

## **YO, ROBOT: DE LA BIOLOGÍA A LA SINGULARIDAD. ¿NUEVAS PREGUNTAS PARA LA FILOSOFÍA DEL DERECHO?\***

José MARTÍNEZ DE PISÓN

CATEDRÁTICO DE FILOSOFÍA DEL DERECHO  
UNIVERSIDAD DE LA RIOJA

**SUMARIO:** I. Introducción. II. La ciencia ficción como ciencia prospectiva. III. De la biología a la singularidad. IV. Nuevas realidades, nuevas cuestiones.

**RESUMEN:** La evolución científica y el progreso de la sociedad han ido siempre acompañadas de un replanteamiento de las categorías jurídicas y políticas sobre las que se fundamenta. En la actualidad, se está anticipando una nueva revolución tecnológica, impulsada por los avances en la genética, la robótica y la nanotecnología. De modo invariable, esta nueva transformación de la humanidad está implicando una nueva revisión de dichas categorías. La ciencia ficción, en el cine –con películas como Yo, robot–, en la literatura, en las series de TV, ya están exponiendo un nuevo relato del futuro y, por tanto, un nuevo significado de esos conceptos. ¿Podrá la reflexión iusfilosófica adaptarse a estos nuevos tiempos y anticipar e, incluso, orientar y encauzar el debate del siglo XXI?

**PALABRAS CLAVE:** Ciencia Ficción, Yo, Robot, singularidad, transhumanismo, Filosofía del Derecho, roboética

**ABSTRACT:** Societal progress and scientific developments have always gone hand-in-hand with a rethinking of the underlying legal and political categories. Today, a new technological revolution is on the horizon, driven by advances in genetics, robotics and nanotechnology. Inevitably, this new transformation of humanity involves a new review of these categories. Science fiction, in cinema –with films such as I, Robot– in literature and in TV series, is already offering a new story of the future and, therefore, offering new meaning to these concepts. Can legal-philosophical reflection adapt to these new times and anticipate and perhaps even guide and channel the debate in the 21st century?

**KEYWORDS:** Science fiction, I, Robot, Singularity, Transhumanism, Philosophy of Law, Roboethics

---

\* Este texto es una versión elaborada de la charla que impartí en la Universidad Pompeu Fabra el 27 de febrero de 2017 con motivo del homenaje al prof. Ernesto Garzón Valdés por su noventa aniversario. Mi más sincero agradecimiento al prof. Jorge Malem, organizador de esta interesante jornada en la que pudimos estar con el maestro y con numerosos amigos.

## I. Introducción

*Yo, robot* (2004), película de ciencia ficción dirigida por ALEX PROYAS, toma como punto de partida el libro del mismo nombre de ISAAC ASIMOV. En realidad, es una versión bastante libre de este delicioso texto que, para su época (1950), planteaba la realidad de la vida humana en una sociedad distópica: un «mundo feliz» en el que los robots son compañeros y están al servicio de los humanos. Un mundo en el que se produce también una simbiosis entre la máquina y la biología produciendo así un nuevo modelo de humano, los *cyborg*. Y un mundo en el que la feliz cooperación entre ciencia y naturaleza se ve alterada por el desarrollo superior de la inteligencia artificial –personificado en la película por el macro ordenador VIKI– que acaba por promover una revolución de los robots con el objetivo de salvar la humanidad. La sociedad perfecta y feliz deviene, por mor de la propia evolución de la robótica, en una tiranía sobre los humanos, materializando así muchos de nuestros miedos en torno al progreso.

ASIMOV recopiló en su libro *Yo, robot* (1950), una serie de cuentos que tenían por objeto la relación entre humanos y robots, siempre complicada, bajo la atenta mirada de la robopsicóloga SUSAN CALVIN. En cada uno de ellos, el escritor plantea tramas diferentes a partir de la modificación de la conducta de los robots producida por alguna alteración de su cerebro positrónico. Todos constituyen un ameno relato de un futuro apenas previsible en la década de los años 40 y 50 del siglo pasado, pero que, sin embargo, puede estar mucho más cerca en este siglo XXI.

El libro de ASIMOV tiene una reconocida fama, sobre todo, porque en él se formula por primera vez las *tres leyes de la robótica*. Una especie de código moral del robot que debe regular su conducta, el ejercicio de sus superiores capacidades y que, en definitiva, es el reflejo del temor de los humanos ante la inteligencia artificial. Estas leyes son: 1.- Un robot no debe dañar a un ser humano o, por su inacción, dejar que un ser humano sufra daño; 2.- Un robot debe obedecer las órdenes que le son dadas por un ser humano, excepto cuando estas órdenes se oponen a la primera Ley; y 3.- Un robot debe proteger su propia existencia, hasta donde esta protección no entre en conflicto con la primera o segunda Leyes»<sup>1</sup>. Aunque fue el escritor y científico quien formuló estas leyes siempre reconoció que el autor de las mismas fue J. CAMPBELL, en 1942.

ASIMOV escribió numerosos relatos futuristas no sólo de robots o de sociedades distópicas, sino también de una humanidad desarrollada que ha iniciado los viajes interestelares y ha colonizado todo el universo. Incluso, en los mismos ha entremezclado sus historias de robots con sus personajes más conocidos dando lugar a cuentos y tramas de lo más inverosímil. Su fecunda imaginación ha inspirado numerosas películas y series televisivas de ciencia ficción<sup>2</sup>.

---

<sup>1</sup> ASIMOV, I., *Yo, robot*, trad. de M. Bosch, 9ª reimp., Barcelona, Edhasa, 2017, pág. 8.

<sup>2</sup> ISAAC ASIMOV fue un prolífico escritor e investigador científico. Publicó numerosos libros sobre ciencia e historia, así como novelas y recopilaciones de relatos de ciencia ficción ganando por primera vez los premios Hugo por esta última especialidad. En su producción literaria puede distinguirse dos etapas. En los años 50 del siglo XX, las primeras novelas de ciencia ficción entre las que destaca la *Trilogía o Ciclo de Trantor* en los que describe un universo colonizado por seres humanos que está llegando a su final, motivo por el cual los protagonistas inician un ignoto viaje a la búsqueda del origen, una Tierra inhabitable y apocalíptica, que conduzca a un universo nuevo en paz y armonía bajo la estela de las leyes de la psichistoria descubierta por

En sus cuentos, historias y relatos jugó también con la interpretación y aplicación de las tres leyes poniendo, incluso, en cuestión su virtualidad y su eficacia protectora de la humanidad. Por eso, al final de su vida, introdujo una cuarta ley: la Ley Zeroth. Esta ley afirma: «Un robot no hará daño a la Humanidad o, por inacción, permitir que la Humanidad sufra daño». Cerró así el círculo del código moral de conducta de los robots, formulando los primeros principios de una *roboética*, tan actual hoy en día.

Los relatos de ASIMOV son premonitorios de un futuro que es ya presente para nosotros o lo será en las próximas décadas. Viajes interestelares, nuevos mundos galácticos, pero también robots humanoides, una inteligencia artificial que supera al cerebro humano, y la incorporación al ser biológico de novedades técnicas que aumentan sus capacidades. Ya lo notamos de una forma imperceptible a través de los artefactos que vamos incorporando a nuestros cuerpos, ya sea en la forma de *weareables* o de *insideables*. Es posible que, en un futuro cercano, convivamos en la misma sociedad la vieja especie junto con humanos renovados artificialmente o cyborgs, con robots humanoides, con una inteligencia artificial muy superior. Es posible que, como en la serie de ASIMOV, pueda alargarse la vida humana durante dos, tres o cuatro siglos. Para algunos, la inmortalidad es posible. Las condiciones de la sociedad, del planeta, del universo, si todo esto se produce, habrán mutado drásticamente. Tal revolución bien merece una reflexión pausada sobre estos fenómenos y sus consecuencias.

