

EL IMAGINARIO SIMBÓLICO DE LA IA

Ciencia-ficción: un espejo para el futuro de la humanidad

El interesante debate sobre IA y ética que propicia la ciencia-ficción

Carme Torras



[Araya Peralta](#)

Las tecnologías digitales se han convertido en parte de nuestras vidas cotidianas y cada vez más actúan como intermediarias en nuestros lugares de trabajo y en las relaciones personales, incluso sustituyéndolas. El internet de las cosas, las redes sociales, los programas capaces de aprender interactuando con humanos, los asistentes artificiales y los robots de compañía, los juegos de ordenador, los juegos con impacto social, los asesores financieros semiautomatizados, las páginas web que ofrecen la inmortalidad digital, etcétera. Estas herramientas pueden, en un breve lapso de tiempo, modificar el mercado de trabajo, acabar con la reputación de cualquiera, transformar un barrio urbano, cambiar nuestras relaciones personales -no sólo en el trabajo, también dentro de la familia y en nuestras relaciones más íntimas- o magnificar lo que uno deja atrás cuando se muere, que ahora incluye lo que se conoce como huella digital.

La creciente interacción con máquinas *inteligentes* no es sólo un paso más allá en la transformación social que arrancó con la revolución industrial. Aunque estas nuevas tecnologías de la información también liberan los humanos de tareas repetitivas y les proporcionan más tiempo para invertirlo de forma creativa y divertida, la diferencia es que ahora se adentran en ámbitos que siempre se han considerado exclusivos de los humanos,

como la toma de decisiones, las emociones y las relaciones sociales, y esto puede comprometer los valores humanos al mismo tiempo que claramente da forma a la sociedad y a nuestra manera de vivir.

El estado de cosas actual, sobre el que no tenemos experiencia previa, genera una serie de preguntas éticas que no fueron pertinentes para otros tipos de máquinas del pasado; tampoco podemos predecir con fiabilidad como acabará influyendo este nuevo marco en la evolución de la humanidad. Esto ha llevado a la confluencia de la inteligencia artificial (IA) con las Humanidades, en un debate ético que está empezando a dar sus primeros frutos, no sólo con el establecimiento de regulaciones y estándares, sino también con iniciativas en ámbitos como la docencia universitaria, la mejora profesional y la conformación de la opinión pública. Curiosamente, la ciencia ficción a menudo tiene una función predictiva notable, destacando los pros y los contras de los posibles escenarios del futuro. En palabras del famoso escritor Neal Stephenson [1]: «la buena ciencia ficción ofrece una descripción plausible y bien estudiada de una realidad alternativa donde irrumpe una atrayente innovación tecnológica. El universo de la buena ciencia ficción se caracteriza por una coherencia y una lógica internas que resultan convincentes para científicos e ingenieros».

Confluencia entre la IA y las Humanidades

La comunidad de investigación en IA ha visto la necesidad de trabajar conjuntamente con científicos sociales, psicólogos, abogados, filósofos y antropólogos para analizar las implicaciones sociales y éticas de las tecnologías que se están desarrollando y la forma como se están desplegando. Se han creado equipos multidisciplinares y se ha dado luz verde a proyectos conjuntos de investigación. Las administraciones e instituciones han querido tomar el pulso a la opinión pública y se han establecido comités con el objetivo de que contengan la situación y generen directrices legales y estándares comunes.

Considerando este último punto, varias organizaciones y sociedades profesionales están desarrollando regulaciones para los programadores en IA, diseñadores en robótica, empresas y usuarios. En abril de 2019, el Grupo de Expertos de Alto Nivel sobre IA de la Comisión Europea presentó el informe Directrices éticas para una IA fiable y, ese mismo año, la Asociación de Estándares IEEE publicó el informe Diseño alineado éticamente: priorizar el bienestar humano con sistemas autónomos e inteligentes, que había sido sometido a debate público desde diciembre de 2017. El resultado de la consulta fueron unas trescientas páginas con contenidos muy relevantes que posteriormente revisaron un millar de expertos de ámbitos muy diversos, desde los negocios, pasando por la universidad hasta las aseguradoras. Anteriormente, había habido otras iniciativas pioneras más modestas, como la Declaración de Barcelona para el desarrollo y el uso adecuados de la IA en Europa y la Declaración de Montréal para un Desarrollo Responsable de la Inteligencia Artificial, ambas durante el 2017.

