



UNIVERSITAT DE
BARCELONA



Observatori de
Bioètica i Dret
Universitat de Barcelona

FLACSO
ARGENTINA

Revista de Bioética y Derecho

Perspectivas Bioéticas

www.bioeticayderecho.ub.edu – ISSN 1886 –5887

PERSPECTIVAS BIOÉTICAS

La ética de la automatización en vehículos y enfoques alternos para problemas morales actuales

L'ètica de l'automatització en vehicles i enfocaments alterns per a problemes morals actuals

The ethics of vehicle automation and alternate approaches to current moral problems

ÓSCAR FREDERIC TEIXIDÓ DURÁN *

*Óscar Frederic Teixidó Durán. Grado en Ciencia y Tecnología de Alimentos, Máster en Protección Integrada de Cultivos (especialidad biotecnología de cultivos), Departamento de Producción Vegetal y Ciencia Forestal, Universidad de Lérida (España). Curso en Filosofía Científica, Universidad Nacional de La Plata (Argentina). Email: oscarburgo@hotmail.com. ORCID: <https://orcid.org/0000-0001-8557-0894>.



Copyright (c) 2023 Óscar Frederic Teixidó Durán

Esta obra está bajo una licencia de Creative Commons Reconocimiento-No Comercial-Sin Obra Derivada 4.0 Internacional.

Resumen

En el presente trabajo se utiliza el problema ético de la automatización en transportes y vehículos dotados de inteligencia artificial (IA) para un análisis crítico de las diferentes posiciones éticas y morales clásicas. Debido a la intersección de la IA con los campos de investigación de la bioética, la nomoética y la tecnoética, existe un hilo conductor en sus fundamentos. Además, se juzga necesario para este propósito, así como por rigurosidad, hacer uso de una cantidad mínima de herramientas formales y exponer un proyecto de una teoría ética actualizada, sistémica y materialista, que sobrepase a sus sistemas rivales. Así, se pretende ofrecer las bases para una teoría alterna más fecunda y explicar cómo las posiciones éticas tradicionales pueden ser obsoletas, si no directamente inaplicables, para afrontar la automatización inteligente y problemas morales similares, en la actualidad.

Palabras clave: automatización; deber; derecho; sistemismo; inteligencia artificial (IA); bioética; nomoética; tecnoética; tecnología; automóviles; necesidades; consecuencias.

Resum

En el present treball s'utilitza el problema ètic de l'automatització en transports i vehicles dotats d'intel·ligència artificial (IA) per a una anàlisi crítica de les diferents posicions ètiques i morals clàssiques. A causa de la intersecció de la IA amb els camps de recerca de la bioètica, la nomoètica i la tecnoètica, existeix un fil conductor en els seus fonaments. A més, es jutja necessari per a aquest propòsit, així com per rigorositat, fer ús d'una quantitat mínima d'eines formals i exposar un projecte d'una teoria ètica actualitzada, sistèmica i materialista, que sobrepassi als seus sistemes rivals. Així, es pretén oferir les bases per a una teoria alterna més fecunda i explicar com les posicions ètiques tradicionals poden ser obsoletes, si no directament inaplicables, per a afrontar l'automatització intel·ligent i problemes morals similars, en l'actualitat.

Keywords: automatització; deure; dret; sistemisme; intel·ligència artificial (IA); bioètica; nomoètica; tecnoètica; tecnologia; automòbils; necessitats; conseqüències.

Abstract

In the present paper we use the ethical problem of automation in transport and vehicles equipped with artificial intelligence (AI) for the critical analysis of the different classic ethical and moral positions. Due to the intersection between the AI with bioethics, nomoethics and technoethics research fields, entails a drastic thread to its foundations. Additionally, we judge necessary for our purpose, as for rigurocity, to make use of a minimum amount of formal tools and expose a project of an updated, systemic and materialist, ethical theory that surpasses their rival systems. Thus, we try to provide the basis for a more fruitful alternative theory and explain how the traditional ethical positions may be currently obsolete, if not directly inapplicable, to cope intelligent automation and similar moral problems, at the present time.

Paraules clau: automation; duty; right; systemism; artificial intelligence (AI); bioethics; nomoethics; technoethics; technology; automobiles; needs; consequences.

1. Introducción: situación de la problemática

Una de las problemáticas morales actuales (Boullosa, 2014), que suscita el interés tanto de filósofos como de científicos y tecnólogos a indagar en ella, es la que se refiere a los vehículos automatizados, en particular (pero no en exclusivo) a los vehículos tipo turismo autónomos. Esto, además y cómo se verá más adelante, puede llegar a tener aún mayor importancia con la actual pandemia de COVID-19 (Coronaviridae Study Group, 2020). No obstante, la cuestión penetra en un campo donde coinciden tecnología y ética, siendo un problema tecnoético (Sánchez, 2017); como la bioética trata la biología y aplicaciones suyas –de genómica, embriogénesis, la vida animal, etc.– en ética (Ferrater y Cohn, 1981), o la nomoética trata el derecho jurídico –en jurisprudencia, la ley penal, etc.– y su ética.

Lo anterior es porque, lógicamente, los vehículos son una tecnología y su automatización con inteligencia artificial o IA (Russell y Norvig, 1995) (Boucher, 2018), en sus efectos sobre las personas, genera problemas éticos y morales (es decir, problemas sobre cómo evaluar reglas de conducta –ética– y sobre las reglas mismas –morales–). Acarrea también problemas sobre la vida, en relación a la tecnología y uso, el grado de tecnificación de la vida humana al integrar la IA en hábitos como lo es la conducción. Por lo que este tema es parcialmente bioético, entre otras áreas (Dennett, 1979). Y al requerirse de legislación para poder zanjar estatalmente problemas en hechos tan ordinarios como la conducción, y hacerlo de un modo ético, también se solapa parcialmente este asunto con la nomoética.

Puede definirse exactamente esta problemática, y representarse pictóricamente de forma ideal (los tres conjuntos de indagaciones éticas en equilibrio), de la siguiente manera (Fig. 1)*, resaltando la tecnoética como particularmente consustancial al presente debate:

*Notación conjuntista: definición ($:=$), inclusión no total (\subsetneq), pertenencia (\in) e intersección (\cap).

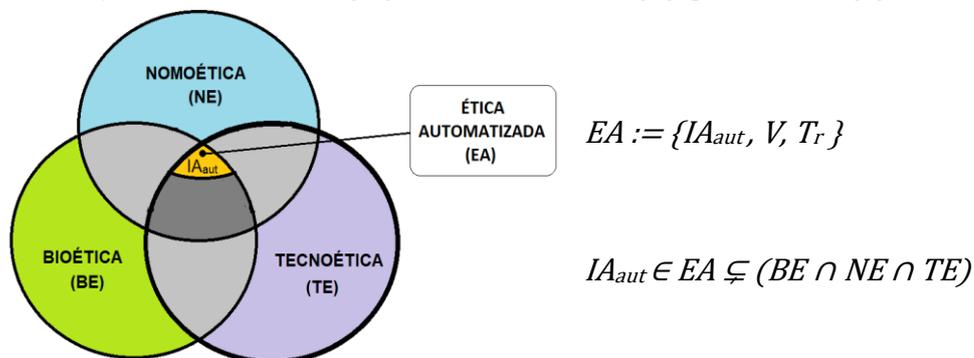


Figura 1: campos éticos que intersectan en un terreno común, donde existe la ética de la automatización como sub-campo, definida por tratar un dominio de estudio genuino: en V vehículos tipo turismo y Tr transportes en relación a la IA_{aut} o inteligencia artificial autónoma.

Todo esto, teniendo en consideración los fines de la ética y la tecnología (Bunge, 2012): la primera, se postula para dirigir el buen obrar, i. e. estudiar y pautar la conducta general en sociedad (“lo moral”) hacia el ‘bien’ de las personas; mientras que la segunda, trabaja con reglas y artefactos que cumplan eficientemente con unos fines (Quintanilla, 2017). Ambas son técnicas, prácticas, o “disciplinas de medios para fines”, pero solo una de ellas es la que, por definición, se debería supeditar a la otra: la tecnología a la ética (“lo ético”). Entonces, por lo postulado, los problemas tecnológicos sobre el bienestar son problemas tecno-éticos, y junto a la legislación y las modificaciones biológicas, deberán realizarse estas actividades supliendo el fin moral de la mejor forma. Esto es: facilitar la vida con la tecnología pero respetando el bienestar, elaborando una tecnología amigable (Quintanilla, 2019) y que garantice ‘lo bueno’ (Moore, 1922, pp. 4-6) (definición la cual se esbozará).

Pero; ¿cuáles son esos problemas éticos concretos? ¿Hacen falta realmente normas de conducta social, i. e. reglas morales, para la automatización? ¿Acaso no será suficiente la sola IA (Russell, 1996) para evitar muchos de los problemas de los vehículos autónomos?

2. Planteamiento de la discusión

La cuestión moral radica precisamente en la parte más delicada, por su peligro, que habitualmente ha tenido todo vehículo confeccionado: los accidentes (DGT, 2021). Antes y aún hoy, la tecnoética de los vehículos –sean barcos, coches, trenes o aviones– se suele limitar a su buena conducción. Entendiendo por “buena” aquella que se ciñe a postulados éticos de no causar daños o minimizarlos, siguiendo para ello las pautas que se prueban haberse aprendido y memorizado, pasando ciertos exámenes teóricos (DGTtest, 2021) y prácticos. Un transporte muy automatizado eliminaría ese asunto: sin conductor o humano en el comando, quizá no haría falta tanto control. El propio vehículo ya lo tendría todo diseñado en su chip con una IA dispuesta en sus circuitos. Tales vehículos, sobre todo los de tipo turismo, verían luz comercial en el futuro. Un futuro poco lejano (Fuentes, 2021).

Pero, ¿qué asegura la maniobra de un vehículo con IA ante un accidente o problema moral casual? O más precisamente formulada esta cuestión: ¿cómo debe estar diseñada la IA de un vehículo para afrontar ciertas situaciones dañinas, aleatorias y poco previsibles? Se da bien por supuesto que un diseñador, con mínima empatía y racionalidad, siempre buscará proteger al conductor del máximo daño posible: así pues, todos los accidentes contra materiales y estructuras, objetos inertes, se planificarían para detectarse y que el vehículo protegiera a los ocupantes de su interior, o a su carga material si se detecta “vital” o “más valiosa” esta que los daños posibles en un choque con más materia inerte. Aquí entrarían en discusión ciertas ideas económicas (Rand,

2007) –y por supuesto, morales– de corte individualista (Rothbard, 2002): unos contratos “libres” y pasivamente trazados de los trabajadores para con unos dueños sobre sus mercancías, podría implicar oponerse por parte de los dueños y su capital a la prioridad de la vida de sus trabajadores respecto a su mercancía. Pero si la carga es materia no-vital, y si peligran personas, lo ético sería en tal problema salvar siempre la integridad de las personas (Smith, 2010). Eso sería, en general lo mejor, si se busca el bien (a definir aún; pero en todo caso no se habla solo de “bien” como el mero placer, virtud o felicidad: si no sobre el bienestar factible y biopsicosocial). Pero esta tarea se torna más compleja cuando, contra quienes se accidenta, también son personas. Sin entrar aquí en qué es ser “persona”; i. e. sujeto de derechos o consideración moral, sí se postulará que en accidentes con personas intervienen factores referentes a la integridad fisiológica, psicológica y social en un conflicto mutuo (Shujman et al., 2010).