## II. La ciencia ficción como ciencia prospectiva

La literatura y el cine, a los que hay que añadir en los últimos años las series de televisión<sup>3</sup>, hace tiempo que lanzan señales sobre el futuro. Un futuro imaginado y muy variado que va desde la realidad de un mundo apocalíptico a la realización de un mundo feliz, ordenado, en paz y pleno de bienestar para todos. En estas realidades imaginadas, plasmamos los seres humanos nuestros sueños y nuestras utopías, nuestro anhelo fáustico de superación, pero también nuestras pesadillas y nuestros miedos. Unos mundos en los que el buen y el mal uso de la tecnología devienen, precisamente, en la realización de esos sueños o en la aparición de esas pesadillas.

---

Hari SELDON. En esos mismos años, ASIMOV inicia también una serie de novelas y de relatos sobre robots donde aparecen también sus personajes más conocidos: el detective Elijah BAALEY y el robot R. Daneel OLIVAW. Casi treinta años después, continuó estas novelas de ciencia ficción de la serie Fundación. En ellas, mezcló su visión futurista del planeta Tierra, de la galaxia y del universo con sus relatos sobre los robots transmitiéndonos una visión muy actual de la sociedad y de la humanidad. No es de extrañar que, con su fecunda imaginación, ASIMOV inspirara a directores de cine como S. KUBRICK, G. LUCAS, S. SPIELBERG o R. SCOTT.

<sup>3</sup> Es demasiado ingenuo menospreciar a las series como vehículo de expresión de las opiniones e ideas políticas de nuestro tiempo y de nuestras proyecciones de futuro. Como afirma D. MOÍSI, «las series no sólo se han convertido en el equivalente de lo que fueron los folletines en el siglo XIX: una dimensión esencial de la cultura. Sus guionistas son, al menos en lo que respecta a los mejores, comparables a quienes fueron los grandes novelistas del pasado, desde BALZAC hasta FLAUBERT, pasando por DICKENS. No se conforman con analizar fríamente la realidad: la perciben y la adivinan, gracias al poder de su intuición y al coraje y lucidez de su imaginación. De hecho, estos guionistas se han convertido en los mejores analistas de las sociedades y del mundo actual, por no decir en futurólogos más fiables». Vid. MOÍSI, D., *Geopolítica de las series o el triunfo global del miedo*, trad. de S. Moreno, Madrid, Errata naturae, 2017, pág. 11.

Estas representaciones del futuro las dejamos en manos de la fantasía o de la imaginación y las encuadramos en la «ciencia ficción», como si fuese un apartado menor de la ciencia y de las formas del conocimiento en comparación con las ciencias experimentales, con el uso del método científico inaugurado por I. NEWTON o con el avance de la química, de la física, de la astronomía, de la medicina, etc., en el siglo XX. Con la comprobación de las teorías de M. PLANCK y de A. EINSTEIN.

Por eso, en la medida en que este tipo de literatura refleja alguno de los mundos futuros posibles no son sólo objeto de la fantasía o de la imaginación de sus autores, más o menos acertadas, sino que pueden ser considerados como algo más, como una anticipación más real, como un acto de «prospectiva», de análisis de desarrollos de realidades presentes, de una anticipación y proyección que entra dentro de la capacidad de planificación y de organización del ser humano. De ahí que sea más oportuno hablar de «ciencia prospectiva». Pues, la mente del autor de «ficción», liberada de las ataduras de la realidad, realiza una auténtica proyección de mundos imaginables a la luz de las revoluciones tecnológicas, de las investigaciones científicas y de los avances en sus diferentes ramas, de las experiencias humanas e históricas, de sus inquietudes y de sus obsesiones, etc. Todo ello da lugar a un variopinto abanico de posibilidades y de proyectos de futuro.

Considerada la ciencia ficción como ciencia prospectiva, presenta un futuro variado en el que se mezclan mundos apocalípticos de violencia junto con paraísos que quieren reflejar una deseada edad dorada, plena de bienestar, paz y supuesta felicidad. En todo caso, los humanos ya no son humanos al estilo de siempre: o bien, han desarrollado alguna habilidad o poder especial (*X-Men*), o bien conviven con robots y con seres medio humanos, medio máquinas: los *cyborgs*. Por supuesto, la humanidad no desaparece, salvo que sea por mor de algún desastre natural o provocado por la voluntad del hombre mismo.

Pero, tomemos algunos ejemplos de la ciencia prospectiva. *Metrópolis* (1927), dirigida por FRITZ LANG, es una película del cine mudo. Es una obra de arte que ha inspirado a otros directores y es una reflexión sobre la naturaleza humana, sobre la relación entre lo natural y lo artificial, y sobre la explotación del hombre en la sociedad industrial. Todo un aviso sobre las consecuencias del fordismo y de la alianza entre el capitalismo salvaje y la revolución tecnológica. No sólo por la imagen a un tiempo cruel y conmovedora de las masas alineadas, yendo al trabajo deshumanizado, sino también por el recurso del poder capitalista al omnipotente robot para volver a amansar a la humanidad explotada. Sólo el amor asegura la reversión a una sociedad más humana. ¿Es la anticipación de FRITZ LANG una premonición del futuro de la humanidad y del uso de las transformaciones tecnológicas?

Se dice que *Metrópolis* inspiró a RIDLEY SCOTT en su *Blade Runner* (1982)<sup>4</sup>. El contexto ya no es una Nueva York deshumanizada por la organización fordista del trabajo y la explotación humana, sino la ciudad de Los Ángeles de 2019, caótica y superpoblada, en un mundo apocalíptico, a la que acuden los nuevos robots (los *replicantes*) a pedir cuentas a su creador por darles una vida finita y limitada. ¡Cuando son los replicantes los que sostienen la sociedad putrefacta de los seres humanos!

---

<sup>4</sup> Mientras escribo estas páginas, se presenta en España la secuela *Blade Runner 2049*.

En efecto, *Blade Runner*, basada en cuento de PH. K. DICK, ¿*Sueñan los androides con ovejas eléctricas?*?, describe un futuro distópico en el que robots humanoides (los replicantes de la serie Nexus-6), con superiores capacidades que los seres humanos, se rebelan al saber que están condenados a morir. Son perseguidos por los *blade runner*, especie de policías de robots, de los que Deckard (HARRISON FORD), a pesar de estar apartado del cuerpo, es llamado de nuevo por ser el más implacable.

Pero, lo interesante es que, a lo largo de la película, son varias las escenas que plantean la condición misma de lo humano. ¿Un Deckard, que se muestra implacable en la persecución de los replicantes y en su «retiro», esto es, en su liquidación, o estos robots, que se preguntan por la razón de su existencia y que luchan por sobrevivir a su terrible destino? Son diferentes los momentos en los que se cuestiona la división hombre-replicante (natural-artificial). En ellos, queda en evidencia la falta de sentimientos y de escrúpulos del *blade runner* Deckard (¿no será un replicante más?), así como el reconocimiento de la presencia de la inteligencia emocional (o sea, lo humano mismo) en los robots humanoides, en particular, en Roy, el líder de estos revoltosos. Entre medio, la frustración de Rachel, la última generación de Nexus a la que se ha incorporado recuerdos y emociones, pero que, enamorada de Deckard, descubre que es otra replicante.