La regulación legal es necesaria para desarrollar una IA ética, pero aún es más importante la educación a todos los niveles, a fin de desarrollar concienciación crítica y formar la opinión pública. Un objetivo crucial es el colectivo de diseñadores de widgets de IA. No es

sorprendente que en el ámbito universitario, voces conocidas como la de la de profesora Barbara J. Grosz defiendan integrar la ética en el plan de estudios de informática. «Hacer del razonamiento ético un elemento central de los estudios hará que el alumnado no sólo aprenda a pensar sobre la tecnología que puede diseñar, sino a plantearse si debería diseñarla» [2]. En este sentido, asociaciones prestigiosas como la ACM y el IEEE incluyen 18 áreas del saber en sus curriculums de informática, entre los que hay *Aspectos sociales y práctica profesional* para que «el alumnado desarrolle el hábito de pensar sobre los aspectos sociales, éticos, legales y profesionales pertinentes».

En Europa, la Universidad de Oxford incluso ha ido más allá en la confluencia con las Humanidades, y ofrece un título en Informática y Filosofía, centrado en los intereses comunes entre la inteligencia artificial, la lógica, la robótica, la realidad virtual y la ética. Este tipo de grados mixtos pronto proliferarán en todo el mundo; en concreto, en Catalunya varias universidades ya tienen iniciativas que van en esta dirección.

La función de la ciencia ficción

La ciencia ficción ha sido una fuente de inspiración para los científicos y tecnólogos desde su nacimiento. Trabajos clásicos como los de Shelley, Verne y Huxley, entre muchos otros, ya concebían experimentos, investigaciones, descubrimientos e innovaciones que más tarde han hecho realidad, como la exploración espacial, las réplicas humanas y la ingeniería genética. Hoy, la ciencia ficción se ha ganado la atención pública a raíz de las tecnologías digitales; en especial la IA y la robótica. En los últimos años han sido habituales los contactos entre escritores y directores de cine de ciencia ficción, por un lado, y científicos y desarrolladores de tecnología, por otro lado.

Como ejemplo, en 2012 se estableció el Centro para la Ciencia y la Imaginación en la Universidad Estatal de Arizona. La idea surgió a raíz del provocador discurso que el escritor Neal Stephenson [3] hizo en presencia del rector, en el que afirmaba que hoy los científicos han perdido la capacidad de pensar y hacer «cosas grandes», como las que en el pasado habían inspirado el programa espacial Apolo o el microprocesador. El rector respondió que quizás los responsables de la situación eran los escritores de ciencia ficción, que habían dejado de sugerir futuros estimulantes que llevaran a los científicos a querer hacerlos realidad. El resultado fue el nacimiento del Centro, que incluye equipos de investigadores en Ciencias y Humanidades que trabajan conjuntamente en proyectos que moldearán nuestro futuro.

Uno de los proyectos del Centro es la continuación de una iniciativa de la compañía Intel, el proyecto *Tomorrow*, en el que se encargó a cuatro escritores que escribieran relatos de ciencia ficción sobre los posibles usos que en el futuro se podrían dar a los productos de la compañía en fotónica, robótica, telemática y sensores inteligentes. El libro es de acceso abierto [4]. Desde entonces, y fruto del trabajo del Centro, han aparecido varios volúmenes con propuestas de solución para los mayores retos a los que la humanidad se enfrenta hoy, explorados a través de la escritura y las artes visuales.

Son innumerables las iniciativas surgidas en esta dirección: talleres, simposios, exposiciones, performances...; incluso revistas científicas de gran impacto han publicado dossiers especiales dedicados a la producción en ciencia ficción. Una pionera fue la prestigiosa revista *Nature*, que para conmemorar el cincuenta aniversario de la hipótesis de Hugh Everett III sobre los universos paralelos fue publicado un volumen titulado *Many Worlds* [5], con artículos tanto de investigadores en mecánica cuántica como de escritores de ciencia ficción. La introducción ya decía muy claramente qué función puede tener la ciencia ficción enumerando las ventajas y los inconvenientes del desarrollo científico. «La ciencia ficción seria se toma la ciencia en serio. [...] La ciencia ficción no nos dice lo que nos llevará el futuro, pero sí que nos ayuda a hacernos una idea y nos plantea como podríamos sentirnos cuando el modo habitual que tenemos de mirar el mundo sea sustituido por otro».

