



# Robotiuris 2017:

Ética y responsabilidad en el diseño de robots e inteligencia artificial

Resumen de la sesión del 16 de noviembre de 2017

**Ponentes:** **Javier Torre de Silva**, Socio del Departamento de Derecho Público y Sectores Regulados de CMS-Albiñana y Suárez de Lezo. Letrado del Consejo de Estado (en excedencia). Premio Nacional de Licenciatura en Filosofía, **Amparo Grau**, Profesora Titular de Universidad, acreditada como Catedrática. Departamento de Derecho Financiero y Tributario. Facultad de Derecho. Universidad Complutense de Madrid y **Eduard Fosch-Villaronga**, Investigador Posdoctoral en la Queen Mary University of London, Miembro del Grupo de Estandarización de la SPARC y de la Foundation for Responsible Robotics.

Patrocinado por:



Fundación para la investigación sobre el Derecho y la Empresa

[www.fidefundacion.es](http://www.fidefundacion.es)

Javier Torre de Silva, Amparo Grau y Eduard Fosch compartieron con los asistentes a esta sesión **la perspectiva ética en el diseño y responsabilidad de los robots**.

**Javier Torre de Silva** comenzó señalando que la ética en los robots es una necesidad y no una opción. En la medida en que delegamos la toma de decisiones en sistemas de inteligencia artificial, las **exigencias éticas que aplicamos a nuestras propias decisiones deben extenderse a las que tomen dichos sistemas**. Así pues, la ética que debemos inculcar a los sistemas de Inteligencia Artificial (IA) tiene que estar en línea con nuestros propios sistemas de valores. El problema que subyace con respecto a la ética y la IA es ¿cómo se enseña esto a una máquina? La ética no se basa sólo en un conjunto de normas morales y su aplicación consiguiente. Cualquier aproximación que se limite a una mera aplicación de normas generales resultará insuficiente.

Existe una Carta sobre robótica elaborada con la asistencia de la Unidad de Prospectiva Científica (STOA) de la *DG European Parliament Research Service* en el seno del Consejo de Europa en la que se insiste en la necesidad de que las máquinas gocen de parámetros éticos, pero estas reglas, aunque inspiradoras, no solucionan el problema.

Es difícil encajar o traducir a código fuente por ejemplo la ética de Aristóteles o Kant, que no se reducen a elencos de normas. Además, la ética aplicada a un ordenador no es la misma que la ética de un humano. No se persigue el libre desarrollo de la personalidad de la máquina, sino el del humano que interactúa con ella. Lo verdaderamente fundamental es que la máquina permita al hombre ser ético no que la máquina sea ética en sí misma.

La única forma en la que un sistema de IA puede interiorizar valoraciones éticas es mediante una **combinación de métodos deductivos** (basados en reglas) e **inductivos** (basados en la experiencia y en el falseamiento de errores), de forma similar a como aprenden los niños.

Probablemente, si logramos inculcar nuestros valores a los sistemas de IA, nos sorprenderán los resultados, pues las decisiones de tales sistemas serán probablemente mucho más éticas que las que tomamos nosotros en la vida real.

**Amparo Grau** comenzó su ponencia poniendo de relieve el hecho de que algunos de los problemas que plantea la robótica no son nuevos, pero el de las **relaciones entre los robots y las personas** sí.

Es innegable que la ciencia y la tecnología avanzan porque las personas quieren que se sigan desarrollando. Es necesaria una robótica socialmente responsable que contribuya al logro de los Objetivos de Desarrollo Sostenible, fijados en la Agenda 2030 de las Naciones Unidas. En el panorama actual en el que nos encontramos, de plena incertidumbre con respecto a muchos aspectos de la IA, necesitaremos que sean los legisladores y los poderes públicos quienes establezcan sistemas de gobernanza que sean inclusivos y justos. Para Amparo Grau con respecto a los robots

cabe preguntarse si los sistemas de IA tienen que ser éticos o vale sólo con que lo parezcan, así como si los robots son considerados un fin o un medio.

El rasgo esencial en la conducta moral de los hombres es querer libremente lo que se hace, pero una máquina no se puede guiar por estos parámetros puesto que lleva a cabo la tarea que le ha sido encomendada por un programador a través de unos códigos de programación. Hoy en día ya existen robots que están planteando problemas y dilemas de carácter ético, tales como los *sex robots*, los robots de asistencia o los drones militares. Es importante fijar un marco seguro para la IA, pero no hay acuerdo todavía sobre cómo regularlo y hasta dónde, dejando un margen para la responsabilidad social empresarial.

A juicio de **Eduard Fosch-Villaronga**, interactuamos con los robots de formas muy distintas, entre ellas, física y cognitivamente; y así es como debería ser reflejado en los instrumentos jurídicos que gobiernan a los robots. No reconocer la **interconectividad** entre el diseño de los robots, la regulación y los humanos que interactuarán con ellas, conllevará a la insuficiencia reguladora. Esto va a ganar importancia sobre todo con el aumento de la robótica de servicio, la cual está dirigida a la población general.

Eduard Fosch-Villaronga señala que, cuando se habla de “ética y responsabilidad en el diseño de los robots”, se hace referencia al comportamiento de las máquinas, de sus diseñadores y a la interacción que las mismas van a tener con los usuarios. A ello, la Unión Europea aboga por la **creación de un código de conducta para los ingenieros**, y el establecimiento de políticas sobre la utilización de los robots desde el punto de vista de los usuarios, y así poder ayudar en la transición de una robótica centrada en lo industrial a otra de servicio.

Fosch piensa que a los principios planteados por las instituciones europeas les falta una base técnica, pues conceptos como “reversibilidad” van a carecer de sentido en robots con cierta apariencia física (*embodiment*). Por ejemplo, si una persona lleva un exoesqueleto atado al cuerpo y cae, dicha acción difícilmente va a poder ser revertida. Eduard argumenta que si las futuras políticas europeas que gobiernen sistemas ciberfísicos no tienen en cuenta la naturaleza dual de los mismos (cíber y física) y cómo afecta una parte a la otra, corren el riesgo de ser completamente inaplicables.

*#Robotiuris17Fide, enero 2018.*

*Crónicas redactadas por la abogada Marta Pérez Cañón*