

# WorkétlcA

Foro para la aplicación ética de la IA en el trabajo

❖ Fundación Máshumano

Hacer. Crecer. Ser.

Nota de conocimiento 7.

## Vehículos autónomos e inteligencia artificial

¿Puede una máquina ser un agente moral?

La inteligencia artificial está transformando progresivamente muchos ámbitos de nuestra vida y del trabajo. Uno de ellos es la movilidad autónoma, que plantea preguntas relevantes no solo desde el punto de vista tecnológico, sino también ético, social y laboral.

En una nueva sesión del foro **WorkétlcA** con nuestro Consejo Abierto, hemos reflexionado sobre este reto a partir del caso de los **vehículos autónomos**. A primera vista, este avance puede asociarse a cuestiones como la eficiencia del tráfico, la reducción de la siniestralidad o la mejora de la autonomía de personas que no pueden conducir por sí mismas. Pero también nos sitúa ante una pregunta de fondo: **¿puede una máquina tomar decisiones morales?**

### Análisis del caso

La conducción autónoma se define como la capacidad de un vehículo para realizar todas o parte de las tareas de conducción de forma sostenida, con distintos grados de intervención humana. En su nivel más avanzado, el vehículo tomaría el control completo y se prescindiría del conductor.

Esta posibilidad tiene una primera consecuencia evidente en el mundo del trabajo: el impacto sobre los empleos vinculados a la conducción profesional. La automatización de ciertas tareas puede transformar profundamente estos perfiles y obliga a pensar cómo acompañar esos cambios desde una mirada responsable.

Sin embargo, **el interés ético del caso va más allá del impacto laboral**. Los vehículos autónomos nos obligan a preguntarnos **qué ocurre cuando una máquina debe actuar en situaciones de riesgo**, especialmente si hay vidas humanas en juego.

## El experimento Moral Machine

Para abordar esta cuestión, se tomó como referencia **Moral Machine**, un experimento desarrollado en 2016 por investigadores del MIT Media Lab, inspirado en el conocido dilema del tranvía.

El experimento planteaba distintos escenarios en los que un vehículo autónomo debía priorizar entre posibles víctimas: pasajeros o peatones, personas jóvenes o mayores, humanos o animales, entre otros supuestos. Millones de personas de más de doscientos países participaron registrando sus preferencias.

Los resultados mostraron que las respuestas variaban ampliamente según el país y estaban relacionadas con factores culturales y económicos. Esto abre preguntas complejas: ¿deberían los vehículos autónomos adaptarse a las preferencias morales de cada país? ¿A las de cada fabricante? ¿A las de cada usuario? ¿Es posible consensuar criterios comunes?

También aparecen riesgos importantes: estereotipar una cultura, congelarla como si no evolucionara, excluir las diferencias individuales o no resolver adecuadamente tensiones entre valores en un mundo cada vez más global.

## No hay una “respuesta correcta” automática

Una de las claves del caso es que **Moral Machine no ofrece una respuesta correcta sobre cómo debe actuar un vehículo autónomo**. Más bien muestra cómo pensamos las personas cuando nos enfrentamos a dilemas morales.

Nuestras decisiones no son neutras: están influidas por sesgos, emociones, normas culturales interiorizadas y distintas formas de valorar lo que está en juego. Elegir implica siempre tomar una posición ética.

Por eso, estas cuestiones no deberían quedar circunscritas únicamente a ingenieros y fabricantes. Los resultados de Moral Machine **podrían ayudar a tecnólogos a pensar con más profundidad sobre la ética de la IA**. El diseño de sistemas de inteligencia artificial que pueden afectar a la vida de las personas exige una reflexión más amplia, capaz de incorporar distintas perspectivas y de deliberar sobre los valores que queremos proteger.

## Cálculo no es decisión moral

Un vehículo autónomo no “elige” moralmente en sentido humano. Lo que hace es ejecutar cálculos según la forma en que ha sido programado, por ejemplo, para minimizar daños o priorizar determinados factores.

Pero esa programación no es neutral. El peso que se dé al pasajero, al peatón, a los daños materiales o a otros criterios ya implica una interpretación previa. Por tanto, aunque la máquina no tenga conciencia moral, las decisiones humanas incorporadas en su diseño sí tienen consecuencias éticas.

Aquí surge la pregunta central: **¿una máquina puede considerarse un agente moral?**

### **¿Qué es un agente moral?**

Un agente moral es aquel que puede discernir entre el bien y el mal, tomar decisiones y asumir responsabilidad por sus acciones.

En el caso de la inteligencia artificial, puede hablarse de sistemas que actúan, calculan y producen resultados, pero eso no significa que comprendan lo que hacen. La IA no tiene conciencia, intención propia, libertad ni comprensión real del bien y del mal. No sabe qué está decidiendo ni puede asumir responsabilidad moral por sus acciones.

Por eso, la postura mayoritaria es que la IA **no es un agente moral pleno**. Puede operar en espacios donde hay implicaciones morales, pero la responsabilidad sigue recayendo en las personas y en las instituciones que diseñan, entrenan, implementan y supervisan estos sistemas.

### **Tres niveles para pensar la ética de la IA**

El caso permite distinguir tres planos de reflexión:

- La ética de las personas cuando diseñan, utilizan o interactúan con sistemas de inteligencia artificial.
- La ética incorporada en las máquinas a través de principios, valores y criterios definidos por seres humanos.
- La posibilidad —todavía abierta y discutida— de que existan agentes morales autónomos capaces de razonar éticamente, comprender sus decisiones y asumir responsabilidad.

Esta distinción ayuda a evitar confusiones. Aunque hablemos de “máquinas que deciden”, en realidad seguimos hablando de decisiones humanas trasladadas a sistemas técnicos.

### **Innovación con responsabilidad**

Los vehículos autónomos pueden abrir posibilidades importantes para la movilidad y la seguridad vial. Pero también plantean desafíos éticos que no pueden resolverse únicamente desde la eficiencia técnica.

El reto no consiste solo en preguntarnos qué puede hacer la tecnología, sino **qué valores incorporamos en su diseño, quién responde por sus consecuencias y cómo deliberamos colectivamente sobre decisiones que afectan a la vida de las personas.**

En definitiva, el caso de los vehículos autónomos nos recuerda que la inteligencia artificial puede calcular, optimizar y ejecutar, pero la responsabilidad ética sigue siendo humana.

### Lectura recomendada

- Extracto-resumen del capítulo 6, “¿Qué ética para las máquinas inteligentes?”, en *¿Ética o ideología de la inteligencia artificial?*, de **Adela Cortina**. Paidós, 2024

**Fátima Álvarez.** Dirección experta WorkétlCA

Con el apoyo de la Red de Empresas Máshumano