Esta clase de problemas son los que tienen un marcado debate, porque pese a ser quizá marginales, son dilemas: toda solución parece contraer insatisfacción en una misma clase máxima de valores, y por ello todo parece, al menos a priori, erróneo hacia el bienestar o el “bien” y “lo bueno”. Luego, no se habla aquí de accidentes evitables: se busca discutir qué programar en dichos vehículos cuando peligre el bienestar de unas personas y otras en corto tiempo –menos aún que el del ojo humano– (2 segundos de media total de tiempo para el P.I.E.V. o Percepción, Intelección, Emoción y Volición; Berardo y Irureta, 2008). Puede ser difícil creer que se logre programar un vehículo para que, en un accidente dado en pocas centésimas de segundo, registre su IA (Bostrom, 2003) que no ocurren más daños que los interiores y actúe para minimizar estos a tiempo.

En el momento en que alguien se desplaza por una carretera y por algún infortunio el coche que conduce va a impactar contra algo o alguien/algunos (Lorenzo et al., 1999): si se supone que en unas centésimas antes del choque, la IA del vehículo detecta un choque seguro, cosa ya posible con la tecnología actual, ¿qué debe ejecutar esta? ¿Debe buscar la mejor manera automatizada de minimizar el daño y riesgo de otros o el propio? Y si un vehículo falla en su mecánica, no puede frenar, y la IA detecta que va a estrellarse contra personas: ¿a quién debe seleccionar el coche que dañará sí o sí, para causar el mínimo perjuicio posible, si solo puede modificar o intentar modificar la trayectoria inminente? A esto precisamente se dedica a “jugar” (Laird et al., 2001), en forma de encuesta masiva, el M.I.T. (Massachusetts Institute of Technology) en una web (*Moral Machine*) con pruebas en forma de dilemas éticos¹. Su teórico objetivo es analizar en datos estadísticos las respuestas de las personas en eventos morales, con una muestra internacional elevada². Las propias personas pueden configurar los dilemas para idear nuevos problemas, y ver qué

¹ El enlace virtual a tal sitio web donde se encuentra el juego moral citado, del M.I.T., es el siguiente: <http://moralmachine.mit.edu/> (“As an outside observer, ...”) [Consulta: 30 marzo 2022].

² Se tiene constancia de que los primeros resultados de este programa online ya fueron publicados como un estudio científico revisado en la revista Nature (Awad et al., 2018).

conductas morales están más o menos extendidos en la sociedad actual para zanjar dilemas éticos y *a fortiori* poder servir de auxilio, cuando se vayan a aplicar tecnologías automatizadas (en elegir qué operaciones mecánicas diseñar para su IA ante unos dilemas dados). Es posible hacerle algunas objeciones, aparte de las críticas a este tipo de planteos, acerca de dilemas cuando se usan para evaluar morales en sociedad (Kahane, 2015).

Se puede decir que los problemas aleatoriamente dados por la web (al menos hasta hoy) son siempre de deberes. Esto es, tratan a los pasajeros del coche y a quienes puedan atropellar por error, sin referir al bien del sujeto propio (“yo”) y su situación o perspectiva (la perspectiva propia como humano en el comando o la perspectiva propia como peatón) dentro del problema: aunque en teoría sean problemas planteados para una generación de coches inteligentes tipo turismo, eventualmente de uso global, en estos dilemas no se explicita la posición del agente que rige la operación de la IA (puede leerse en la pantalla inicial: ‘Como observador externo...’ -v. la nota 1-). Se simula que la situación de uno mismo es externa a todo evento dilemático moral que luego se propone. Siendo en este caso, una elección donde el “yo” no está presente, y los “ellos” o el “tú” sí; cuando lo más sistémico sería tratar el peligro de todos los posibles implicados, más cuando se sabe la influencia del estrés y estados emocionales en la toma de decisión moral de un ser humano (Valesolo y DeSteno, 2006) (Youssef et al., 2012). Las cuestiones planteadas en esta web, donde se dan tales conflictos morales, son del tipo “deber-deber”³ y no de derecho-deber (tampoco de derecho-derecho). No se consideran todas las perspectivas que cabría reunir.

Agotar toda la problemática moral de la automatización de vehículos en esa perspectiva sola, sería un error tal y cómo se representa en los dilemas de la citada web. En esta se baraja además diferentes situaciones donde se supone que el jugador (“yo”) controla la operación de una IA representada en un coche inteligente supuesto, que detecta a personas contra las que se está a punto de atropellar sin remedio y ajeno a consecuencias sobre uno mismo. Entonces, puede escoger la mejor combinación entre muertes en el interior del vehículo o la de personas externas. Y detecta si esas personas son embarazadas, jóvenes, adultos, deportistas, obesos, criminales, intelectuales, etc., para poder sopesar y decidir⁴.

Dejando de lado cómo sería factible (Flórez, 2019), en las pocas milésimas de segundo en las que un error y un accidente se suelen producir, detectar este y analizar, procesar y aplicar sobre

³ Para entender los dilemas acotados de tal modo según una oposición de normas, divididas en deberes y/o derechos, cabe partir de que hay sociedades conformadas por individuos y terceras personas respecto a estos, todos con sus rasgos morales a cumplir; derechos y deberes. Una tesis que no es exclusiva pero sí muy enraizada en el sistemismo social, principio ontológico general aplicado a lo social: la sociedad es un tejido compuesto a su vez por individuos, todos ellos materiales, sin reducciones holistas ni individualistas

⁴ Es patente que los principios subyacentes a este “juego virtual” no son solo consecuencialistas, sino además utilitarios, al asumir el calcular según si hay más o menos años de vida, más o menos aportes útiles a la sociedad, etc. No se deja abierta la cuestión ni tampoco se decanta por un consecuencialismo de corte objetivo en sus valores, sino que se asumen factores relativos generalmente a la felicidad y la productividad.

los datos relevantes (*i. e.* grado de vejez, productividad, delitos cometidos, situación física, o etc.); lo cierto es que, varias respuestas morales pueden darse a dilemas morales como los planteados en esa web. Pero aquí empiezan las mayores dificultades.

3. Las propuestas morales clásicas

3.1. Utilitarismo

Una respuesta clásica a qué evaluar en problemas dilemáticos de deberes hacia otros, como los dados en la web *Moral Machine*, podría darla la teoría utilitarista (Bentham, 1836) general de J. Bentham y S. Mill: evaluar el bien como lo útil o utilidad, *i. e.* felicidad que toda persona puede tener –y es “lo bueno” para uno cumplir–. Llevado al límite, esta idea implicaría pensar sólo en la utilidad inmediata: respecto al juego de la automatización (Keeling, 2019) anterior, se primará a deportistas sobre obesos porque el deportista vivirá posiblemente más años y mejor, el joven frente al mayor por disfrutar y trabajar el primero más años, al empleado frente al desempleado al tener el primero una mayor productividad social, etc.; y siempre considerando que, entre iguales, el mayor número de sujetos priva.

Pero las teorías utilitarias suelen disponer de normas para lograr un bien global, a largo plazo (Mill, 2014, pp. 26-27): para impedir la rebelión de los desempleados se les tratará como iguales a los trabajadores. Aunque, de no haber nada que perder –y quizás, incluso de ayudar a disminuir la criminalidad–, podría defenderse que siempre privara el decente sobre el criminal o ante el que tuviera más antecedentes que otro (aunque es dudoso si tener antecedentes puede servir, siquiera, como indicador de inmoralidad potencial). En cualquier caso, el utilitarismo representa una teoría moderna que bebe de diferentes fuentes clásicas para intentar una especie de “ciencia ética” *per se*, un proyecto preciso.

Se puede representar el enfoque utilitarista como sigue, en un modelo simple (Mosterín, 2000) de evaluación ética de “hechos morales” o estados de cosas que involucran acciones de humanos y sus valores: el utilitarismo buscaría efectos medibles o al menos evaluables empíricamente, para impulsar siempre aquellos actos que aporten mayor bien (Sidgwick, 1981), que sería la felicidad, lo útil (sea ya inmediatamente o bien por reglas). Raramente se consideran los márgenes de error en la conducta falible de las personas agentes.

Definidos A y Y como dos actos diferentes guiados por normas diferentes, con sus respectivas consecuencias de bien B' y B , siendo B' más elevado (\succ), y cuyo cómputo o sumatorio (Σ) da resultados X' y X respectivamente. Estos se evalúan éticamente, siendo más

proporcionalmente éticos cuanto mayor es X , en unidades de utilidad. Puede ponerse así X en función $f(x)$ de lo ético como propiedad evaluativa Z , como $Z(X)$. Ya que el bien B se iguala con lo útil o el disponer de la mayor cuota de felicidad, componente subjetivo, tales valores se designarán como S . Esto da lugar, para dos evaluaciones éticas finales lógicamente diferentes (Z' y Z) a estos gráficos cualitativos A y Y , donde $A > Y$ (Fig. 2).

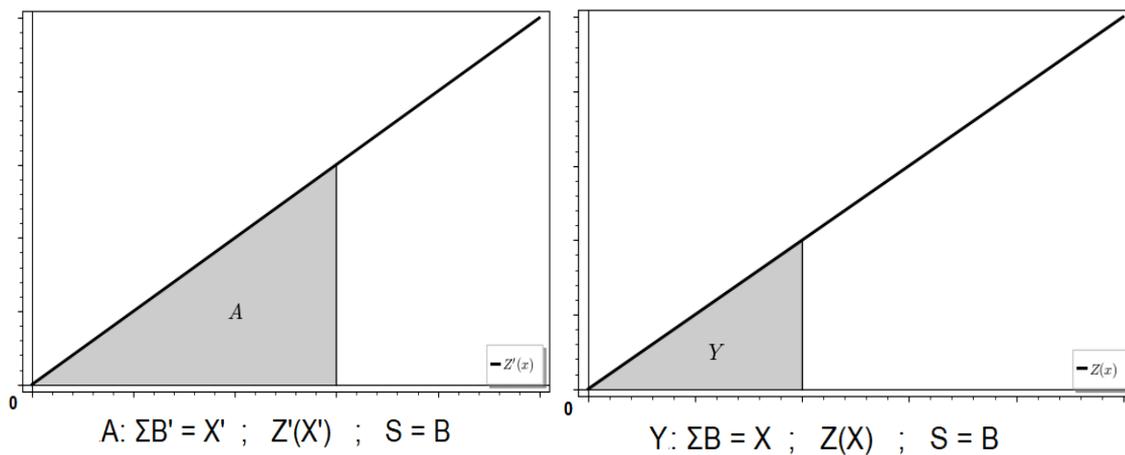


Figura 2: Gráficos comparativos de utilidades A y Y , con A mejor éticamente que Y .