Sin duda, el interés de la película se eleva en la escena antesala del final, lucha entre lo humano y lo artificial, el *blade runner* y el replicante, Deckard y Roy. Parece que el fiel de la balanza se inclina a favor de los robots, más inteligentes y poderosos. En la escena, Roy-replicante, sabiendo que llega su final, en lugar de dejar que Deckard-humano caiga desde un edificio y se estrelle en el suelo, lo agarra y le salva la vida, mientras se le acaba la vida programada por su creador. Recordando sus accidentadas experiencias, dice la famosa frase: «Todos esos momentos se perderán en el tiempo como lágrimas en la lluvia. Es hora de morir». Empieza a llover, mientras, según algunos, él también llora demostrando lo mucho que tiene de humano. Al final, ¿en qué reside la cualidad de lo humano? ¿Quién es, realmente, el ser humano? ¿Deckard o Roy?»<sup>5</sup>

Al comentar esta escena final, J. DE LUCAS afirma: «Precisamente lo que sucede en la película es que esa diferencia ya no existe entre lo humano y su *réplica* y esa es la propuesta de Roy: hacerse humano es una tarea, no una herencia, y la piedad y el amor se pueden adquirir en ese proceso de aprendizaje, mediante el tiempo vivido que es la humanización. En esa tarea, Roy es, posiblemente, *demasiado humano*»<sup>6</sup>.

Son numerosos los ejemplos de películas o series de televisión que en los últimos años tratan la *humanización* de los robots. En este sentido, me parece oportuno referirme a una de las películas de los últimos años que constituye un paradigma de lo que estoy afirmando: *Ex Machina* (2015). Y entre las series *Westworld* (2016).

*Ex Machina*, escrita y dirigida por ALEX GARLAND, cuenta la historia de un programador que es elegido por el presidente de la empresa más importante en inteligencia

---

<sup>5</sup> *Blade Runner* ha dado lugar a numerosas interpretaciones, tanto en relación a cuestiones generales o técnicas, como filosóficas e, incluso, jurídicas. En este sentido, puede consultarse NAVAJAS, S., *EL hombre tecnológico y el síndrome Blade Runner*, Córdoba, Berenice, 2016. Desde la perspectiva jurídica, LUCAS, J. DE, *Blade Runner. El Derecho, guardián de la diferencia*, Valencia, Tirant lo Blanch, 2003.

<sup>6</sup> LUCAS, J. DE, *Blade Runner. El Derecho, guardián de la diferencia*, cit., pág. 35.

artificial para que aplique el *test de Turing* en una robot humanoide. La trama de la película muestra el éxito total en el desarrollo de la inteligencia artificial y de robots humanoides. AVA, la ginoide inteligente, es capaz de cautivar mental y físicamente al programador, de urdir una estratagema para salir de su prisión y, a la postre, de escaparse aplicando realmente una estratagema finalista y utilitarista. *Ex Machina* muestra las enormes posibilidades de que, en un tiempo bastante corto, los robots humanoides superen en inteligencia a los humanos.

ALAN TURING es conocido en el mundo científico por importantes aportaciones. Una de ellas es el descubrimiento de las claves para descifrar Enigma, esto es, los códigos secretos de los alemanes en la II Guerra Mundial con los que enviaban instrucciones a los submarinos que atacaban a los convoyes de Estados Unidos. Se calcula que su contribución ayudó a adelantar la conclusión de la guerra entre dos y cuatro años<sup>7</sup>. El *test de Turing* es una prueba de habilidad que se hace a una máquina con inteligencia artificial para intentar demostrar si ha alcanzado el nivel de desarrollo de la inteligencia humana<sup>8</sup>. El test, tal y como es explicado por el mismo autor, es un test sumamente controvertido. Para muchos, no garantiza el éxito de su resultado. No obstante, es considerado, hoy por hoy, una de las pruebas más relevantes para conocer el nivel de desarrollo de la inteligencia artificial.

*Westworld*, la serie de televisión protagonizada por ANTHONY HOPKINS, va en esta misma línea de demostrar los avances y los desarrollos en inteligencia artificial dejando incluso atrás al *test de Turing*. *Westworld* es un parque temático en el que los seres humanos se divierten, violentan, liberan sus pasiones a costa de los anfitriones, robots que son reutilizados y reparados después de cada historia, al tiempo que, en alguno de ellos, van surgiendo recuerdos, así como una especie de conciencia que hace que acaben «rebelándose». En todo caso, es un ejemplo de cómo la ciencia prospectiva, en este caso, de la mano de una serie de televisión, muestra las similitudes entre lo natural y lo artificial y las posibilidades de un futuro en el que ambos, humanos y robots, acabemos coexistiendo.

Podrían ponerse más ejemplos del reflejo en películas y en serie de televisión de la capacidad prospectiva de guionistas y directores, pero no haría más que redundar en el sentido de lo que quiere exponerse. Desde el deseo de una sociedad como una edad dorada, de orden y bienestar general (*Minority Report*), a la superación de las capacidades humanas (*X-Men*, *Terminator*), la confusión entre realidad y mundo virtual (*Matrix*, *Her*) o la búsqueda del origen y de revuelta de la inteligencia artificial (*2001: una odisea en el espacio*), etc. Y de series prospectivas (*Black Mirror*, en la que el capítulo *San Junípero*, de la tercera temporada, el espíritu, la vivencia, la conciencia, la experiencia de las protagonistas acaban en una cápsula y en un gran ordenador gestionado por brazos robóticos con el premio de poder continuar su amistad). En todo caso, estas obras expresan los sueños, los deseos, los

---

<sup>7</sup> A pesar de ello, ALAN TURING fue acusado de homosexual en 1952, fue condenado de acuerdo a la dura legislación inglesa y sometido a «curas» inhumanas hasta que se suicidó en 1954. La reina Isabel II rehabilitó su figura casi sesenta años más tarde, en 2013.

<sup>8</sup> El test consiste, sucintamente, en la evaluación que debe realizar un ser humano de sus conversaciones con una máquina. Si después de un tiempo no ha sido capaz de distinguir si habla con una máquina o con un ser humano, la inteligencia artificial ha superado la prueba. No se evalúa el conocimiento, sino la capacidad para generar respuestas lo más parecidas posibles a las de un ser humano.

miedos e inseguridades de los seres humanos. Ellas mismas se convierten en textos de reflexión filosófica sobre lo que somos y lo que podemos ser.

### III. De la biología a la singularidad

Pero estas reflexiones y estas visiones futuristas no están hechas en el aire. Hay un trabajo y un desarrollo del conocimiento y de la ciencia que avala estas construcciones de la imaginación prospectiva. ASIMOV es un escritor de cuentos y novelas futuristas, pero, además, es un científico con un amplio conocimiento, igualmente acreditado a su vez por sus numerosas publicaciones<sup>9</sup>. Sabía sobre lo que fantaseaba. Lo mismo se puede decir de TURING, no sólo por sus éxitos en los inicios de la inteligencia artificial, sino porque sus sugerencias han supuesto importantes aportaciones en el desarrollo de los primeros ordenadores.