Enseñar y fomentar el debate ético a través de los relatos de ciencia ficción

Enseñar ética en IA dista considerablemente de enseñar otras materias en Ciencias de la Información e Ingeniería. No es cuestión de que el alumnado aprenda unos contenidos concretos, sino de hacer que sean conscientes de las implicaciones sociales y éticas de su futuro trabajo y de entrenarlos a analizar y debatir sobre estas cuestiones. A menudo, los alumnos ya tienen múltiples valores que son discrepantes, y el objetivo no es que acaben pensando todos lo mismo, sino concienciarlos para que cultiven la capacidad de reflexionar y debatir. Además, los alumnos en el ámbito de la tecnología no son filósofos. Aunque hay teorías éticas de las que deberían tener nociones básicas, los textos filosóficos son a menudo demasiado abstractos, y da mejor resultado adoptar una vía pragmática.

Según Sullins [6], las teorías éticas pertinentes para la IA y la interacción robot - humano son: consecuencialismo o utilitarismo (maximizar el número de personas que disfrutan de los máximos beneficios), deontologismo (actuar sólo según máximas que podrían convertirse en leyes universales), ética de la virtud (confiar en el carácter moral de individuos virtuosos), justicia social (todos los seres humanos merecen ser tratados como iguales; debe haber una justificación sólida para el trato diferencial), el bien común (vivir en comunidad implica ciertas restricciones para el individuo), ética religiosa (las normas derivan de una autoridad espiritual) y ética de la información (reglamentos y códigos que rigen la creación, la organización, la divulgación y el uso de la información).

Como ninguna teoría por sí misma puede dar respuesta a todos los aspectos éticos que se derivan del diseño y el uso de las innovaciones técnicas, la opción más pragmática es adoptar una aproximación híbrida. Esta ética híbrida es defendida por Wallach y Allen [7] como combinación de teorías que van de arriba a abajo (esto es, las que usan principios racionales para derivar las normas éticas) y las que van de abajo hacia arriba (es decir, las que infieren las normas generales a partir de situaciones concretas).

Ahora bien, ¿dónde encontramos estas situaciones concretas? Stephenson [8] afirma que «lo que las narraciones de ciencia ficción hacen mejor que ninguna otra cosa es dar no sólo

una idea de innovación tecnológica concreta, sino también ofrecer un marco coherente de cómo aquella innovación será integrada en la sociedad, la economía y la vida de la gente». Así, algunos cursos sobre ética y tecnología recurren a los relatos de ciencia ficción para ejemplificar situaciones concretas que permitan visualizar el conflicto ético. Los temas explorados en las obras clásicas de Asimov, Dick, Bradbury, Orwell, Huxley, Hoffman, Shelley, Capek, Wells, Sturgeon, Silverberg o Keyes, como las tres leyes de la robótica, las niñeras robóticas, el debate sobre seguridad versus libertad, la pérdida de la privacidad, el totalitarismo tecnológico, los sustitutos emocionales, las réplicas humanas, la incidencia en el mercado laboral, la responsabilidad moral, la pérdida de control por parte de los humanos, los sesgos de la alta tecnología, la distancia entre la manipulación y la automatización o el mejoramiento humano y el posthumanismo son de gran relevancia a raíz del desarrollo de la IA.

Dada esta relevancia, es natural que los docentes de ética en los grados tecnológicos recurran a este tipo de relatos para ejemplificar situaciones a las que el alumnado pueda ser sensible o a las que tendrán que hacer frente en la práctica profesional, a fin que reflexionen o debatan. Tras cinco ediciones del curso *Ciencia ficción y ética computacional* en la Universidad de Kentucky y dos en la Universidad de Illinois, Chicago, Burton *et al.* [9] afirman que, en su experiencia, «recurrir a la ficción para enseñar ética permite al alumnado hablar y razonar con seguridad sobre cuestiones difíciles y emocionalmente connotadas sin llevar el debate al terreno personal». Además, destacan que los relatos son alentadores para los estudiantes, y dan cuenta de descubrimientos muy interesantes hechos a lo largo de los años que merecen una atenta lectura.