E igual valdría si los gráficos y sus áreas respectivas fueran inversas en su análisis moral: si en vez de cuantificar más o menos bien se cuantificara utilitariamente más o menos mal. De la misma forma se elegiría el que menor mal y mayor bien supusiera, o en el caso del utilitarismo de reglas, lo mismo pero siguiendo ciertas reglas generales a obedecer.

Hay diversas objeciones aplicables al enfoque utilitarista (Rawls, 2006, pp. 9-11, 34-44 y 246): en primer lugar se tiene el hecho de que, aún si no es necesario, apenas hay leyes utilitarias expresadas, algo no indispensable pero deseable en una teoría con dos siglos de vida que propone el cálculo, aún si es cualitativo, del bien. En segundo lugar, al tomar el subjetivismo de lo 'útil', se hace ambiguo tratarlo; sea en felicidades, placeres o intereses, a menos que se haga *a priori*; pero esto impediría la medición de utilidades *a posteriori*, empírica y consecucionalmente. En tercer lugar, una vaguedad en el utilitarismo puede ser bastante peligrosa al poder tomar como útiles ciertas arbitrariedades discriminativas pero aceptadas y con mucha utilidad en una sociedad, como priorizar a un género sobre otro en sus derechos. Aparte, no se suele considerar la falibilidad del obrar utilitarista, su margen de error. Y finalmente, un utilitarismo desenfrenado que no considere los propios derechos (Nozick, 1974, p. 41) lleva a conflictos como tener que escoger el sacrificio de la vida propia por el deber sobre la utilidad en terceros, si aquellos requiriesen de felicidad constante; el "monstruo moral utilitario". Esto puede ser, en ética de la automatización, el dilema que enfrentaría una IA de un vehículo que debe elegir entre sus pasajeros con poca

felicidad a cubrir o peatones con gran demanda de felicidad: ajeno a más factores, el vehículo quizás debería implosionar para desintegrarse antes de dañar a los peatones, y acabar así con la vida del humano al comando y pasajeros en pos de terceros (sacrificio).

3.2. Deontologismo

Pero hay otras respuestas bastante populares, como el deontologismo (Kant, 2003) de I. Kant, que pretende ser *a priori* y formal: de norma universal o imperativo categórico, sin más desviaciones en las situaciones que puedan darse, que lo ya teorizado como universal. Sus defensores no pueden atender así a fines dinámicos según circunstancias cambiantes, solo pueden valorar por igual en toda situación según una ley moral universal, máxima, ergo buscar el bien siempre por la sola voluntad (Kant, 2007) y no por sus consecuencias.

Y de existir una oposición entre fines, debería dejarse seguir el curso de las vidas en juego, o bien priorizar algunos fines selectos considerados como apriorísticamente superiores frente a otros. Las consecuencias no importan, si no solamente la regla de la voluntad.

Por lo anterior, puede (no) representarse evaluativamente el deontologismo kantiano en dos gráficos *A* y *Y* donde cierto bien como meta *B* se logra con las reglas *A* e *Y* en valores metafísicos aprioristas, sean subjetivos o no $S \vee O = B$, con unidades arbitrarias (Fig. 3).

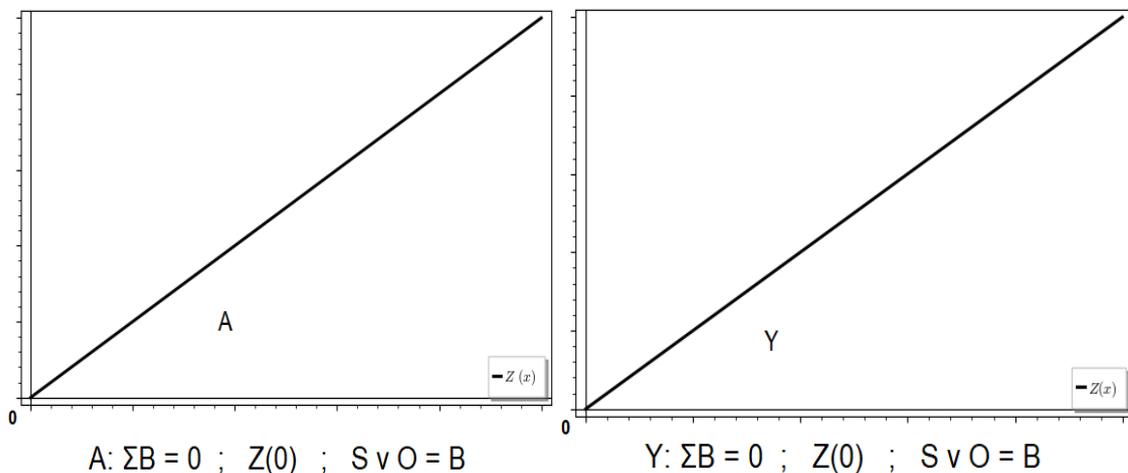


Figura 3: Gráficos deontológicos de consecuencias, nulas al no evaluarse efectos, ergo $A = Y$.

El deontologismo kantiano conlleva varios errores (Schopenhauer, 2016). Por un lado, repite la carencia del utilitarismo “sacrificador”: puede no tener en cuenta al propio individuo en sus derechos. Pese a intentar partir de la autonomía como principio, debe ser el sujeto quién se gane tal autonomía por sus deberes satisfechos. Aparte, es problemático saber cuál es la universalidad de la ley absoluta en el terreno moral, dado que en ética no se derivan leyes *a priori*

inmutables, como sí se hace quizás en la matemática (por ejemplo, en geometría), de considerarse en el sentido kantiano; y aunque fueran de tal modo kantiano tomadas y deducidas semejantes ideas en una teoría moral, entonces por ser sólo matemáticas y universales carecerían de hechos referidos: por lo que no serían asunto fáctico ni mucho menos, ético o moral. Bajo este error subyace el fallo ontológico de confundir la formalidad de enunciados como $3 \cdot 3 = 9$ con los enunciados fácticos o de hecho como $F = m \cdot a$, y posibles reglas como “para sobrevivir, no saltar desde 10 m a $g = 9,8\text{ m/s}^2$ ”; error propio de la filosofía kantiana al sostener sus sintéticos *a priori* fenomenistas (Bunge, 2004, p. 447), lo cuál parece ser un callejón sin salida. También determinar qué es la buena voluntad para poder alcanzar la máxima *a priori* es difícil dado que: a veces se define la volición y voluntad vagamente; y definir objetivamente la voluntad es difícil hoy, hasta saber qué es la voluntad psicológica y neurobiológicamente, tarea aún en plena elaboración (Gutiérrez-Soriano et al., 2017, p. 11). En resumen, todo ello da pie a trazar máximas éticas arbitrarias (Mosterín, 2013, pp. 49-50). Finalmente, la inaplicabilidad a la realidad sella su tumba: cuando dicha ética trata de absolutizar el obrar en un mundo que es dinámico y cambiante, el “corsé kantiano” resulta demasiado rígido, poco práctico. Sobre el tema central del artículo, sería sin duda absurdo programar la IA de automóviles para que, por principio, no hagan nada ante un dilema entre personas a escoger: ya que la IA, de programarse para suponer una imaginativa voluntad (en ella misma fingida ficticiamente o con referencia a los ocupantes del vehículo, “delegando” su voluntad a la IA), debería solo obrar para lograr el bien *a priori*, y de tener que escoger sacrificar algo, no movería ficha, ya que no debería obrar por las consecuencias. Pero entonces sería inútil tener una IA en dilemas de accidentes de tránsito si no va a reaccionar porque está programada para no elegir en dilemas, al sopesar un criterio *a priori* para no intervenir en decisiones tales (véase la precedente Fig. 3). Se prefiere que haya cualquier muerte con tal de preservar impoluta una “voluntad” que conlleve la regla máxima moral (se reitera, sea esta fingida en la IA o como forma de delegar los ocupantes su autonomía).

3.3. Ética de las virtudes

También se puede fijar la atención en la clásica ética de las virtudes (Aristóteles, 2005), el sistema original de Aristóteles y recuperado contemporáneamente por E. Anscombe (Anscombe, 2006). Tal teoría toma por “lo mejor” la virtud de alguien como un sujeto agente u obrador, sobre aquellos valores entendidos como los que determinan quién es o no virtuoso: se sigue un modelo ideal de la persona benévola. Esta es una teoría clásica intuitiva y sugerente al centrarse directamente en cómo es quién bien obra y actúa.

Pero no deja de contener una mezcla de todos los errores antes señalados: por un lado, dado que en último término, y pese a intentarse racionalizar, los valores que se toman por virtudes son valores ambiguos, se pueden tomar tanto valores necesarios como valores que dependan de contextos sociales subjetivos así: por ejemplo, en un entorno cultural determinado, se darían juicios que no podrían extenderse como reglas en circunstancias objetivas y realistas alternas, mínimamente regulares (sin por ello tampoco pretender unas leyes absolutas). Al poder ser tales virtudes subjetivas, ya que no están definidas de forma suficientemente impersonal y “realista”, *i. e.* objetiva, se puede caer *a fortiori* dentro del saco de valoraciones personales y las arbitrariedades (junto a: útil, felicidad, interés, etc.).

Por otro lado, de dar razones apriorísticas sobre qué es la virtud y el ser virtuoso, también acarrearía los problemas previos. Los vehículos autónomos tipo turismo quizás ejecutarían cómo operar según una base de datos de la axiología media predominante en la sociedad, y lo actualizarán cada cierto tiempo para adaptarse y regirse por estos valores sin cálculos consecuenciales, con tal de ser reconocidos los automóviles como sujetos virtuosos, o de nuevo, estar programados para delegar teóricamente tal “virtuosidad” a sus ocupantes y no la IA misma. Conlleva los anteriores problemas del deontologismo.