Desde entonces, los avances en inteligencia artificial cabe calificarlos de asombrosos. Desde que TURING ideara su famoso test, por el que defendía la posibilidad de que las máquinas fuesen inteligentes y que, algún día, pudiesen superar al mismo ser humano, se han promovido proyectos que han desarrollado y superado sus ideas. En la actualidad, no es ningún absurdo la posibilidad de que, en un futuro cercano, se cumplan sus profecías. En efecto, los avances empezaron en la misma década de los años 50s del siglo XX. Aunque fueron lentos al principio, en los últimos años, los cambios están siendo espectaculares<sup>10</sup>.

Siguiendo las pautas de TURING se construyeron las primeras máquinas de inteligencia artificial. Poco a poco, se fueron superando las dificultades. IBM se propuso el desarrollo en investigación de estas máquinas hasta el diseño y fabricación de los ordenadores. Hoy en día, no podemos vivir sin su presencia. Nos acompañan y nos ayudan desde las tareas domésticas o laborales hasta la solución de los problemas más complejos.

El afán de superación de la mente humana está en la base de los progresos en inteligencia artificial. Algunos ejemplos de emulación son bien conocidos por el gran público. En 1997, *Deep Blue*, de IBM, lograba algo impensable: ganar al campeón del mundo de ajedrez GARRI KASPAROV. A pesar del éxito de la computadora, realmente, se trató del éxito en el almacenamiento de datos y jugadas de ajedrez sin que *Deep Blue* evidenciara otro tipo de avance en inteligencia. Fue un triunfo del *hardware* y de la capacidad de memoria. No obstante, fue un momento importante en el desarrollo de la robótica, pues se considera que el ajedrez es el juego que refleja mejor la capacidad racional del ser humano. Lo mismo sucedió con otro ordenador de IBM, *Watson*, que consiguió ganar un popular concurso estadounidense, *Jeopardy!*, en el que se pone a prueba los

---

<sup>9</sup> Tras una primera etapa de escritos de cuentos de ciencia ficción, ISAAC ASIMOV sacó a relucir sus conocimientos científicos y su formación en ingeniería. Estos son algunos ejemplos de su labor como conocedor y divulgador de los avances científicos: *Momentos estelares de la ciencia* (1959), *Breve historia de la química* (1965), *¿Hay alguien ahí?* (1967), *El universo* (1971); más adelante, siguió publicando: *El cometa Halley* (1985), *Nueva guía de la ciencia* (1986), *Alpha Centauri. La estrella más próxima* (1991). A toda esta producción científica, hay que añadir sus libros de historia de la humanidad.

<sup>10</sup> Sobre estos avances puede verse un breve resumen en ORTEGA, A., *La imparabla marcha de los robots*, Madrid, Alianza Editorial, 2016, págs. 77 y ss. y 85 y ss.

conocimientos de los concursantes<sup>11</sup>. Como afirma ORTEGA, *Watson*, a pesar de su éxito, igual que *Deep Blue*, «no deja de ser una máquina de Turing sin lo que llamaríamos vida mental»<sup>12</sup>. El éxito consistió en construir una máquina capaz de entender el lenguaje ordinario y de responder a las cuestiones planteadas. A pesar de todo, fue, sin duda, un importante avance.

Los siguientes pasos en el desarrollo de la inteligencia artificial parecen materializar la prospectiva de la ciencia ficción. Los casos anteriormente citados son ejemplo de la capacidad de almacenamiento de conocimientos e, incluso, de respuesta. El siguiente paso ha sido el de que, además, la IA desarrolle sus propios procedimientos de aprendizaje. Esto es lo que parece que ha hecho el programa *AlphaGo* –sobre todo, en su versión *AlphaGo Zero*– diseñado por investigadores chinos en colaboración con Google DeepMind, en el que se mezclan técnicas clásicas de almacenamiento de datos con modelos de autoaprendizaje. El programa *AlphaGo* superó a maestros del juego Go –juego milenario chino que se considera el antecedente del ajedrez– e, incluso, se le concedió un nivel alto en los rankings mundiales de dicho juego. De nuevo, A. ORTEGA lo explica magníficamente: «La novedad de AlphaGo estriba esencialmente en el método, que ha revolucionado el mundo de la IA, pues puede cambiar muchas cosas en la vida real. Sus algoritmos contienen tres capacidades inspiradas en el ser humano: el aprendizaje por refuerzo; las redes neuronales o redes de profunda convolución, como el cerebro; y la memoria selectiva, como hace el hipocampo. AlphaGo se nutrió al principio de centenares de miles de partidas registradas. Y luego empezó a mejorar jugando sin cesar contra sí misma»<sup>13</sup>. En suma, esto es la inteligencia artificial: máquinas que evolucionan solas y que aprenden de sí mismas (*machine learning*).

Dicho con las palabras de J. KAPLAN, «los últimos avances en robótica, percepción y aprendizaje automático, impulsados por la aceleración de las mejoras en tecnología informática, están posibilitando una nueva generación de sistemas que compiten con las capacidades humanas o las sobrepasan»<sup>14</sup>. Por tanto, el futuro previsto en la literatura, en las películas y en las series de televisión, calificadas como de ciencia ficción, son ya cosas del presente.

En este contexto resultan de sumo interés las tesis transhumanistas y la posición de RAY KURZWEIL sobre el advenimiento de una nueva *singularidad* en el desarrollo de la especie humana. La teoría de KURZWEIL de que la «singularidad está cerca» es una tesis sumamente controvertida, pero, no por ello, deja de tener interés en el desarrollo de la

---

<sup>11</sup> El primer concurso tuvo lugar el 30 de marzo de 1964 sobreviviendo, en diferentes épocas, hasta hace pocas fechas. En febrero de 2011 tuvo lugar «The IBM Challenge» en el que compitieron dos de los concursantes con más éxito y el ordenador *Watson*, desarrollado por IBM. Una competición hombre-máquina. *Watson* ganó en tres programas.

<sup>12</sup> ORTEGA, A., *La imparable marcha de los robots*, cit., pág. 86.

<sup>13</sup> ORTEGA, A., *La imparable marcha de los robots*, cit., págs. 87-88.

<sup>14</sup> KAPLAN, J., *Abstenerse humanos. Guía para la riqueza y el trabajo en la era de la inteligencia artificial*, trad. de S. Arilla, Zaragoza, Teell Editorial, 2016, p. 1.



historia que estamos analizando<sup>15</sup>. Es más, es una pieza clave para entender el terreno en el que nos movemos.

La noción de «singularidad» procede del mundo de las matemáticas. Así como durante décadas el concepto de «paradigma» desarrollado por TH. KUHN<sup>16</sup> ha servido para explicar la evolución de la ciencia y de sus profundas transformaciones, en la actualidad, el término «singularidad» constituye un potente instrumento teórico cuyo objetivo es posibilitar la predicción del futuro –incluso, de escenarios relativamente cercanos–, de la humanidad. Como se ha afirmado, el origen del significado «singularidad» procede de las matemáticas y, en este campo, quiere hacerse referencia a comportamientos extraños o inesperados producidos por determinadas funciones. Con el tiempo, se ha aplicado a la innovación tecnológica para explicar su creciente desarrollo y su influjo en el mundo.

RAY KURZWEIL señala que fue el matemático y físico JOHN VON NEUMANN quien, en los años cincuenta del siglo XX, utilizó el término para referirse a la evolución tecnológica y a su repercusión en el modo de vida de las sociedades humanas y que fue, treinta años más tarde, VERNOR VINCE, matemático también y científico de la computación, quien confirió el significado prevalente finalmente. Realmente, ha sido el propio KURZWEIL quien ha popularizado el uso de este concepto, que no deja de ser una categoría carente de ambigüedades.