La ciencia ficción contemporánea hace referencia a muchas de las cuestiones éticas de las novelas clásicas, pero enfocadas desde una perspectiva actual: por ejemplo, las preocupaciones de hoy en día a raíz de las tecnologías digitales, el uso intensivo de los teléfonos móviles, el extendido uso de las redes sociales, la toma automática de decisiones guiada por artilugios de inteligencia artificial, los juegos de inmersión en realidades virtuales, los algoritmos que hacen uso del Big Data... Es por todo esto que se generan debates interesantes [10]. La figura 1 muestra algunas series y películas centradas en estas cuestiones. La serie *Black Mirror* es una obra maestra que, en cada capítulo, lleva una tecnología particular a sus consecuencias extremas, y la película *Her* retrata a un hombre que se enamora del sistema operativo de su ordenador, trasladando el título *The Sandman* de Hoffmann a un paradigma digital contemporáneo. En cuanto a robótica, destacaría la serie *Real Humans*, donde robots casi-humanos coexisten con humanos y a menudo compiten con ellos, y la película *Surrogates*, donde cada ciudadano tiene un avatar controlado desde casa que puede mover por la ciudad y que interactúa con humanos. La película *Robot and Frank* muestra la relación entre un Frank envejecido y su cuidador artificial, y merece especial mención por su realismo y valor educativo: fue la base para un curso en línea sobre ética y robótica a través de la plataforma 'Teach with Movies' [11], entre otros.



Figura 1. Series y películas de ciencia ficción que tratan cuestiones éticas generadas por las tecnologías digitales, especialmente la IA y la robótica.

En cuanto a libros pensados para la enseñanza de cursos de ética en grados tecnológicos a través de la ciencia ficción, resaltaría los de Murphy [12], Nourbakhsh [13] y Torras [14]; ver en la figura 2. El editado por Murphy cuenta los principios clave de la IA; los relatos que incluye -de Asimov, Vinge, Aldiss i Dick- tratan sobre las cuestiones de la telepresencia, la robótica especializada en la imitación del comportamiento, la deliberación, la validación, la interacción robot-humano, el valle inquietante, el procesamiento de lenguas naturales, el aprendizaje automático y la ética. Cada narración es precedida por una nota introductoria, 'Mientras lees', y va seguida de una discusión sobre las posibles derivadas, 'Después de la lectura'.