4. Balance de la situación

A menudo, en dilemas de vehículos autónomos, se suele creer que el utilitarismo es más intuitivo y preciso (Singer, 1995), al fin y al cabo se ideó con potencial científico al buscar computar bienes –aunque pocas veces se muestre formalmente–. El utilitarismo subyace en los dilemas de *Moral Machine* del M.I.T. sobre coches automatizados con IA, al asumir datos en número y sugerir calidad, y al posibilitar que las personas hagan sus “cálculos” sobre ellos. No dejan de analizarse dilemas, algunos elaborados por los propios participantes, en donde se debe (y se sugiere) elegir entre (La comunidad, 2018): deportistas con gran esperanza de vida u obesos con algo más corta esperanza de vida; expresidarios que pueden volver a reincidir o gente sin expedientes más fiable; entre gente rica y contribuyente o gente pobre que probablemente no volverá a contribuir a la sociedad; entre animales (de diferentes especies) o humanos; y entre jóvenes con mucha vida y felicidad por delante o ancianos con posible escasa vida, etc. También varias combinaciones de estos casos. El vehículo realmente debe, con su IA, no solo “elegir” sin más, sino hacerlo en pocos milisegundos y registrando al momento los datos de todos los individuos cercanos, sabiendo perfectamente su estado. Es un programa curioso para ver hasta dónde llega el utilitarismo en sus últimas consecuencias. Sin embargo tiene las deficiencias citadas por severas idealizaciones y asunciones muy parciales en su filosofía.

Y se ha visto que la teoría utilitaria no solo es una posición habitual, sino que seguramente sea de las más usadas en tecnoética y bioética, quizás junto al principalismo bioético (Clouser y Gert, 1990); luego es ineficaz emplear tal teoría sobre ciertos problemas, con diversidad de errores teóricos a solucionar. Y que quizá nunca pueda zanjar por definición de su propia doctrina sobre lo útil (Bunge, 2010, p. 121), valor subjetivo. Además, con independencia de la teoría, utilitaria o no, algunos dilemas planteados de forma idealizada y muy simplista pueden constituir “pseudo-dilemas”. El debate sobre dichos dilemas, en exceso irreales, y a veces también propuestos sobre vehículos con IA, debería omitirse.

Todo esto lleva a proponer la siguiente tesis pesimista: las teorías clásicas y hoy en día predominantes, esencialmente son (al menos las principales y analizadas aquí) obsoletas y no sirven (o no ellas solas) para afrontar los problemas morales de más actualidad, junto a los análogos de futuro cercano. Los dilemas con vehículos autónomos con IA son clave para ilustrar tal obsolescencia. Y los problemas planteados con premisas o que sugieran ideas de las teorías éticas y morales precedentes, no suelen ser fecundos. Quizás deban potenciarse otras teorías, aunque puedan tener puntos comunes con las teorías clásicas.

5. Alternativa propuesta de metaética y ética normativa

5.1. Proyecto

En este punto se sugiere una propuesta o proyecto de ética moderna, la cual luego podrá aportar respuestas a los dilemas sobre vehículos autónomos. Para que tal sistema teórico pueda cumplir con sus metas en sociedad, debería encararse de forma práctica, esto es, de forma normativa o reglamentada (un sistema técnico; Quintanilla, 1983, p. 122), a la par siendo sistémica y lo más objetiva posible (Bunge, 1972). Esto no excluye la ética teórica, explicativa o descriptiva, pero esta sería dirigida a la praxis misma. La objetividad, i. e. impersonalidad, se postula como hipótesis que pasaría por tratar necesidades humanas –y posiblemente otras necesidades animales (Fraser, 1999)– (indagadas por las ciencias naturales, biosociales y sociales) como aquello de mayor valor⁵. Tal objetividad no sería ontológica (Nagel, 1986, p. 139): sin animales no hay valores; residiría más en lo práctico, en las consecuencias reales y materiales, inapelables subjetivamente, de no satisfacer ciertos tipos de valores básicos para la especie humana (*Homo sapiens*).

⁵ Conviene señalar un matiz: el agua es valiosa o de valor para los animales, pero el agua no es un valor *per se* o de suyo, el valor es la ficción que un ser animal complejo puede elaborar para evaluar tal recurso.

Supervivencia y salud son los dos tipos de necesidades básicas, precediendo la supervivencia a la salud, y sus valores objetivos respectivos se tomarían en todo aspecto biopsicosocial como lo más primario o bueno, el bienestar, común a todo ser humano y cosmopolita (Cortina, 2021, p. 131); pero sin oponerse a valores subjetivos como los de placer, felicidad o realización (en un nivel axiológico de menor prioridad). Los requerimientos serían cambiantes, con falibilidad material para ser cumplidos y dependientes del conocimiento disponible. En definitiva, la ética no debería (ni puede) ser ciencia, pero sí puede (y debe) ser “científica”, como la medicina: práctica y con base las ciencias, pero sin reducciones (Haack, 2017).

De esta manera, la discusión metaética asume cierto realismo normativo (Nagel, 1986, p. 139), solo como una objetividad a la que atender para prescribir normas racionalmente –siendo este proyecto en *stricto sensu* un consecuencialismo moderado, más no de corte subjetivista–. También se toma un naturalismo epistemológico moderado (Diéguez, 2012, p. 240), *i. e.* un cientifismo “light” en el que las ciencias se consideran de las mejores formas de conocimiento pero no únicas, en igualdad a la filosofía, autónoma pero basada en ellas para teorizar o guiar la acción. Y supone el materialismo al pensar a los animales como particulares con rasgos cambiantes, pero también el ficcionismo (Romero, 2020) y un constructivismo (no excluyentes del materialismo, puesto que la ficción construida solo tiene existencia convenida, luego irreal) dado que todo valor es per se una invención de seres animales, en una forma más o menos precisamente construida – más no siempre trazada arbitrariamente– para evaluar la satisfacción de sus deseos y necesidades. Y se parte de la aludida posición sistémica pluralista: varias categorías de valor y de derechos y deberes en los sistemas sociales, jamás reduciendo a lo sólo biológico o lo sólo cultural.

Esto solo supone los trazados rápidos de un proyecto ético alternativo a los sistemas clásicos, una “ética técnica”. Pero son suficientes por ahora con tal de resumir y ofrecer una opción sólida a las teorías predominantes en el debate actual. Puede expresarse todo lo dicho, de un modo más ordenado e ilustrativo –aunque todavía programático–, como sigue.

- ◆ Como se postuló al hablar del bien y necesidades al principio de este artículo; se dividiría entre tipos de valores humanos (Bunge, 2002, p. 43), subjetivos y objetivos, *S* o *O*, y se obraría ante todo sobre lo objetivo u *O*; siendo *S* terreno laico para con esos valores, donde no interceder innecesariamente (y solo hacerlo si se dificulta realmente el satisfacer valores *O*, primarios). Cabría realizar algunos cálculos de valor, pero previos a aplicar las reglas *in situ* en los problemas morales. Y se haría sobre necesidades básicas, con auxilio de varias ciencias y técnicas: bromatología, medicina, higiene, psicología, sociología, etc.
- ◆ Conociendo lo anterior, tanto en las evaluaciones éticas como en posibles dilemas se debería tener en cuenta la mutabilidad de las circunstancias y seres animales, materiales, así como el

posible error práctico. *Ergo* se tiene que atender a la falibilidad humana (pues son sólo los seres humanos quiénes pueden ejercer responsabilidad, el deber) y por ende considerar margen error “ ϵ ” de todos los artefactos diseñados, incluidos los artefactos conceptuales y normativos. Y cabrá planear el obrar moral sin necesariamente tener que computar esperanzas ni hacer futurismos acerca de unas felicidades (Margot, 2007), sino tratando necesidades bajo reglas que sirvan en todo problema moral de un mismo tipo.

- ◆ Del mismo modo, al postular la sociedad como un sistema con individuos y colectivos, esto es, moralmente con cargas (deberes) y merecimientos (derechos) (Rawls, 2003) de algún modo N distribuidos, podría aclararse más el rol normativo del agente, su situación y perspectiva en dilemas, y tratar su bienestar: fisiológico, psicológico o sociocultural. De ahí que los valores básicos de supervivencia y salud se puedan tomar a su vez como biopsicosociales, y en cierto modo cosmopolitas, por su globalización básica en humanos; universalidad de valores que, no obstante, sigue siendo relativa a especie (no absolutista).

Pueden representarse ahora un caso ilustrativo, sin pretender resolución alguna, como el siguiente ejemplo donde dos sujetos –uno es el propio agente– tienen un conflicto en sus sumas de necesidades varias (de cualquier tipo biopsicosocial) a satisfacer, mediante reglas derecho y deber; “yo” y “él/tú”, en unidades de cumplimiento de normas (Fig. 4).

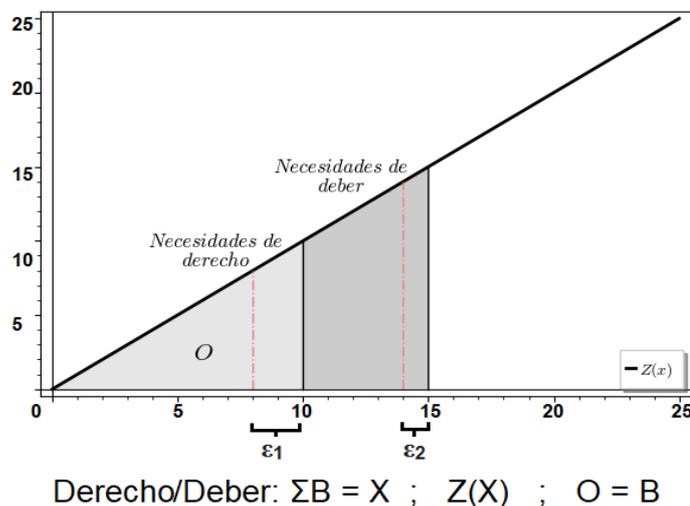


Figura 4: Un cómputo de bienes según valores O , valores sobre necesidades, con un bienestar total X evaluado éticamente $Z(X)$ sobre los derechos (=10) y los deberes (=15), en juego aquí, y con ϵ errores (podría no lograrse cumplir en cada grupo de reglas con un deber y con dos derechos –y sus valores objetivos asociados–, respectivamente).

El anterior gráfico (Fig. 4) podría ser bien un modelo ideal aplicable a algún evento moral donde evitar generar necesidades de una persona en peligro, y uno mismo, bajo el trato de ciertas normas: pero, cómo se ha afirmado, solo sería analizable teóricamente, no en la praxis. Esto es, en ética se deben investigar sistemas morales generales: reglas en derechos y deberes (de diferentes

tipos) sobre grupos enteros de problemas, sin promover hacer cálculos morales en global, en los problemas críticos in situ –sea dilema o no–. No se ofrecerán en este trabajo unas leyes morales firmes y generales de justicia más allá de asumir cierta equidad entre derechos y deberes (cómo se expondrá), pero justificar cualquier norma moral siempre requerirá evaluar según necesidades. Y en este sentido, sí cabría hacer algo análogo al ejemplo aludido; pero como investigación. Desde luego, sería contraproducente ponerse a computar necesidades satisfechas o no en medio de un atraco o en un típico dilema del tranvía (Thomson, 1985) (que, contrario a la idea habitual, no plantea necesariamente un dilema deber-deber sino también y ante todo, derecho-deber, puesto que en el planteamiento original se situaba al lector como al comando del tranvía).