En el momento presente, la idea de la singularidad, según KURZWEIL, hace referencia «a un tiempo venidero en el que el ritmo del cambio tecnológico será tan rápido y su repercusión tan profunda que la vida humana se verá transformada de forma irreversible»<sup>17</sup>. Aclara que este tiempo y su repercusión en la vida humana no es ni utópico ni distópico. La cuestión central es que la evolución de la tecnología creada por el hombre se está acelerando a tal ritmo que está superando los límites de lo humano. La conclusión de este cambio exponencial será, como dice el subtítulo de su libro, la superación de lo biológico: «La Singularidad constituirá la culminación de la fusión entre nuestra existencia y pensamiento biológico con nuestra tecnología, dando lugar a un mundo que seguirá siendo humano pero que trascenderá nuestras raíces biológicas. En la post-Singularidad, no habrá distinción entre humano y máquina o entre realidad física y virtual. Si se pregunta sobre lo que seguirá siendo inequívocamente humano en un mundo así, la respuesta es simplemente esta cualidad: la nuestra es la especie que inherentemente busca expandir su alcance físico y mental más allá de sus limitaciones actuales»<sup>18</sup>. De hecho, la singularidad no es otra cosa que el siguiente paso en el proceso evolutivo del planeta y de la humanidad: desde que surgieron los primeros átomos y el mundo se desarrolló gracias a las leyes de la física y de la química hasta el surgimiento de la biología, la neurociencia y la tecnología (las seis eras de la evolución, según este autor).

Por encima de todo, la tesis de la singularidad de KURZWEIL tiene un especial atractivo pues parece realizar los anhelos fáusticos del ser humano: la superación de las

---

<sup>15</sup> Vid. KURZWEIL, R., *La singularidad está cerca. Cuando los humanos trascendamos la biología*, trad. de C. García, Berlín, Lola Books, 2012.

<sup>16</sup> KUHN, TH., *La estructura de las revoluciones científicas*, trad. de A. Contín, Madrid, FCE, 1982.

<sup>17</sup> KURZWEIL, R., *La singularidad está cerca. Cuando los humanos trascendamos la biología*, cit., pág. 7.

<sup>18</sup> KURZWEIL, R., *La singularidad está cerca. Cuando los humanos trascendamos la biología*, cit., pág. 10.

limitaciones biológicas, la fusión de lo humano y lo artificial, la conversión en cyborg, incluso, la trascendencia del ser vivo, nuestra inmersión en lo universal, en fin, el logro de la inmortalidad, etc. Todo lo que aparecen en las obras de la ciencia ficción. En el cine, en la literatura y en las series de televisión.

Pues bien, KURZWEIL fundamenta su controvertida teoría en numerosas apreciaciones que no han pasado desapercibidas. Pero, al menos, hay dos ideas clave en su visión de la singularidad: en primer lugar, la *ley de rendimientos acelerados* o ley de Moore como explicación del crecimiento exponencial del conocimiento, y, en segundo lugar, su aplicación a tres ciencias, la genética, la nanotecnología y la robótica (GNR). Según esto, la evolución tecnológica se está acelerando a ritmo exponencial desde finales del siglo XX, lo que permite predecir que, para dentro de pocas décadas, la singularidad se habrá cumplido.

De hecho, el propio KURZWEIL establece una serie de predicciones. En una primera fase, que se materializaría a finales de la década de 2020, la ciencia habrá «completado el proceso de aplicación de la ingeniería inversa al cerebro», de manera que tendremos un mapa del mismo y permitirá crear sistemas no biológicos que cumplan con los patrones de inteligencia, incluida la emocional. La siguiente fase permitirá cargar los datos del cerebro biológico –conocimiento, conciencia, experiencia, etc.– en «un sustrato no biológico pensante adecuado para ello». El tercer momento, «el más emocionante de todos», supondrá el paso de lo biológico a lo no biológico. «Esto ya ha comenzado con la benigna introducción de dispositivos tales como implantes neuronales... Se harán progresos con la introducción de nanorobots en el torrente sanguíneo... Más tarde, nanorobots más sofisticados harán de interfaces con nuestras neuronas biológicas para mejorar nuestros sentidos, proporcionando con ello realidad virtual y aumentada... También ayudarán a nuestra memoria y realizarán otras tareas cognitivas rutinarias. Entonces seremos *cyborgs*...». Todo ello se materializará en la década de 2040, en la que la parte no biológica será más potente que la biológica<sup>19</sup>. En definitiva, la singularidad es el siguiente paso en el proceso evolutivo de la especie humana.

A la vista de estos y otros comentarios de *La singularidad está cerca. Cuando los humanos trascendamos la biología*, uno no sabe muy bien dónde está ciencia ficción o la ciencia prospectiva. Lo más interesante es que estas ideas han tenido una enorme repercusión y no sólo en el cine y en las series de fantasía. Por supuesto, la ha tenido en el mundo de la robótica, de la ciencia y de la filosofía.

Hoy, constituyen una de las piezas clave del pensamiento transhumanista<sup>20</sup>. Además, marcan la línea de acción de poderosos grupos de interés y de multitud de personas fieles a esta filosofía. Por no ir más lejos, importantes empresas de Silicon Valley (entre ellas, Google y la NASA) crearon, a primeros de la década de 2010, la Universidad de la Singularidad (*Singularity University*) de la que RAY KURZWEIL es su presidente, además de fundador<sup>21</sup>. Aparte de sus cursos y estudios sobre el futuro, redes y sistemas informáticos, biotecnología y bioinformática, nanotecnología, etc., esto es, las disciplinas básicas de la

<sup>19</sup> KURZWEIL, R., *La singularidad está cerca. Cuando los humanos trascendamos la biología*, cit., págs. 432 y ss.

<sup>20</sup> Véase la obra de NATASHA VITA-MORE, NICK BOSTROM y DAVID PEARCE, estos dos últimos autores de la Declaración Transhumanista.

<sup>21</sup> Web de la Universidad de la Singularidad: <https://su.org>

«singularidad», editan un *newsletter* a través del cual divulgan los avances en todas estas materias<sup>22</sup>.

Igualmente, estas ideas han tenido un reflejo importante en proyectos de investigación y en programas sobre el futuro. Es el caso de la iniciativa «2045» (AVATAR 2045), impulsada en 2011 y financiada por el millonario ruso DIMITRI ITSKOV, con el que pretende implementar el ideario de la singularidad. Concretando aún más el proyecto de KURZWEIL, AVATAR 2045 establece cuatro etapas<sup>23</sup>: 1.- AVATAR A (2015-2020), etapa en la que se logre desarrollar un robot con forma de cuerpo humano o avatar. 2.- AVATAR B (2020-2025), en la que se logra trasplantar el cerebro humano a un avatar. 3.- AVATAR C (2030-2035), en la que se transfiere la personalidad de un ser humano a un cerebro artificial y a un avatar. 4.- (2040-2045), en la que se logra convertir el avatar en un holograma. Se logra así la inmortalidad del ser humano<sup>24</sup>.