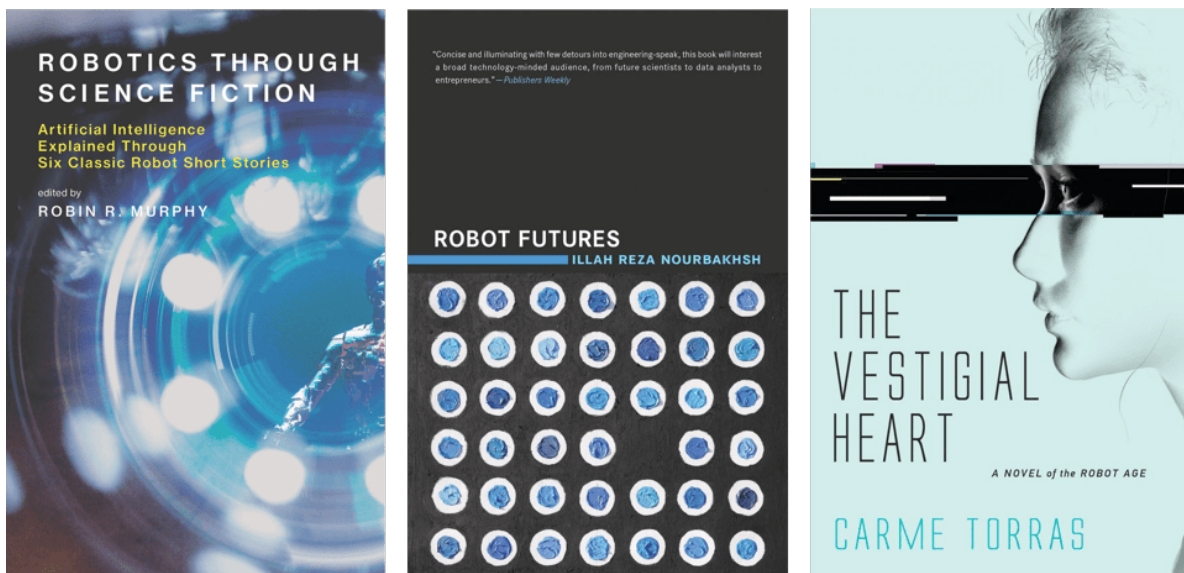


Figura 2. Tres libros publicados por MIT Press especialmente pensados para estudiantes tecnológicos y basados en relatos de ciencia ficción.

Los otros dos libros a los que me referiré difieren del primero en dos aspectos: por un lado, los autores son también autores de los relatos de ciencia ficción utilizados como ilustración y, por otro lado, los libros no incluyen explicaciones sobre principios técnicos, pero buscan explícitamente fomentar el debate sobre las implicaciones sociales y éticas de las tecnologías de la información y la comunicación.

Con la descripción de posibles escenarios de futuro, Nourbakhsh genera preguntas sobre hacia dónde estamos avanzando, y lo hace sin adoptar ni un punto de vista regulador ni un tono marcadamente pedagógico. Como famoso investigador en robótica, el autor hace comentarios muy lúcidos sobre temas como las estrategias de marketing en la red; consecuencias del diseño no efímero; juguetes robóticos voladores que operan siguiendo la mirada; telepresencia multimodal y multicontinental habilitada mediante un robot e incluso la posibilidad de adoptar diferentes formas físicas gracias a los nanorobots. Nourbakhsh examina la tecnología en que se basa cada widget y describe las consecuencias sociales de cada escenario. Pero también ofrece una réplica: una robótica diseñada para fortalecer los lazos cívicos y comunitarios.

En virtud de mi investigación en asistentes artificiales [15], me interesé progresivamente por las implicaciones sociales y éticas de las tecnologías que estamos desarrollando, concretamente con el objetivo de encontrar maneras de enseñar ética a los estudiantes con perfiles tecnológicos. Esto me llevó a probar mi habilidad para escribir ficción: en la novela *La mutación sentimental* (Torras, 2018) exploré como el hecho de ser criado por niñeras artificiales, o de tener maestros robóticos y de compartir trabajo y ocio con programas de IA, afectaría los hábitos intelectuales, emocionales y sociales de las generaciones futuras. El *leitmotiv* de la novela es una cita del filósofo Robert C. Solomon (1977): «son las relaciones que hemos construido las que a su vez nos modelan» [16]. Él se refería a las relaciones humanas con nuestros padres, maestros y amigos, pero la frase se puede aplicar también a los asistentes artificiales y a todo tipo de dispositivos interactivos, si es que tienen que estar tan presentes en nuestras vidas.

La trama de la novela es la siguiente: Celia, una chica de trece años congelada a quien criogenizaron porque sufría una enfermedad terminal, es curada y llevada de nuevo a la vida en pleno siglo XXII para ser adoptada. A parte de los recuerdos, acarrea algo más del pasado: sensaciones y emociones que ya no existen en el nuevo mundo. Esto es lo que más atrae a Silvana, una mujer de mediana edad que trabaja como masajista emocional tratando de recuperar las sensaciones que los humanos han perdido. Los sentimientos de Celia son también un preciado material de investigación para Leo, un bioingeniero que está diseñando una prótesis de creatividad para el misterioso Dr. Craft, propietario de la importante compañía de robótica CraftER.

Siguiendo una sugerencia de la editora de MIT Press Marie L. Lee, incluí en el libro un apartado final con 24 preguntas de contenido ético y apuntes para generar debate sobre las situaciones que aparecen en la novela. Se publicó conjuntamente con una guía en línea para el profesor y una presentación de 100 diapositivas para impartir un curso sobre ética en robótica social e inteligencia artificial. El libro cubre seis temas: cómo diseñar el «asistente» perfecto; la importancia del aspecto y de la simulación de emociones en los robots para ser aceptados; la función de los programas de IA en el lugar de trabajo y en el entorno educativo; el dilema entre la toma de decisiones automáticas y la libertad y dignidad humanas, y la responsabilidad implícita de alojar cierto sentido moral en los programas informáticos y los robots.