Esta postura materialista, técnica y sistémica, cuya base teórica ha sido expuesta pero cuyo sistema moral vertebrado por una máxima N no ha sido aún presentado; ya puede proveer de varias reglas generales y principios desde los cuales los dilemas con sistemas automatizados pueden partir, para tratar de aportar soluciones y respuestas a problemas éticos. Por ejemplo: póngase en la situación recurrente de un posible accidente de tráfico con peatones y un coche inteligente, con una IA autónoma, y suponiendo que el sujeto agente y propio es el conductor y no algún miembro de los peatones citados en peligro.

En primer lugar, el valor a perseguir sería el máximo bienestar, ergo la suma de bienes como necesidades en juego y pertinentes en el problema⁶. No lo sería el deseo de los sujetos peatones o del propio agente durante el choque ni la posible felicidad que se pueda generar in situ y/o en un futuro incierto de un modo u otro. Y en un dilema tal se evaluarán estas necesidades, sin suponer conocimientos cuasi-místicos sobre una IA que cargue en ínfimo tiempo, por ejemplo, todo el historial penal de los peatones que pueda atropellar.

En segundo lugar, al proponerse cierto sistemismo social desde el que imperar reglas morales, en este evento se deberán tener en cuenta de cualquier forma deberes y derechos implicados; escoger privilegiar al humano al comando sobre otros o escoger sacrificar a este sobre terceros tiene que sopesarse según el rol que juega cada sujeto en la situación. No es una cuestión, así, de simples balances numéricos; no será despreciable ni el derecho ni el deber por uno u otro. Aquí se aceptará, a modo de exposición, las normas de derechos y deberes en total simetría o equidad (Del Río, 2018) derecho-deber como regla máxima N, a lograr en cualquier problema moral. Por ende, se debería tener en igual consideración derecho y deber: es elección del agente si sacrificarse o privilegiarse. Ya que el coche con IA actuará autónomamente, lo óptimo como

⁶ Las necesidades son la raíz sobre las que se idean los valores objetivos, por lo que las necesidades son sus portadores, y a su vez (por ello mismo) son también portadoras indirectas de derechos-deberes en sociedad.

regla en tal suposición de justicia equitativa, sería: (a) ofrecer venta de vehículos de líneas de fabricación con un sistema operativo ya diseñado que siempre sacrifique al agente moral conductor o siempre al paciente moral peatón, advirtiéndolo legalmente –lo cuál no implica que se publicite como algo benévolo de tal tecnología, sería solo una opción–, o (b) la venta de vehículos inteligentes que sean configurables a elección, antes de la conducción misma y modificables en algún punto de esta (y en cuyos registros cabría esperar que constara legalmente tal decisión). En (b) se evitaría reducir deberes del usuario como principio de autolimitación, si se teme que tal tecnología pueda distanciar en exceso el deber de la decisión consciente⁷ (Hueso, 2020). Por esto último, quizás el trabajo industrial debería revisarse en esta dirección (b).

En tercer lugar, las reglas morales de las que se deriva el adecuado trato de la IA automatizada en transportes, deberían probarse en su eficacia empíricamente, aún de forma indirecta respecto a la filosofía y la bioética. Esto es, cabe realizar estudios sociales que generalicen, aunque sea de forma tentativa y a un bajo nivel teórico, tendencias sobre ciertas reglas de derechos y deberes (combinadas en alguna disposición normativa, basada en una posible máxima de justicia *N*) que indiquen una ayuda sobre aumentar el bienestar en problemas en vehículos autónomos: con un modelo como el anterior descrito (Fig. 4) pero más general. Y luego de evaluar las reglas eficaces en tales accidentes de tráfico, se procedería a legislar nomoéticamente. En total se debería: (i) lograr datos sobre el estado de las necesidades en esta clase de eventos, (ii) analizar el grado de compromiso distinto que tiene una comunidad respecto a otra en sus valores y normas de derecho y deber sobre una máxima moral *N*, (iii) contrastar estadísticamente los datos para ver si significativa y causalmente hay mayor bienestar en la comunidad con más compromisos éticos con *N*, que será a su vez (iv) en ética juzgada evaluativamente como eficaz –o no– para luego (v) legislar o modificar la legislación imperativa vigente, acorde a lo ético.

5.2. Otras aplicaciones en la ética de la automatización en vehículos

Un tema donde aplicar la propuesta ética dentro de la problemática de la IA automatizada, es en atribuir culpas –que se pueden definir (Moreno-Romero, 2020) como los juicios públicos sobre los deberes transgredidos por alguien– y tenencia de responsabilidades ante un accidente. Para

⁷ Principio de autolimitación que podría enunciarse como sigue: ante un dilema moral, el sistema autónomo de decisión debe ser desactivado o reducido de forma eficiente para que el usuario pueda asumir su propia responsabilidad, tomando así más acciones individuales, de tal forma que su intervención en la toma de decisión de la IA sea rastreable y más transparente (Hueso, 2020). Es el caso de la opción programable (b); en efecto, el dispositivo puede desactivarse o adaptarse en todo momento para que su usuario pueda actuar.

ello cabría definir antes a los sujetos morales, en general, que deben tener derechos y deberes en alguna forma bajo la ética básica presentada. La ética, como se dijo anteriormente, debe definir teóricamente la 'persona'. En este proyecto de ética sistémica, materialista, consecuencialista y de moral como técnica construida sobre las ciencias; al tratar normas y valores, debe poner en categoría moral a todo animal con cognición y autoconciencia presentes que puedan sentir y valorar, ser sujeto de derechos.

Esto es: mamíferos, vertebrados complejos en general, y algunos invertebrados (Diéguez, 2014). No se profundizará más, pero puede trazarse un límite allí aun si es aproximado.

Si se parte de tal ética técnica, y suponiendo aquí de nuevo una máxima de justicia como equidad derecho-deber: en un accidente, dado que la máquina no es un ser consciente por definición, queda excluida de culpas (Wallach y Allen, 2009, pp. 22-26, 136, 207). Luego, si (1) el humano al comando ha seguido todo deber, teniendo pues el derecho a conducir tal vehículo, el accidente no ha podido ser su culpa. Ergo, el grado de falibilidad o riesgo legal que se aceptaba en el automóvil vendido, de ser este desmedido, indica una culpa en (2) el diseñador o su equipo. De fallar dicho margen de error, la culpa es del fabricante (que no del científico estudioso de leyes sobre las que el tecnólogo diseñador construye).

Pero eso también depende de la legislación y el mercado –desde la nomoética–: el sistema de regulación de venta, estándares legales que se apliquen en el comercio, etc. Entonces la culpa sería del tecnólogo fabricante y (3) del vendedor solo si y solo si no se ha seguido la ley. En el caso haberse seguido, y sin presiones de mercado, la culpa ulterior es (4) de quien en su derecho de legislar correctamente no ha cumplido con sus obligaciones y ha dilatado en exceso lo que era recomendable (seguramente, recomendación de científicos). A parte, (5) el propietario reconocido del vehículo también puede haber incumplido algún deber comunicado del que sea así responsable para usar tal artefacto –desde la tecnoética– como: no actualizar el software, no cargar bien el vehículo, tocar el volante durante un trayecto (en caso de automatización elevada) o, ya siendo sujeto agente, usar mal el coche –su responsabilidad–, llevar el vehículo por zonas por las cuales no fue diseñado, etc. Implicando así moralmente, por equidad, que tampoco tiene derecho a tal vehículo. Las actualizaciones además, pueden circunscribirse solo en el ámbito de ciertos aspectos de la automatización, que suelen ser graduados⁸. Y de no ser avisado el propietario de tales deberes, la culpa sería, o de nadie si es por un infortunio, o bien “lo sería” del sistema comunicante, como herramienta fallida. Sobre esto; en efecto, de haber un error azaroso en la

⁸ Cabe mencionar que pueden existir varios grados de automatización (se suelen hablar de 0 a 5 grados de automatización). A lo largo del texto se ha supuesto, en general y con tal de simplificar, un intervalo binario 0-1, sin graduaciones: 'o automatizado o no automatizado', situación un tanto idealizada para poder plantear mejor todos dilemas y comparaciones morales en las aristas más extremas de estos asuntos.

herramienta de comunicación, no podría adjudicarse una culpa con rigor ni justicia (Knight, 2013): cabría revisar los canales de información habituales en la vida cotidiana –desde la bioética– lo cual inevitablemente borra toda perspectiva de decisión al sujeto.

Por otro lado, partiendo de la opción del sacrificio escogido, se puede plantear la posible reparación civil de automóviles diseñados y comprados deliberadamente para proteger a terceros peatones, respecto al vehículo y a sus ocupantes en un infortunio imprevisto por encima del riesgo legal. De este modo se cubriría, en caso de choque y supervivencia de las partes, los daños producidos en el transporte que requieran una erogación económica considerable. Tal reparación civil puede ser jurídicamente definida, quizás, estableciendo un sistema de seguros estatalmente planificado (se evitan así enriquecimientos de clase arbitrarios). Este desarrollo tiene relevancia nomoética, aún sin abandonar la tecnoética.

Una síntesis de la discusión sobre la automatización de vehículos⁹, podría ser (Tabla 1):

Opción de producción y venta	Programación moral según opción
Venta de vehículos sin control propio	Programados para el sacrificio propio
	Programados para la muerte de terceros
Venta de vehículos bajo control propio, con el principio de autolimitación aplicado (O bien un control facultativo muy estrecho; por ejemplo, al tocar el volante, aunque tal opción usada por la empresa T.E.S.L.A., grado 2 de IA automatizada, es problemática; Godwin, 2021)	Programable en todo momento para el sacrificio propio o no (posición ideal)
	Programable en todo momento para la muerte de terceros o no (posición ideal)
¿Tercera vía hipotética?	Programado/programable para optimizar bienestar propio y/o de terceros

Tabla 1: Opciones para el desarrollo ético equitativo de producción de IA en automóviles.

En el presente se están usando *de facto* vehículos automatizados con IA de tipo militar, drones de ataque (modelo Zala Aero Kub) en la guerra Ucrania-Rusia (Army Recognition Group, 2022). Esto intensifica el debate sobre el rol de tales tecnologías en conflictos¹⁰.

⁹ Otro tema que cabría discutir más extensamente sobre la ética de la automatización de IA en vehículos, en sus consecuencias sociales, podría ser el hipotético mercado negro que se daría de existir personas con poder que buscaran truncar vehículos en posesión de terceros para que en dilema priorizase siempre su vida.

¹⁰ Fragmento añadido en la última revisión de este documento, en vista de la reciente invasión a Ucrania por parte de Rusia a finales del mes de febrero en 2022, de irremediable y dramática actualidad. Agradezco aquí a Antonio Diéguez Lucena haberme dado acceso a esta información.