Pero, incluso, los tecnólogos más escépticos, como NICK BOSTROM, ven plausibles alguno de los desarrollos de estos proyectos. En efecto, BOSTROM, en su libro *Superinteligencia. Caminos, peligros, estrategias*, pretende meter algo de sordina a las tesis de KURZWEIL analizando las bases de su propuesta, en particular, el desarrollo de la *ley de rendimientos acelerados* o ley de Moore y los caminos para el avance de la inteligencia artificial o superinteligencia, que son fundamentales para el advenimiento de la singularidad. Pues bien, después de todo, reconoce que la inteligencia artificial llegará e, incluso, superará el nivel humano a lo largo del siglo XXI. Recoge una encuesta realizada a expertos en IA por NILS NILSSON en la que se les pregunta por la probabilidad de que esto suceda: el 10% cree que para 2030; el 50% para 2050 y el 90% para 2100<sup>25</sup>.

Por supuesto, no faltan quienes nos ponen sobre aviso acerca de las calamidades que sobrevendrán del triunfo de las tecnologías. El título del libro de EUGENY MOROZOV, uno de los popes del «éxito del silicio», es muy significativo: *La locura del solucionismo tecnológico*<sup>26</sup>. Una crítica acerada al recurso de la tecnología de la información, de la cibernética y la robótica, y de su presencia en nuestras vidas. El libro es todo un trabajo de predicción de las consecuencias de la implantación de estas tecnologías y del dominio de los algoritmos en el desarrollo de las sociedades: lo bueno y lo malo. La siguiente afirmación, recogida en la introducción, es todo un aviso a navegantes:

«De allí la premisa de este libro: el objetivo de Silicon Valley de meternos a todos en una camisa de fuera digital fomentando la eficacia, la transparencia, la certeza y la perfección –eliminando, por consiguiente, sus contrapartes negativas: la fricción, la opacidad, la ambigüedad y la imperfección– resultará demasiado caro en el largo plazo. Por distintas razones ideológicas que se explicarán más adelante, ese alto costo permanece

<sup>22</sup> <https://singularityhub.com>

<sup>23</sup> El proyecto tiene su sitio web: 2045.com

<sup>24</sup> Por supuesto, estas propuestas encajan con el proyecto de KURZWEIL tanto con su idea de la singularidad como con la empresa de escaneo del cerebro humano. Conviene recordar que uno de sus últimos libros se titula *Cómo crear una mente. El secreto del pensamiento humano*, trad. de C. García, Berlín, Lola Books, 2013.

<sup>25</sup> BOSTROM, N., *Superinteligencia. Caminos, peligros, estrategias*, trad. de M. Alonso, Zaragoza, TEEL, 2016, pág. 19.

<sup>26</sup> MOROZOV, E., *La locura del solucionismo tecnológico*, trad. de N. Viviana, Buenos Aires, Katz, 2015.

oculto a los ojos del público, y así seguirá mientras nosotros, en esta búsqueda irracional del Edén de silicio, no examinemos radicalmente nuestro embelesamiento ante un conjunto de tecnologías que suelen meterse en la misma bolsa bajo la engañosa etiqueta de 'internet'. Por lo tanto, en este libro intento incluir como variable de la ecuación los verdaderos costos de ese paraíso tan esperado, y explicar por qué ha sido muy difícil rendir cuentas sobre ellos.

La imperfección, la ambigüedad, la opacidad, el desorden y la oportunidad de errar, de pecar, de hacer lo incorrecto: todos son elementos constitutivos de la libertad humana; cualquier esfuerzo dirigido a erradicarlos también erradicaría la libertad»<sup>27</sup>.

Humano, demasiado humano<sup>28</sup>. Todo un ataque a lo que llama solucionismo, pero que está en la base del optimismo tecnológico de las tesis de KURZWEIL.

#### IV. Nuevas realidades, nuevas cuestiones

Mientras se resuelven las alternativas a las dudas planteadas por EUGENY MOROZOV y otros sobre el uso de la tecnología y su proyección en nuestras vidas, la realidad es que, de forma imperceptible, se van cumpliendo alguna de las predicciones del tipo de las de RAY KURZWEIL. El uso de smartphones y de tablets, la creciente oferta de software «útil» que desarrolla aplicaciones sobre mapas, salud, organización y gestión del trabajo, ocio y videojuegos, etc., se amplía constantemente. Los bots y los robots inundan nuestra vida y condicionan nuestras relaciones sociales. A su vez, la inteligencia artificial, en pequeñas dosis, nos permite mejorar nuestro cuerpo y nuestro físico, y superar nuestras limitaciones. Nos hacemos cyborg gracias al uso de gafas, audífonos, marcapasos, etc. *Weareables* e *insideables* aumentan nuestras capacidades físicas y psíquicas, estrechando los límites entre lo natural y lo artificial<sup>29</sup>.

Al mismo tiempo, la robotización avanza rápidamente en nuestras vidas, en el seno de la sociedad y en el empleo y en la economía. Encontramos inteligencia artificial o robots en todas aquellas facetas que pueden ser objeto de automatización. No sólo eso, sino que imperceptiblemente, cada vez más, nos apoyamos en las máquinas en nuestras vidas, en nuestros éxitos profesionales, en el ocio, etc. Hoy por hoy, la inteligencia artificial está muy presente en nuestras vidas y representa todavía, a pesar de que se alzan voces que previenen futuros riesgos, un enorme potencial de desarrollo y de bienestar para los seres humanos. Como afirma A. ORTEGA, «los robots, de momento, son únicamente una parte de este

---

<sup>27</sup> MOROZOV, E., *La locura del solucionismo tecnológico*, cit., pág. 16. El libro es también un ataque a internet y su preponderancia en nuestras vidas: «Esta propensión a creer que internet es una fuente de sabiduría y asesoramiento sobre políticas es lo que transforma un conjunto de cables y enrutadores de red poco interesantes en una ideología seductora y apasionante; quizá la superideología de nuestros días» (pág. 45). Con esta premisa, el libro es un ataque a los grandes centros de poder tecnológico: Facebook, Google, Youtube, Amazon, etc., e, incluso, Wikipedia. Nada es gratis, ni libera, pero siempre hay gente que se enriquece con estos desarrollos.

<sup>28</sup> MOROZOV no es el único gurú tecnológico que acaba por revolverse contra el dominio de lo que llama «solucionismo» que no es otra cosa que esa utopía cibernética que aparece en las obras de ciencia ficción. Un caso similar es el de JARON LANIER. De sus publicaciones, merece la pena destacar *¿Quién controla el futuro?*», trad. de M. Pérez, Barcelona, Debate, 2014.

<sup>29</sup> Sobre esto, puede verse BRONCANO, F., *La melancolía del ciborg*, Barcelona, Herder, 2009, en particular, el cap. I.

inmenso potencial tecnológico, que produce novedades casi a diario, y en el que empiezan a mezclarse las artes digitales con las de la robótica, y con la enorme promesa que suponen los avances en la biología. Si sabemos aprovecharlos, mejorarán nuestras vidas, desde la ayuda en casa a la eliminación de trabajos inhumanos o los progresos en la medicina, entre otros, incluidos el juego y el ocio (que se verá facilitado). Una inteligencia artificial básica (como el texto predictivo en nuestros móviles) ya se puede adelantar a nuestras intenciones cuando utilizamos aparatos, muchos de los cuales se vuelven cada vez más pequeños, portables e intuitivos»<sup>30</sup>.

¿Les suena todo esto? Es lo que nos ha anticipado los mensajes del cine, la ciencia prospectiva y los gurús del transhumanismo. Las máquinas están bien presentes en la sanidad, ya sea en el cuidado de personas o en la mejora de la salud, en la educación, en las fábricas y en las oficinas, en el servicio doméstico, en las misiones militares, etc. Y están para mejorar nuestras vidas.

