tres cosas muy concretas: la tabla de pesos tal cual se ejecuta en el vehículo, el procedimiento con el que se actualiza ese software “por el aire” y un registro cifrado de lo que decide el coche en cada situación crítica. Esa especie de caja negra digital permitirá reconstruir los hechos ante un tribunal sin que el sistema sea, a efectos prácticos, una

caja negra moral. Si en un siniestro se demuestra que el programa operaba con otra matriz o que la versión no estaba validada, la responsabilidad del fabricante será directa.

Precisamente la cuestión de la responsabilidad es la gran piedra de toque mientras no exista autonomía plena. Mientras el nivel de automatización se quede por debajo del nivel 5, es decir, mientras el coche reclame al conductor que tome el control en ciertos momentos, la ley debería mantener la culpa principal en el ser humano. No obstante, el fabricante seguirá respondiendo por fallos de diseño, por hardware defectuoso o por cualquier divergencia entre la matriz que declaró y la que realmente corría en el vehículo. Para que no haya zonas grises, el propio sistema debe registrar al milisegundo cuándo pidió la ayuda del conductor y cuándo este la aceptó o la ignoró.

Por último, el reglamento ha de crear un mecanismo ágil de actualización. Cada vez que un fabricante cambie su red neuronal o quiera ajustar un peso de la matriz que altere el reparto de riesgos, deberá notificarlo y someterlo a un ensayo acelerado –simulaciones masivas, pruebas cerradas– que confirme que dicho cambio no empeora la seguridad de ningún grupo de usuarios. Una agencia técnica supranacional, similar a la que existe en aviación, podría encargarse de esa validación continua.

La única incógnita que se queda es el momento en que los vehículos logren la autonomía completa y las cifras de siniestralidad se reduzcan drásticamente. Qué ocurrirá entonces con el derecho a conducir de forma manual es una cuestión que la norma deja abierta: será necesario un debate específico, apoyado en datos y participación pública, para decidir si la conducción humana seguirá permitiéndose en vías públicas o quedará restringida a entornos controlados. Ese trabajo queda pendiente y requerirá otra fase de desarrollo legislativo.

7.5. Solución y cuestiones existentes:

En conjunto, el marco ético propuesto articula tres pilares que lo sustentan: la arquitectura híbrida del software, el utilitarismo negativo como brújula moral y la regulación como guardarraíl externo. La capa de reglas preprogramadas fija los límites inquebrantables (velocidades máximas, distancias de seguridad, prioridad de usuarios vulnerables); dentro de ese perímetro, la red neuronal optimiza en tiempo real la maniobra que minimiza el daño previsto según la matriz de

impacto pública. Esa matriz, a su vez, se homologa y revisa de forma periódica por el regulador, que exige registros cifrados de cada decisión para poder auditarla. El resultado es un circuito cerrado: la ley define los valores, el software los aplica y la “caja negra” confirma que, en la práctica, se han respetado.

Quedan, sin embargo, áreas fértiles para seguir trabajando. Habrá que pulir métodos de explicación automática que traduzcan en lenguaje sencillo por qué el coche eligió una maniobra concreta; mejorar los sistemas de simulación acelerada que verifiquen cada actualización antes de su despliegue masivo; y, sobre todo, abrir un proceso participativo que decida, con datos y deliberación pública, cuál debe ser el lugar del conductor humano cuando la autonomía sea plena. Además, la matriz de impacto deberá evolucionar con nueva evidencia clínica (por ejemplo, sobre lesiones en peatones mayores o ciclistas) y con cambios en la infraestructura urbana. Este marco ético es, por tanto, un punto de partida robusto, pero no un puerto de llegada: está diseñado para aprender, adaptarse y seguir incorporando la voz de la sociedad a medida que la tecnología y nuestras expectativas avancen.

8. Conclusiones

8.1. Conclusiones

A lo largo de este Trabajo Fin de Máster se ha llevado a cabo un recorrido exhaustivo que parte de los primeros prototipos radiocontrolados de Francis Houdina, continúa con los hitos académicos del Navlab de Carnegie Mellon y los DARPA Grand Challenges, y desemboca en las actuales flotas comerciales de Waymo, Cruise o Baidu.

Este contexto histórico ha permitido justificar la pertinencia del estudio: comprender cómo la maduración simultánea de sensores, algoritmos IA y conectividad ha situado al vehículo autónomo en la antesala de una adopción masiva, pero también en el centro de un debate multidisciplinar sin precedentes.

8.2. Modelo de los beneficios del coche autónomo

La tesis confirma que la seguridad vial es uno de los principales motores de la autonomía: hoy, el factor humano interviene en más del 90 % de los accidentes de tráfico. Este dato, reiterado por organismos internacionales, dota de legitimidad ética a la investigación: salvar vidas se convierte en el primer principio de diseño.