5.3. Algunas aclaraciones metaéticas pertinentes

La propuesta de ética sugerida, aplicable a los dilemas con vehículos autónomos, no sería pues de ningún modo utilitaria, pero sí consecuencialista y técnica. La preceden sistemas teóricos análogos (Bunge, 1989) (Romero, 2018, pp. 71-79): la ética como práctica ajustada al conocimiento moderno y sus vertientes, cuyo fin definido es primordialmente –aunque tampoco solamente– satisfacer valores objetivos, sobre necesidades¹¹. Y se logra la eficacia moral a partir de teorías aproximadas, realistas, que junto a una axiología precisa y no-arbitraria (Teixidó; 2019, 2021) ayuden a trazar prescripciones. No cabe confundir esta metodología con basar los valores y/o normas en sólo enunciados de hecho a modo justificación, que supone cometer la falacia ser-deber ser (Hume, 2005) –o su inverso, derivar enunciados de hecho a partir de los normativos; la falacia moralista–. En este caso se deducen juicios de valor y normas sobre enunciados de hecho, y junto a un sistema de valor ya tomado. Por lo que, al existir valoración de principio a fin no hay salto lógico (Zavadivker, 2004) y sí un asiento fáctico junto a una axiología rigurosa. Si se busca alguna justificación ad infinitum en tal axiología sobre por qué sería no-arbitrario priorizar los valores de necesidades biopsicosociales, solo cabe decir que no se toma un fundacionismo en ética, como tampoco en ninguna ciencia, tecnología u otra filosofía: no se requieren de unas justificaciones últimas, basta con partir de hipótesis sobre rasgos impersonales de las cosas. Y la hipótesis aquí, en resumen, es buscar provisionalmente el bienestar en necesidades humanas reales como ‘lo bueno’¹² en ética. Sin restar a valores deseables pero secundarios, y para todo evento moral, que se entenderá como sistémico¹³.

Además, se vuelve a reiterar, como también previamente se ha mencionado: cabría trazar y justificar todavía una moral articulada bajo una norma rectora superior N , *i. e.* máxima de justicia, que vertebré todo derecho y deber en general de un modo más preciso. Y que pueda evaluarse como el más razonable, eficiente y pragmático en cuanto a bienestar real, en este y otros temas morales (constatable mediante investigaciones como el proyecto de contrastación empírica y de “cálculo” esbozado en la anterior sección –que, como se ha afirmado, esta vez es objetivo y previo a toda acción específica, realizado para indagar en evaluar sistemas morales sólo, y no para computar *in situ* o *in media res* dentro de un evento moral–). Se parte de la existencia de derechos

¹¹ Necesidades que, como se ha indicado al mentar las ciencias que las estudian, no se reducen tampoco de forma simplista a lo biológico o a lo biomédico, si no que abarcan lo psicológico y lo social.

¹² De este modo se logra evadir además la también archiconocida falacia naturalista (que puede pensarse que contiene también la falacia ser-deber ser), pues se define algo normativo por otros aspectos normativos.

¹³ Aunque el sistemismo como filosofía ontológica, trasladada a la sociedad y a la ética en particular, implica bastante más aparte de considerar a las personas en niveles sociales (individuo y colectivo) y en variables diversas (biológicas, políticas, culturales, etc.), como por ejemplo implica también: trazar las diferentes dinámicas sociales y sus mecanismos, localizar sus componentes, su estructura, etc. Distribuir normas a individuo y colectivo en derechos-deberes es también algo sistémico, pero no es exclusivo del sistemismo.

y deberes en los sistemas sociales, acorde tanto con una ontología sistémica como con el derecho jurídico actual. Sólo se ha asumido aquí la equidad como mejor máxima a modo de hipótesis, una muy provisional, que podría para nada ser tal y requiere de más investigación. Ahora, también se insiste en que cualquier proyecto ético fértil sí deberá partir como base hipotética pero más rigurosa, de, al menos, las necesidades y su trato en normas de valor como sistema moral mínimo.

6. Consideraciones adicionales

Toda investigación ética y moral sobre la automatización de vehículos debería ir en el camino en que la IA esté programada para lograr el mayor nivel de bienestar entre las posibilidades. Cabe considerar siempre derechos y responsabilidades sin vagos cálculos de felicidades, que también a veces se tratan de forma demasiado idealizada en falsos dilemas. De escoger bajo dilema con una IA configurada, debe hacerse de forma limitada: pudiendo elegir directamente si sacrificar o privilegiar en todo momento (v. Tabla 1). Y cabe teorizar siempre en consonancia a los avances científicos disponibles con tal de enriquecer adecuadamente una filosofía seria (Mosterín, 2013, p. 6) (Fig. 5).

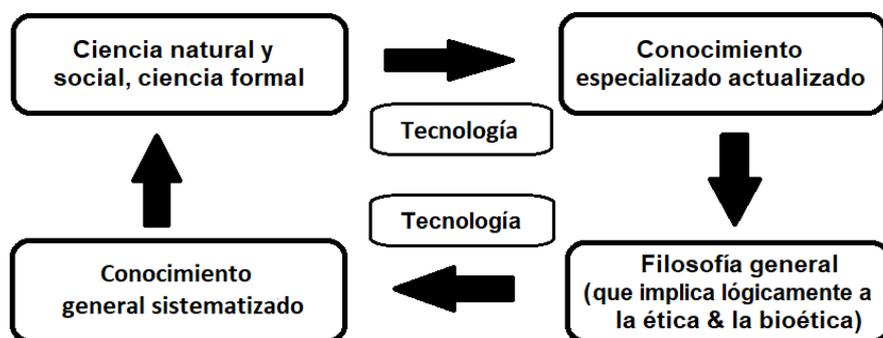


Figura 5: Influencias causales básicas (flechas) entre las ciencias y la filosofía, en el avance del saber concreto y general, junto a las áreas colindantes de la tecnología, que ayudan en el proceder de las ciencias y la filosofía en generar conocimientos. Esto no es de extrañar, pues la ética se ha concebido aquí como una técnica filosófica.

En disciplinas éticas, una propuesta sólida tenderá a proveer precisión a las reglas que teóricamente deban servir para maximizar el bienestar en derechos y deberes, consistentes con las ciencias actuales, así como tenderá a borrar semánticamente los personalismos (Bunge, 2008, pp. 146-149) en favor de mayor realismo. Esto, se repite, no debe sugerir la ingenua idea de hacer cálculos en medio de una catástrofe para saber exactamente qué hacer: las teorías éticas en su moral deben ya promover reglas o normas en torno a una máxima hipotética N , de eficacia previamente evaluada y aún si tentativa, y deben servir para obrar allí donde sucedan problemas

y dilemas morales locales (aunque se tiene en cuenta que no es habitual, en el primer mundo¹⁴, afrontar muchas situaciones extremas).

La automatización con IA de transportes, y su problemática ética (tecnoética sobretodo), sirve en definitiva para señalar factores que habitualmente son bastante omitidos. Puede usarse para cuestionar la eficacia de las teorías éticas predominantes hoy en día, obsoletas para los problemas modernos los de la automatización y análogos (edición genética, tecnología blockchain, etc.), cuando no totalmente desconectadas de la realidad material.

La problemática puede complicarse si se combinan temas, desplazando el eje más a la bioética en la ética de la automatización: ¿qué cabría hacer ante vehículos automatizados que transporten, por ejemplo, alimentos transgénicos con nutrientes clave, o vacunas antivirales de RNAm (ácido ribonucleico mensajero), ahora muy manidas en la pandemia de COVID-19 (Kaur y Gupta, 2020), como carga que pueda peligrar en un accidente? Aceptando su seguridad y eficiencia respectiva (Pellegrino et al., 2018) (Pilishvili et al., 2021), y si se imagina aquí que tanto el antiviral como la modificación genética realizada en los alimentos tuvieran capacidad para sanar de facto cierto aspecto en una población, que en breves perecerá sin dichos productos, ¿priorizaría en un posible accidente la integridad del camión o avión, por necesidades mayoritarias, frente a unas pocas víctimas –los transportistas o los vehículos colindantes–? ¿O se deberían salvar las últimas, lo que implica destruir el transporte y su carga? ¿Cómo decidiría la IA del camión de alimentos transgénicos? Y respecto la nomoética, ¿qué legislación usar para regular tales eventos?

De hecho ya suceden eventos a muy poco tocar de tal dilema, con el COVID-19 respecto a automatizar el sistema de refrigeración (a -70 °C) de vacunas RNAm (Gore et al., 2021) (Ríos, 2021). Evitar este tipo de daños es crucial y de impacto cosmopolita en pandemias.

Aparte, una cosa es previsible (o al menos, es fácil de hipotetizar de cara al futuro tecnológico cercano) y no siempre se considera. La respuesta a los dilemas sobre esta área de estudio, sea entre deberes o derechos en una disposición u otra, pasará probablemente por tener en cuenta que las nuevas generaciones de automóviles con IA reconocerán, no solo su propio estado, sino el estado simultáneo de otros vehículos con IA. Posiblemente se hará uso de sistemas de red veloces (Diario Siglo XXI, 2019) por satélite u onda corta, conectadas permanente o semi-permanentemente (Fig. 6), formando un sistema de IA por canales de información con funciones

¹⁴ Posiblemente, y de omitir los deberes no-dilemáticos, tales como: pagar los impuestos proporcionales respectivos, atender a heridos en un accidente en el que se puede ayudar y nadie más puede, llamar a emergencias ante ciertas alertas percibidas y no contestadas... o llevar mascarilla y vacunarse ante una pandemia mundial; las situaciones morales más extremas, de dilema, que se pueden vivir hoy en Occidente suelen ser a lo sumo intentar obrar con bastante riesgo propio en un posible crimen, ataque terrorista o catástrofe para ayudar a otros, intervenir o no en defensa de alguien en sus derechos ante una agresión, etc.

(representables en grafos), logrando datos radar con tecnología LiDAR *-Laser Imaging Detection and Ranging-* (Castaño et al., 2018).

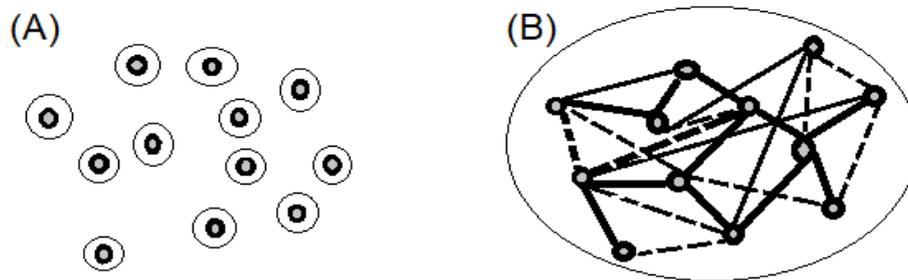


Figura 6: (A) grupo aleatorio de vehículos sin interacción, con LiDAR activados (áreas circulares de radar); (B) sistema idealizado de vehículos interconectados en relaciones estrictas (línea continua: canal abierto) y facultativas (línea discontinua: canal factible), formando sus LiDAR activados un radar grupal (área circular grande de radar colectivo).