Al mismo tiempo, también se investiga y nos cuestionamos por las consecuencias de este fenómeno y por su repercusión en nuestras vidas. Ciertamente, la mejora es evidente y será, con el tiempo, muy superior, pero también se escuchan voces que ponen el acento en efectos no deseados. Por ejemplo, en el mundo laboral. ¿Cuántos puestos de trabajo ha destruido o destruirá el proceso de industrialización en la industria o en los servicios? Se hacen hay estudios en los que se evalúa sobre los puestos de trabajo que desaparecerán o los que permanecerán.

Igualmente, es objeto de debate el uso de robots, de drones, etc., en el ejército y, en particular, en la guerra. Aparentemente, desde la perspectiva de quien los utiliza, su uso es positivo pues evita la pérdida de vidas propias, pero ¿sucede lo mismo con el ejército enemigo? Los drones pueden convertirse en asesinos implacables, pero, ¿quién se responsabiliza de esas vidas? Y ¿si hay un fallo? Ciertamente, el uso de robots en las guerras no es un tema pacífico.

La robotización de la sociedad y el proceso de cyborgización de nuestras vidas tienen consecuencias filosóficas, éticas, jurídicas y políticas. El mismo RAY KURZWEIL da en el clavo cuando, al anunciar que «la singularidad está cerca», afirmó: «Con cuerpos versión 3.0 (capaces de moldearse según diferentes formas a nuestro gusto) y nuestros, en gran medida, no biológicos cerebros (no sujetos a la limitada arquitectura que la biología nos ha otorgado), la cuestión sobre lo que significa ser humano será analizada exhaustivamente»<sup>31</sup>. Pero no sólo la noción de lo humano será analizada; también todas las categorías que están implicadas en tan serio debate y éste es esencialmente filosófico y ético, pero tiene profundas repercusiones en la política y en el derecho.

De hecho, en el ámbito jurídico, ya está siendo objeto de controversia alguna de las consecuencias de la paulatina robotización de nuestras sociedades. El caso más problemático es el de la responsabilidad moral y jurídica de los robots o de las máquinas. Sin un debate en profundidad, para KAPLAN, «en la actualidad, podemos encontrarnos con la idea medieval de que los animales son capaces de cometer crímenes ridículos, pero la

---

<sup>30</sup> ORTEGA, A., *La imparable marcha de los robots*, cit., pág. 60.

<sup>31</sup> KURZWEIL, R., *La singularidad está cerca. Cuando los humanos trascendamos la biología*, cit., pág. 390.

interpretación moderna de la capacidad moral está prácticamente limitada a los humanos»<sup>32</sup>. Y, sin embargo, la preponderancia de las máquinas en nuestras vidas bien puede causar males mayores, incluso, la pérdida de vidas humanas, como ha sucedido en diversos accidentes laborales, como puede plantearse con la proyección del automóvil sin conductor o como pasa ya con la utilización de drones en las guerras. ¿Quién es el responsable en estos casos y cuáles son las consecuencias morales, jurídicas y, en consecuencia, económicas?

En opinión de KAPLAN, la respuesta está ya en las mismas categorías del derecho. «Así, la teoría jurídica moderna acepta la idea de que tanto las personas como las empresas pueden tener capacidad moral y, por tanto, ser acusadas de crímenes. ¿Qué sucede con un intelecto sintético? ¿También puede cumplir los requisitos de la responsabilidad moral? Sí que puede. Si posee la suficiente capacidad como para sentir aspectos moralmente relevantes de su entorno y puede elegir acciones, esto lo clasifica como agente moral». Como se ha sugerido, la respuesta está en la creación de una nueva categoría: la de «agentes morales artificiales», lo que permitiría atribuir responsabilidades jurídicas y económicas por las consecuencias de su funcionamiento.

Estas consideraciones nos conducen a una cuestión relevante que ya aparecía en los primeros cuentos de robots de ASIMOV. Y es la de establecer o no un código de conducta para los robots. Lo que en sus historias eran las tres leyes de la robótica que protege a los seres humanos de las facultades y habilidades aumentadas de los robots.

La filosofía moral ya ha dado serios pasos en aras de avanzar, justificar y establecer un código de este tipo. Es lo que se conoce con el nombre de Roboética (*Roboethics*). La Roboética vendría a ser una ética aplicada cuyo objeto es el ámbito científico de la robótica. Según G. VERUGGIO y K. ABNEY, esta nueva ética se construye en torno a tres grandes bloques de temas<sup>33</sup>. En primer lugar, incluye un ámbito de estudio muy relevante y que está planeando en estas páginas: el estudio de los efectos de estos productos tecnológicos (genética, nanotecnología, robótica) en nuestras sociedades. Así pues, tiene que servirnos para lograr el desarrollo y respeto de la dignidad de las personas, tiene que favorecer el perfeccionamiento individual, asegurar las libertades y los derechos fundamentales e impulsar una sociedad más justa, equitativa y solidaria. En segundo lugar, otro importante objeto de análisis son las reglas de conducta que deben incorporarse a las máquinas, especialmente, con el desarrollo de la inteligencia artificial. Dicho de otra forma, el código moral que los programadores deben insertar en los robots y que debe guiar sus acciones, esto es, los límites que deben respetar en sus actuaciones cuyo incumplimiento puede permitir juzgarlos. Incluso, tildarlos de inmorales. Un ejemplo estas reglas serían las conocidas tres leyes de la robótica de ASIMOV. Finalmente, un tercer sentido, no exento de polémica, es el que se refiere al estudio de la posibilidad de autoconsciencia de los robots, de una evolución de la inteligencia artificial que les permita hacer también juicios morales

---

<sup>32</sup> KAPLAN, J., *Abstenerse humanos. Guía para la riqueza y el trabajo en la era de la inteligencia artificial*, cit., p. 54 y para todo lo que viene a continuación el cap. 5.

<sup>33</sup> VERUGGIO, G. y ABNEY, K., «Roboethics: The Applied Ethics for a New Science», en P. LIN, K. ABNEY y G. A. BEKEY, *Robot Ethics. The Ethical and Social Implications of Robotics*, Massachusetts, MIT Press, 2012, págs. 347-364.

como un ser humano, y todas sus consecuencias: tomar decisiones autónomamente, actuar libremente, ser, en suma, conscientes y responsables de sus acciones. Este es el último paso en la evolución de la robótica en la que la inteligencia artificial supera el famoso test de Turing.

Lo cierto es que la sucesiva y lenta penetración de los robots en las sociedades humanas, la realización de la singularidad en sus diferentes fases, etc., origina no pocas cuestiones relacionadas con la filosofía, la ética, el derecho y la política. Y, de paso, con otras disciplinas tanto sociales como experimentales. Realmente, ¿se acerca la singularidad a la velocidad que algunos anuncian? Si es así, ¿se verán afectados todos los seres humanos? ¿En una proporción igual? ¿Cómo afectará a nuestra concepción de lo humano y de lo artificial? ¿El futuro será una sociedad del tipo de las películas *Blade Runner*? ¿Acabarán nuestros replicantes pidiéndonos cuentas por su vida efímera en comparación con los servicios que nos prestan y la superioridad de sus capacidades?