Además, el trabajo cuantifica el ahorro potencial en costes sanitarios y productividad que conllevaría una caída drástica de la siniestralidad, así como la reducción de emisiones gracias a una conducción más suave por la eliminación de arranques y frenadas bruscas; una conducción más eficiente desde el punto de vista de velocidades y rutas; y la eliminación de efectos propios de la conducción, como el efecto acordeón. Tales beneficios se alinean con la justicia distributiva, porque protegen de forma prioritaria a los colectivos más vulnerables como peatones, ciclistas y personas con movilidad reducida o en situaciones desfavorables, que hoy sufren un riesgo desproporcionado en la vía pública, visto por ejemplo en las grandes diferencias a nivel de siniestralidad entre países desarrollados y en vías de desarrollo. Bajo esta óptica, la introducción del vehículo autónomo no es solo un avance técnico, sino una obligación moral coherente con los Objetivos de Desarrollo Sostenible y con el principio bioético de beneficencia.

8.3. Modelo ético

El modelo ético aquí defendido articula un triángulo de coherencia entre técnica, moral y derecho. En primer lugar, la arquitectura híbrida; reglas preprogramadas como barandillas y redes neuronales como agente adaptativo; otorga al sistema la trazabilidad que exige el regulador con la elasticidad operativa que demanda el tráfico real, captando así lo mejor de cada una de las filosofías de software actuales. En segundo lugar, el utilitarismo negativo aporta un criterio cuantificable y público: minimizar el sufrimiento. Este principio se plasma en una matriz de impacto cuyos pesos se pueden auditar, ajustar y comparar entre fabricantes y jurisdicciones sin exponer secretos industriales. En tercer lugar, un marco legislativo supranacional convierte la brújula moral en obligación jurídica: homologa la matriz, exige registros cifrados de cada maniobra crítica y fija un procedimiento de actualización continua que impide la dispersión de estándares.

Las ventajas de este planteamiento son múltiples. Técnicamente, evita la rigidez de los sistemas puramente reglados y la opacidad absoluta de las redes neuronales aisladas. Éticamente, hace explícitos los valores que guían la acción del vehículo y permite revisarlos conforme evolucionen los datos clínicos o la sensibilidad social. Regulatoriamente, impone una base común que disipa la “lotería ética” entre marcas y países y otorga a jueces y ciudadanía herramientas objetivas para dirimir responsabilidades.

8.4. Puntos abiertos

Quedan, no obstante, rutas abiertas que merecen atención en futuras investigaciones:

- Mecanismos de explicabilidad que traduzcan el cálculo de “sufrimiento esperado” a un lenguaje comprensible para usuarios y tribunales.
- Sistemas de simulación acelerada que certifiquen cada actualización de la red neuronal y de la matriz antes de su despliegue masivo.
- Procesos participativos estables que revisen, con datos y deliberación pública, los pesos asignados a cada tipo de daño y decidan el lugar de la conducción humana cuando la autonomía sea plena.

- Refinamiento continuo de la evidencia clínica y estadística para ajustar los coeficientes de daño, en especial respecto a usuarios vulnerables como peatones mayores o ciclistas.
- El dilema de si los humanos, llegado el punto de una automatización total de máximo nivel, deberíamos estar autorizados a conducir fuera de entornos controlados como circuitos.

9. Bibliografía

- Abdulgafoor M. Bachani, P. M. (s.f.). *WHO Global Status Report on Road*. Obtenido de Johns Hopkins : https://www.aprso.org/sites/default/files/2025-05/WHO%20Global%20Status%20Report%20on%20Road%20Safety%202023_Abdul%20Bachani.pdf
- ACEA. (2025). *Economic and Market Report*.
- Amodei, D. O. (2016). *Concrete problems in AI safety*. Obtenido de <https://arxiv.org/abs/1606.06565>
- Anderson, J. M. (2016). *Autonomous Vehicle Technology: A Guide for Policymakers*. RAND Corporation.
- Badoni, V. (2021). *A Sense of Responsibility: Lidar Sensor Makers Build on NVIDIA DRIVE*. Obtenido de NVIDIA Blogs: <https://blogs.nvidia.com/blog/lidar-sensor-nvidia-drive/>
- Bimbraw, K. (January de 2015). *Autonomous Cars: Past, Present and Future - A Review of the Developments in the Last Century, the Present Scenario and the Expected Future of Autonomous Vehicle Technology*. ReserachGate.
- Burns, L. D. (2013). *ransforming Personal Mobility*. The Earth Institute, Columbia University.
- Chen, L. S. (s.f.). *Learning affordance for direct perception in autonomous driving*. *IEEE International Conference on Computer Vision (ICCV)*, 2722-2730.
- Chen, T. L. (2021). *High-performance computing platforms for autonomous driving: A survey*. Obtenido de *IEEE Transactions on Intelligent Transportation Systems*, 22(3), 1287-1303.: <https://doi.org/10.1109/TITS.2020.2981234>
- Etrasa. (2019). *¿Sabes qué es el “efecto acordeón” y cómo evitarlo?* Obtenido de <https://www.etrasa.com/sabes-que-es-el-efecto-acordeon-y-como-evitarlo/>
- Goldman Sachs. (2024). *Partially autonomous cars forecast to comprise 10% of new vehicle sales by 2030*. Obtenido de Goldman Sachs:

- <https://www.goldmansachs.com/insights/articles/partially-autonomous-cars-forecast-to-comprise-10-percent-of-new-vehicle-sales-by-2030>
- Green Car Congress . (2017). *New Audi A8 debuting Level 3 autonomous AI traffic jam pilot; parking and remote garage pilots; zFAS controller*. Obtenido de <https://www.greencarcongress.com/2017/07/20170712-a8.html>
- Koopman, P. &. (2017). *Challenges in autonomous vehicle testing and validation*. . Obtenido de SAE International Journal of Transportation Safety, 5(1), 15-24. : <https://doi.org/10.4271/2017-01-0128>
- Li, Y. L. (2021). *Recent advances in steer-by-wire systems for autonomous vehicles: A review*. Obtenido de IEEE Access, 9, 123456-123470.: <https://doi.org/10.1109/ACCESS.2021.3051234>
- Massar, M. R.-I. (2021). *Impacts of Autonomous Vehicles on Greenhouse Gas Emissions-Positive or Negative?* Obtenido de International journal of environmental research and public health, 18(11), 5567.: <https://doi.org/10.3390/ijerph18115567>
- Merat, N. L. (2016). *Highly automated driving, secondary task performance, and driver state*. Obtenido de Human Factors, 58(4), 642-652.: <https://pubmed.ncbi.nlm.nih.gov/23156621/>
- Miller, C. (2025). *GM's Cruise Ditches Origin Robotaxi for a Self-Driving Next-Gen Bolt*. Obtenido de Car and Driver : <https://www.caranddriver.com/news/a61677537/gm-cruise-origin-robotaxi-dead/>
- MIT. (s.f.). *Moral Machine*. Obtenido de MIT: <https://www.moralmachine.net/>
- MIT. (s.f.). *Moral Machine* . Obtenido de Moral Machine: <https://www.moralmachine.net/>
- MITECO. (2016). *FACTORES DE EMISIÓN DE CO₂ y COEFICIENTES DE PASO A ENERGÍA PRIMARIA*.
- Naciones Unidas . (s.f.). *Objetivos de desarrollo sostenible* . Obtenido de <https://www.spglobal.com/automotive-insights/en/blogs/2025-auto-sales-forecast-global>

Popper, K. (1945). *The Open Society and Its Enemies* .

S&P Global. (2024). *2025 Auto Sales Forecast: 89.6M Vehicle Sales Worldwide*. Obtenido de S&P Global: <https://www.spglobal.com/automotive-insights/en/blogs/2025-auto-sales-forecast-global>

SAE. (s.f.). *SAE Levels of Driving Automation*. Obtenido de <https://www.sae.org/>

Schwarting, W. A.-M. (2018). *Planning and decision-making for autonomous vehicles*. Obtenido de Annual Review of Control, Robotics, and Autonomous Systems, 1, 187-210. : <https://doi.org/10.1146/annurev-control-060117-105157>

Shan, T. &. (2020). *LIDAR-based 3D mapping and localization for autonomous vehicles: A survey*. Obtenido de Robotics and Autonomous Systems, 125, 103393. : <https://doi.org/10.1016/j.robot.2019.103393>

Siddahantjain. (2023). *Six Levels of Vehicle Automation: Fully Manual to Fully Autonomous*. Obtenido de <https://medium.com/@siddahantjain50/six-levels-of-automation-fully-manual-to-fully-autonomous-2d528ea50c63>

TensorFlow. (s.f.). *A Neural Network Playground*. Obtenido de <https://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=circle®Dataset=reg-plane&learningRate=0.03®ularizationRate=0&noise=0&networkShape=4,2&seed=0.72267&showTestData=false&discretize=false&percTrainData=50&x=true&y=true&xTimesY=fal>

The End of driving: Autonomous Cars. (s.f.). Obtenido de Motortrend: <https://www.motortrend.com/features/the-end-of-driving-autonomous-cars-future/photos>

Titterton, D. H. (2004). *Strapdown Inertial Navigation Technology*.

United Nations. (2024). *Emissions Gap Report 2023*. Obtenido de UNEP: https://www.unep.org/interactives/emissions-gap-report/2023/#section_0

- Wang, K. &. (2025). *Global, regional, and national burdens of road injuries from 1990 to 2021: Findings from the 2021 Global Burden of Disease Study*. Obtenido de Trauma Center/Department of Emergency and Traumatic Surgery, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology: <https://www.sciencedirect.com/science/article/abs/pii/S0020138325000816>
- Washabaugh, K. (2021). *NVIDIA Blog*. Obtenido de <https://developer.nvidia.com/blog/nvidia-drive-os-5-2-6-linux-sdk-now-available/>
- WAYMO. (2025). *New Study: Waymo is reducing serious crashes and making streets safer for those most at risk*. Obtenido de WAYMO: <https://waymo.com/blog/2025/05/waymo-making-streets-safer-for-vru>