Este sistema, hipotéticamente como una red, registraría con LiDAR posición, estado y entorno de diversos vehículos en un radio dado, a cada momento. Acabarían así los automóviles “individuales” (aunque tampoco esto llevaría a un holismo tecnológico): a la hora de conducir, se constituirían sistemas coordinados de vehículos inteligentes que podría predecir fallos, prevenir conflictos, etc. Aunque, si no se idealiza demasiado, es probable que para su funcionamiento óptimo se deban usar algoritmos con gran capacidad de cómputo, y por ello quizá se deba contar con una unidad auxiliar de IA centralizada.

Pero ciertamente una tecnología de esta índole permitiría también una conectividad con *smartphones* personales y demás sistemas operativos, de uso diario y rápida innovación, como sucede ahora con las realidades aumentada y virtual (Redacción El Heraldo, 2017), entrelazando toda la información. ¿Cómo pensar el futuro donde se deba preocupar uno, al ir a comprar por ejemplo, no solo de los alimentos que se necesiten consumir de forma sana, si no de los requerimientos energéticos, estructurales y etc. que la IA indique en el reloj portable con wifi que requiere el automóvil? En estos supuestos, la tecnoética se desplaza a la bioética y a la nomoética cada vez más. No se desea aquí profundizar más, solo se sugieren ideas acerca de la IA en vehículos, que cabe reflexionar (si no se ha hecho ya): pues eventualmente, de suceder, modificarán nuestras formas de afrontar dilemas y eventos en ética de la automatización; pero también modificarán la vida cotidiana misma.

Ya que, si bien se puede conjeturar y “jugar” a los dilemas partiendo de teorías morales y éticas clásicas, con el conocer si un transeúnte aleatorio muere o no, o con saber cuasi mágicamente quien posee x años o cierta salud; ¿por qué no considerar antes la realidad cercana de que muchos

vehículos inteligentes puedan operar como sistema, aprovechando todas las herramientas electrónicas factibles en este problema, y bajo una ética puntera?

7. Conclusiones provisionales

En síntesis: es razonable sostener que se requieren en la actualidad de nuevos esfuerzos en ética ante los avances próximos en campos como la ética de los artefactos y vehículos automatizados, su problemática, si se quieren tratar con rigor y eficacia. Esto, claro, de ningún modo significa despreciar la tradición en filosofía hasta nuestros días, y de hecho diversos problemas y posiciones tratadas se entroncan en varios clásicos (Romero, 2017).

Aquí, sobre propuestas previas de otros autores y sistemas, se ha sugerido finalmente un proyecto técnico quizás pretencioso por su énfasis en el rigor, pero falsable y que deja de idealizar conductas y fines. Un proyecto que no demanda cálculos imposibles ni pide actos rigurosos pero muy dogmáticos. Un esbozo teórico que evalúa consecuencias de acciones (aunque no en exclusivo, y no se define “bueno” solo por últimas consecuencias; Carcacia, 2021, p. 82) en función de valores objetivos, todos ellos biopsicosociales¹⁵ en los diferentes sistemas sociales, trabajando sobre normas, y considerando siempre factores contingentes como riesgos y márgenes de error en los diferentes actos de derecho y deber. Y considerando los márgenes de error, no sólo en las posibles normas morales a pautar, si no en los mismos fines y los valores objetivos a perseguir, todos pensados como falibles y en último término, solo provisionales hasta adquirir nuevos conocimientos.

Como se ha esgrimido, el presente proyecto ético y su uso en la problemática de vehículos automatizados con IA, sobre todo de tipo turismo, tema usado al mismo tiempo para criticar los sistemas éticos y morales clásicos, no es original y continua una corriente filosófica planteada años ha (Bunge, 2002) (Quintanilla, 2017). Se tiene constancia de que otros proyectos de teoría ética, que dan actualizada cuenta de la responsabilidad sobre la tecnología con IA en base a las ciencias y la equidad cívica o social, se han sugerido en los últimos años. Y han sido desarrollados con mayor extensión, aunque también con menos sistemismo y menor rigor teórico (por ejemplo,

¹⁵ Esto permite entender que, aunque salud y supervivencia sean las necesidades de valor básico a preservar (aunque no exclusivo), y a fortiori, promover; no solamente la prioridad se da en lo biológico sino también sobre lo psicológico y lo social. Una persona que mantenga su biovalores pero para ello deba escoger sacrificar sus psicovalores en unas circunstancias extremas, puede llegar a estar justificada objetivamente tanto en seguir priorizando sus biovalores como ya priorizar sus psicovalores, incluso si eso implica la extinción de una de ambas clases de valores o todas (si los psicovalores extinguen los demás valores, se acabaría con todos ellos ya que moriría el propio organismo, pero se lograría evitar un perjuicio continuado: un suicidio, o eutanasia si un tercero le asiste, pudiendo estar de este modo justificado en ambos casos).

véase: Rodríguez, 2020). Pero, es más, se postula la presente posición como legítima y con posible mayor validez (y eficacia moral) que sus alternativas clásicas o análogas, por lo que se reitera su posibilidad y los aportes que esta puede lograr sobre sus alternativas, en problemas morales modernos. Queda así en manos de los filósofos éticos y morales reflexionar sobre los cambios necesarios y qué línea de investigación continuar para adaptarse a la actualidad. Una actualidad donde tecnología como los vehículos automatizados, si bien rara vez da una solución (National Geographic Society, 2020), sí puede ser de gran ayuda –y más se valora esto ante la pandemia actual–. Siempre con los respectivos tratos tecno, nomo y bioéticos, aunque sea la tecnoética la que encabece el estudio ético de la tecnología. Y dónde, con seguridad, cabe repensar el contenido y los horizontes de la filosofía ética.

Bibliografía

- ◆ Anscombe, G. E. M. (2006). *Filosofía Moral Moderna*. México: UNAM Instituto de Investigaciones Filosóficas. (Versión original 1958).
- ◆ Aristóteles (2005). *Ética a Nicómaco*. Madrid: Alianza Editorial. (Versión original 349 a. C.).
- ◆ Army Recognition Group (14 marzo 2022). Russia using Kalashnikov Zala Aero Kub suicide drones in Ukraine. *Army Recognition*. Recuperado de <https://bit.ly/38bfDRT> [Consulta: 14 octubre 2022].
- ◆ Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J. y Iyad Rahwan, I. (2018). The Moral Machine experiment. *Nature*(563), 59–64. doi: 10.1038/s41586-018-0637-6.
- ◆ Bentham, J. (1836). *Deontología o ciencia de la moral*. Valencia: Mallen y Sobrinos.
- ◆ Berardo, M. G., Irureta y V. A. (2008). *Influencia de la correcta evaluación del tiempo de percepción y reacción*. Universidad Nacional de Córdoba. Recuperado de <https://bit.ly/3uMxEx> [Consulta: 14 octubre 2022].
- ◆ Bostrom, N. (2003). When machines outsmart humans. *Futures* 35 759-764. doi: 10.1016/S0016-3287(03)00026-0.
- ◆ Boucher, P. (2018). ¿Debemos temer a la inteligencia artificial? *EPRS, Unidad de Previsión Científica (STOA) PE 581.948*. doi: 10.2861/61195.
- ◆ Boullosa, N. (28 mayo 2014). ¿Máquinas morales? Ética "automatizada" en coches autónomos. *Faircompanies*. Recuperado de <https://bit.ly/3Ka0m1K> [Consulta: 14 octubre 2022].
- ◆ Bunge, M. (1972). *Ética y ciencia* (2ª Ed.). Buenos Aires: Siglo XX.
- ◆ _____ (1989). *Treatise on Basic Philosophy, Vol. VIII: Ethics*. Dordrecht: D. Reidel Publishing Company.
- ◆ _____ (2002). *Ser, saber, hacer*. México: Paidós.
- ◆ _____ (2004). *La investigación científica* (3ª Ed.). México: Siglo XXI editores.
- ◆ _____ (2008). *Tratado de Filosofía: Semántica I, sentido y referencia*. Barcelona: Editorial Gedisa.
- ◆ _____ (2010). *Las pseudociencias ¡vaya timo!*. Pamplona: Laetoli.
- ◆ _____ (2012). *Filosofía de la tecnología y otros ensayos*. Lima: Universidad Garcilaso de la Vega.

- ◆ Carcacía, I. (2021). *El cambio como idea filosófica*. Research Gate. Recuperado de <https://bit.ly/3iUSc1k> [Consulta: 14 octubre 2022].
- ◆ Castaño, F., Beruvides, G., Villalonga, A. y Haber, R.E. (2018). Self-Tuning Method for Increased Obstacle Detection Reliability Based on Internet of Things LiDAR Sensor Models. *Sensors*, 18(5), 1508. doi: 10.3390/s18051508.
- ◆ Clouser, D. y Gert, B. (1990). A Critique of Principlism. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, 15(2), 219-36. doi: 10.1093/jmp/15.2.219.
- ◆ Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*(5) 536-544. doi: 10.1038/s41564-020-0695-z.
- ◆ Cortina, A. (2021). *Ética cosmopolita: Una apuesta por la cordura en tiempos de pandemia*. Barcelona: Paidós Estado y Sociedad.
- ◆ Del Río, P. (2018). La Simetría de los Derechos Humanos son los Deberes. En Proner, C., Olasolo, H., et al. *70º Aniversario de la Declaración Universal de los Derechos Humanos*. (pp. 23-26). Tirant lo Blanch Editorial. Recuperado de <https://bit.ly/3JSH0hE> [Consulta: 14 octubre 2022].
- ◆ Dennett, D. (1979). Artificial Intelligence as Philosophy and as Psychology, Philosophical Perspectives in Artificial Intelligence. M. Ringle, ed., Atlantic Highlands, NJ: Humanities Press, 57-80.
- ◆ DGT (3 septiembre 2021). 191 personas han perdido la vida en las carreteras este verano. *Revista DGT*. Recuperado de <https://bit.ly/36CAddP> [Último acceso: 14 octubre 2022].
- ◆ DGTtest (2021). Ejemplos de Test de la DGT. *Web DGT*. Recuperado de <https://bit.ly/3NBUArU> [Consulta: 14 octubre 2022].
- ◆ Diario Siglo XXI (22 febrero 2019). Las redes 5G, el Blockchain y la IA serán las principales tendencias tecnológicas de 2019, según BDO. *Diario Siglo XXI*. Recuperado de <https://bit.ly/36NzSop> [Consulta: 14 octubre 2022].
- ◆ Diéguez, A. (2012). La opción naturalista, una respuesta a Francisco Soler. *Naturaleza y Libertad. Revista de estudios interdisciplinarios*, Número 1.
- ◆ Diéguez, A. y Atencia, J. M. (Ed.) (2014). *Naturaleza Humana y Animal*. Madrid: Siglo XXI.
- ◆ Ferrater, M., y Cohn, P. (1981). *Ética aplicada: del aborto a la violencia*. Madrid: Alianza Unviersidad.
- ◆ Flórez, M. (3 enero 2019). El dilema de los coches autónomos. *El País*. Recuperado de <https://bit.ly/36KEfke> [Consulta: 14 octubre 2022].
- ◆ Fraser, D. (1999). Animal ethics and animal welfare science: bridging the two cultures. *Applied Animal Behaviour Science* 65: 171-189. doi: 10.1016/S0168-1591(99)00090-8.
- ◆ Fuentes, V. (4 diciembre 2021). España introduce el coche autónomo en la nueva Ley de Seguridad Vial, y hay novedades para Tesla y Mercedes-Benz. *Motorpasión*. Recuperado de <https://bit.ly/3gnShge> [Consulta: 14 octubre 2022].
- ◆ Godwin, C. (23 abril 2021). Tesla's Autopilot 'tricked' to operate without driver. *BBC News*. Recuperado de <https://bbc.in/3T4r3cX> [Consulta: 14 octubre 2022].
- ◆ Gore, M., y Emsellem-Rope, C. (31 diciembre 2021). Logistic face a bumpy road in the COVID-19 vaccine rollout. *HFW Briefings*. Recuperado de <https://bit.ly/3euOljX> [Consulta: 14 octubre 2022].
- ◆ Gutiérrez-Soriano, J. R., Ortiz-León, S., Follioux, C., Zamora-López, B., y Petra, I. (2017). Funciones mentales: neurobiología. *Facultad de Medicina UNAM*.