La sociedad humanista, hasta la fecha, ha dado lugar a perennes debates sobre los valores que deben inspirarlas, sobre su estructura y su organización. Sobre el significado y la materialización de los valores mismos. Frente a ello, la sociedad transhumanista implica una importante mutación y un nuevo horizonte. De hecho, el transhumanismo anticipa ya etapas posteriores en la evolución de la sociedad humana de la que surgen indefectiblemente importantes dudas y cuestiones. ¿Cómo organizar, estructurar y coordinar el *modus vivendi* de colectivos tan variados como los humanos, con sus viejos deseos de libertad, igualdad, seguridad y solidaridad, con los medio humanos o ciborgs y la inteligencia artificial?

Los defensores más preclaros del transhumanismo vaticinan la convivencia y la colaboración entre tres tipos de colectividades. Eso sí, a nadie se le presionará para que, de acuerdo con su plan de vida, actúe e interactúe libremente. En efecto, RAY KURZWEIL prevé la existencia de diferentes grupos e intereses: desde los que no quieren aprovecharse de la evolución tecnológica y prefieren desarrollar su vida humana (*humanish*) con todas sus ventajas y limitaciones hasta los que creen que en un futuro próximo (2045) será posible extraer, proyectar, inyectar la experiencia y la conciencia de un ser humano en un *pen drive* y volcarlo en un robot humanoide. La última fase de la singularidad no es sino el tránsito de la biología al ser robótico.

El desarrollo de la genética, la nanotecnología y la robótica (GNR), fundamento de la evolución hacia una sociedad transhumanista, hace prever también que el alargamiento de la vida será una realidad, incluso, para aquellos que no quieran la inmersión en lo universal, esto es, la supuesta inmortalidad que predicen los popes del transhumanismo como RAY KURZWEIL y DAVID PEARCE, pero también NICK BOSTROM y otros. En suma, en un futuro, quizás no muy lejano, el ser humano podrá alargar su vida durante décadas e, incluso, siglos, gozará de orden, estabilidad y bienestar, podrá dedicar su tiempo a mejorar sus cualidades y sus habilidades, a perfeccionarse como sujeto, a ser feliz, en suma. Podrá materializar sus sueños personales y colectivos y lograr la utopía social deseada. Incluso, podrá lanzarse a la conquista de la galaxia y del universo, como en los cuentos de ASIMOV. ¿Cómo no van a afectar todos estos cambios a la vida social, a nuestras relaciones con los demás, a nuestros deseos, a nuestros planes de vida, etc.!

Todas estas cuestiones y estos asuntos no pueden quedar circunscritos únicamente a la respuesta proveniente de una disciplina como la Roboética dada la amplitud de las previsible consecuencias. No sé si los promotores deberían desarrollar una filosofía o una ética específica para resolver la multitud de cuestiones que surgen y surgirán. Ya sea la propia filosofía transhumanista o una nueva filosofía o ética de la singularidad. Pero las cuestiones y sus repercusiones son numerosas y muy relevantes y requieren una visión global de realidades sumamente complejas.

Al menos, para empezar, creo que las siguientes categorías de la ética y del derecho quedan radicalmente afectadas por estas realidades. De lo cual, por tanto, derivan importantes interrogantes:

– **Identidad.** Nuestra civilización está basada en una fuerte consideración del «yo» hasta el punto de que, en la actualidad, nuestras sociedades son tremendamente narcisistas; se cultiva la individualidad y el libre desarrollo de la personalidad como motores del desarrollo general. Pero, éste no deja de ser un «yo» corporal, biológico y, por tanto, finito. ¿Qué pasará cuando este «yo» humano derive en un «yo» artificial, robótico, en un cyborg? ¿Podrá mantenerse la existencia de una misma identidad o, por el contrario, las variaciones producto de la tecnología implicarán cambios substanciales de nuestra identidad, de nuestro «yo»? Todo parece indicar que nuestro «yo» biológico no sólo continuará su existencia, sino que además será superado por el «yo» no biológico: «Nuestros no biológicos del futuro serán muchísimo más inteligentes, de manera que exhibirán en un grado muy superior las refinadas cualidades del pensamiento humano»<sup>34</sup>.

– **Libertad.** Pero, si esto es así, el futuro transhumanista afectará también a conceptos tan importantes como el de libertad o el de igualdad. El de libertad porque ya hace tiempo que la teoría de los derechos humanos distinguió entre las diferentes formas de libertad: libertad formal, libertad material y libertad real, dando lugar así a diferentes formas de comprender la democracia. En última instancia, que unos y otros individuos desarrollen sus capacidades y sus posibilidades de prolongar su vida de acuerdo a la disponibilidad, mayor o menor, de acceso a las formas de singularidad, determina el concepto y el ejercicio de las libertades personales. Hay experiencias históricas de que esto ha sido así.

– **Igualdad.** Otro tanto puede decirse de la igualdad. ¿Podremos hablar de igualdad, de igualdad en la satisfacción de necesidades básicas, en el bienestar, en las posibilidades de perfeccionamiento personal, de participación en la vida social, en las responsabilidades públicas, de la igualdad de oportunidades, en suma, si la sociedad del futuro afianza una sociedad dividida entre los *humanish*, los cyborg, las máquinas, los robots humanoides?

– **Responsabilidad.** La sociedad humana ha evolucionado enormemente en la atribución de responsabilidades, elemento clave del devenir moral, jurídico y social. Gracias a este desarrollo, es posible la atribución de consecuencias por los actos realizados. La sociedad de los seres humanos empezó atribuyendo esta responsabilidad a las personas físicas y, pasado al tiempo, también lo hizo a las personas artificiales, a las personas

---

<sup>34</sup> KURZWEIL, R., *La singularidad está cerca. Cuando los humanos trascendemos la biología*, cit., pág. 434.



jurídicas, es decir, a organizaciones, asociaciones, instituciones públicas o privadas. Nada hace dudar que también se dará el paso de resolver las responsabilidades de los daños o de las consecuencias producidas por las acciones de los seres artificiales o de las máquinas.

– **Tolerancia.** Para algunos, la tolerancia es una virtud básica de la convivencia en nuestras sociedades plurales. Para otros, una vez que se han reconocido y positivado en las constituciones los derechos y libertades individuales, la tolerancia debe dejar paso a la protección y materialización de éstos. A pesar de todo, la tolerancia tiene un papel importante en el discurso político en nuestras sociedades multiculturales. Pero, ¿se logrará en un futuro cada vez más próximo que se respeten las diferencias de capacidades derivadas de la aplicación de la tecnología a las personas? La serie de películas X-Men demuestra el recelo, si no la animadversión, hacia quienes poseen poderes muy superiores a los humanos.

– **Derechos humanos.** ¿Se podrá atribuir, en un futuro, derechos, los mismos derechos que a los seres humanos, a entes semi o no biológicos? Las discusiones en la ética y en la teoría de los derechos humanos se orientan hacia el debate de si los seres no humanos pueden tener derechos. Véase los animales o, incluso, seres inanimados. Por ello, nada impide que en un futuro el debate se extienda a la inteligencia artificial, a los robots, etc. Sobre todo, si, al mismo tiempo, de les atribuye el carácter de agentes morales y, como consecuencia de ello, devienen en seres con responsabilidad.

Para concluir, quizás, convendría retener en la memoria la cruda moraleja de una película como *Metrópolis* (1927), a la que ya he hecho referencia. Sus imágenes, su trama y su final feliz. Pero, también esa alianza entre el capitalismo más explotador y la tecnología.