- ◆ Haack, S. (2017). *Scientism and Its Discontents*. Rounded Globe. Recuperado de <https://bit.ly/3LuAloP> [Consulta: 14 octubre 2022].
- ◆ Hueso, C. (2020). SyRI, ¿a quién sanciono? Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020, *La Ley Privacidad, Wolters Kluwer n.º 4*, mayo. Recuperado de <https://bit.ly/3VxHPTe> [Consulta: 14 octubre 2022].
- ◆ Hume, D. (2005). *Tratado sobre la naturaleza humana, Libro III Parte I* (4ª Ed.). Madrid: Tecnos. (Versión original 1740).
- ◆ Kahane, G. (2015). Sidetracked by trolleys: Why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment, *Soc Neurosci* 10(5): 551–560. doi: 10.1080/17470919.2015.1023400.
- ◆ Kant, I. (2003). *Crítica de la razón práctica*. Buenos Aires: Editorial Losada S.A. (Versión original 1788).
- ◆ _____ (2007). *Fundamentación metafísica de las costumbres* (1ª Ed.). San Juan: Pedro M. Rosario Barbosa. 17. (Versión original 1785).
- ◆ Kaur, S. P. y Gupta, V. (2020). COVID-19 Vaccine: A comprehensive status report. *Virus Research*(288), 198114. doi: 10.1016/j.virusres.2020.198114.
- ◆ Keeling, G. (2019). Why Trolley Problems Matter for the Ethics of Automated Vehicles. *Science and Engineering Ethics* 26: 293–307. doi: 10.1007/s11948-019-00096-1.
- ◆ Knight, C. (2013). Luck Egalitarianism. *Philosophy Compass*. Recuperado de <https://bit.ly/3LtCkzc> [Consulta: 14 octubre 2022].
- ◆ La comunidad del taller redacción (12 noviembre 2018). ¿A quién matar? Los principios morales del coche autónomo. *La comunidad del taller*. Recuperado de <https://bit.ly/38iBly8> [Consulta: 14 octubre 2022].
- ◆ Laird, J. E. y van Lent, M. (2001). Human-Level AI's Killer Application, Interactive Computer Games. *AI Magazine*(22) 2. doi: 10.1609/aimag.v22i2.1558.
- ◆ Lorenzo, A. y Menchaca, M. (1999). Biomecánica del accidente de tráfico. *Medynet*. Recuperado de <https://bit.ly/3Lqi7dq> [Consulta: 14 octubre 2022].
- ◆ Margot, J. (2007). La Felicidad. *Praxis Filosófica*. Recuperado de <https://bit.ly/3JYmz2W> [Consulta: 14 octubre 2022].
- ◆ Mill, J. S. (2014). *El utilitarismo* (1ª Ed.). Madrid: Alianza Editorial. (Versión original 1861).
- ◆ Moore, G. (1922). *Principia Ethica*. Cambridge: Cambridge University Press. (Versión original 1903).
- ◆ Moreno-Romero, C. (2020). Desentendimiento Moral y Atribución de Culpa: Encuentros y Desencuentros en el Estudio de la Cognición Moral. *Rev. colomb. psicol.*(29), 125-141. doi: 10.15446/rcp.v29n1.76536.
- ◆ Mosterín, J. (2000). *Conceptos y teorías en la ciencia*. Madrid: Alianza Editorial.
- ◆ _____ (2013). *Ciencia, filosofía y racionalidad*. Barcelona: Editorial Gedisa.
- ◆ Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.
- ◆ National Geographic Society (28 enero 2020). Are Genetically Modified Crops the Answer to World Hunger? *National Geographic*. Recuperado de <https://bit.ly/36FvSGK> [Consulta: 14 octubre 2022].
- ◆ Nozick, R. (1974). *Anarchy, State, and Utopia*. Cornwall: Basic Books.

- ◆ Pellegrino, E., Bedini, S., Nuti, M., y Ercoli, L. (2018). Impact of genetically engineered maize on agronomic, environmental and toxicological traits: a meta-analysis of 21 years of field data. *Scientific Reports*(8), 3113. doi: 10.1038/s41598-018-21284-2
- ◆ Pilishvili, T., et al. (2021). Effectiveness of mRNA Covid-19 Vaccine among U.S. Health Care Personnel. *The New England Journal of Medicine*. doi: 10.1056/NEJMoa2106599.
- ◆ Quintanilla, M. A. (1983). *A favor de la Razón*. Madrid: Taurus.
- ◆ _____ (2017). *Tecnología: un enfoque filosófico* (2ª Ed.), México: Fondo de Cultura Económica.
- ◆ _____ (2019). *Tecnologías entrañables*. Madrid: La Catarata.
- ◆ Rand, A. (2007). *La rebelión de Atlas*. Buenos Aires: Grito Sagrado. (Versión original 1957).
- ◆ Rawls, J. (2003). *Justicia como equidad* (3ª Ed.). Madrid: Tecnos.
- ◆ _____ 2006). *Teoría de la justicia* (2ª Ed.). México: Fondo de Cultura Económica. (Versión original 1971).
- ◆ Redacción El Heraldo (2 octubre 2017). Realidad aumentada será algo cotidiano en 'smartphones'. *El Heraldo*. Recuperado de <https://bit.ly/3CxCRNG> [Consulta: 14 octubre 2022].
- ◆ Ríos, J. D. (16 abril 2021). Invima confirma que se perdieron 693 vacunas COVID por errores en la cadena de frío. *Blu Radio*. Recuperado de <https://bit.ly/36Dl3Fa> [Consulta: 14 octubre 2022].
- ◆ Rodríguez, A. L. (2020). *Inteligencia artificial responsable. Humanismo tecnológico y ciencia cívica*. Valencia: Universitat de València.
- ◆ Romero, G. E. (2017). La filosofía científica y los límites de la ciencia. *Revista Científica Estudios e Investigaciones*, 6(1), 97-103. doi: 10.26885/rcei.6.1.97.
- ◆ _____ (2018). *Scientific Philosophy*. Cham, Suiza: Springer.
- ◆ _____ (2020). ¿De qué se trata la matemática? *Scientia in Verba Magazine*(6), 60-64. doi: 10.5281/zenodo.3841921.
- ◆ Rothbard, M. (2002). *The Ethics of Liberty*. Nueva York: New York University Press. (Versión original 1982).
- ◆ Russell, S. (1996). Rationality and Intelligence. *Computer Science Division*(4), 57-77. doi: 10.1016/S0004-3702(97)00026-X.
- ◆ Russell, S. y Norvig, P (1995). *Artificial Intelligence, A Modern Approach*. New Jersey: Prentice Hall.
- ◆ Sánchez, A. (2017). Tecnoética y Bioética. *Scribd*. Recuperado de <https://bit.ly/3Nwe55n> [Consulta: 14 octubre 2022].
- ◆ Schopenhauer, A. (2016). *Mundo como voluntad y representación I* (2ª Ed.). Madrid: Editorial Trotta. (Versión original 1819).
- ◆ Shujman, A., Erramuspe, P. y Amico, V. (2010). Dilemas y problemas morales. *Escuelas que aprenden, UNICEF*. Recuperado de <https://bit.ly/3wTD78E> [Consulta: 14 octubre 2022].
- ◆ Sidgwick, H. (1981). *The Methods of Ethics* (7ª Ed.). Indianapolis, Estados Unidos: Hackett Publishing Company (Versión original 1874).
- ◆ Singer, P. (1995). *Ética Práctica* (2ª ed.). Cambridge: Cambridge University Press.
- ◆ Smith, C. (2010). *What Is a Person? Rethinking Humanity, Social Life, and the Moral Good from the Person Up*. Chicago: University of Chicago Press.
- ◆ Teixidó, O. (2019). Sobre la axiología de Gustavo E. Romero y Mario Bunge. *Scientia in verba magazine*(5), 57-67. doi: 10.5281/zenodo.3842243 (Versión actualizada en; Teixidó, O., 2022 doi: 10.5281/zenodo.6582745).

- ◆ _____ (2021). Necesidades, valores y normas desde una filosofía científica. *Universidad verdad*, 1(78), 120-135. doi:10.33324/uv.v1i1.396.
- ◆ Thomson, J. (1985). The Trolley Problem, *Yale Law Journal* (94), 1395-1415. doi: 10.2307/796133.
- ◆ Valesolo, P., y DeSteno, D. (2006). Manipulations of Emotional Context Shape Moral Judgment, *Psychol Sci*, 17(6): 476-7. doi: 10.1111/j.1467-9280.2006.01731.x.
- ◆ Wallach, W y Allen, C. (2009). *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press.
- ◆ Youssef, F., et al (2012). Stress alters personal moral decision making, *Psychoneuroendocrinology*(37), 4, 491-498. doi: 10.1016/j.psyneuen.2011.07.017.
- ◆ Zavadivker, N. (2004). *Una ética sin fundamentos*. Tucumán, Argentina: Facultad de Filosofía y Letras, Universidad Nacional de Tucumán.

Fecha de recepción: 6 de enero de 2022

Fecha de aceptación: 20 de octubre de 2022

Fecha de publicación: 14 de febrero de 2023