Cada capítulo de la guía del profesor sigue la misma estructura: arranca con alguna sección destacada de la novela para pasar a introducir después el marco ético de la situación conflictiva, seguida de cuatro preguntas y puntos para estimular el debate. El capítulo se cierra con algunas cuestiones ya tratadas en capítulos anteriores. En la figura 3 aparecen algunas diapositivas de muestra de la estructura de estos materiales docentes.

Teaching materials
Ethics in Social Robotics and AI

Appendix in "The Vestigial Heart":

1. Designing the "perfect" assistant: Chapters 1, 5
2. Robot **appearance** and emotion: Chapters 9,10,12
3. Robots in the **workplace**: Chapter 13
4. Robots in **education**: Chapters 14, 16
5. Human-robot **interaction** and human dignity: Chapters 25, 28
6. **Social responsibility** and robot morality: Chapter 30

Teaching materials
Ethics in Social Robotics and AI

5. Human-robot interaction and human dignity

- 5.1. Highlights from *The Vestigial Heart*
- 5.2. Ethical Background and Discussion:
 - Four questions
 - Hints for a debate on each question
- 5.3. Revisiting Issues
- 5.4. Scholarly References for Further Reading

Designing the "perfect" assistant
1.2. Questions

1.A - Should public trust and confidence in AI/robots be enforced? If so, how?

1.B - Is it admissible that devices be designed to generate addiction?

1.C - Should the possibility of deception be actively excluded at design time?

1.D - Could AI be used to control people?

Human-robot interaction and human dignity
5.2. Questions

5.A - Could automatic decision-making undermine human freedom and dignity?

5.B - Is it acceptable for robots to behave as emotional surrogates? If so, in what cases?

5.C - Could robots be used as therapists for the mentally disabled?

5.D - How adaptive/tunable should programs and robots be? Are there limits to human enhancement by technology?

"Ethics in Social Robotics" based on *The Vestigial Heart* @ MIT Press, 2018 16/104

"Ethics in Social Robotics" based on *The Vestigial Heart* @ MIT Press, 2018 66/104

Figura 3. Muestra de diapositivas de los materiales docentes asociados con el libro *La mutación sentimental*: en la parte superior, correspondencia entre las secciones en que se distribuye el contenido ético y los capítulos de la novela, y estructura de cada sección; en la parte inferior, algunas de las preguntas planteadas.

El libro y los materiales de apoyo docente han sido utilizados no sólo en los grados de Ciencias de la Información e Ingeniería, sino también en los estudios de Filosofía y Administración de empresas de varias universidades del mundo; de momento, principalmente en España y EEUU, pero algunas universidades europeas ya han mostrado interés para poder contar con ellos en los próximos cursos. Además, hay una versión en catalán de los materiales adaptada a la educación secundaria y disponible en la web de Pagès Editors, que también puede servir para generar debate en grupos de lectura, paneles, mesas redondas y otros foros similares de discusión.

A modo de conclusión

La creciente interacción de los humanos con programas de IA en todo tipo de contextos cotidianos genera importantes retos sociales y éticos que muy probablemente modificarán nuestro futuro. Esto exige que los grados técnicos se acerquen a las Humanidades, por lo que alumnos y profesionales tienen que ser conscientes de los posibles aspectos con los que tendrán que lidiar a lo largo de su carrera y es necesario que aprendan a definirlos y a debatir sobre ellos.

Hemos descrito tres propuestas recientes de investigadores en el ámbito de la IA y la robótica: un libro que utiliza relatos clásicos para explicar los principios computacionales y

dos libros que confían en la ciencia ficción contemporánea para transmitir a los alumnos el conocimiento de la tecnoética, con la característica singular que las ficciones que contienen son también obra de los mismos investigadores. Así, los relatos ilustran exactamente los aspectos que los autores quieren tratar, y el material es pedagógicamente motivador tanto para lectores con perfil tecnológico como para los docentes y probablemente también para cualquier lector aficionado a leer sobre alternativas de futuro.

Concluiré diciendo que la mayoría de la literatura de ciencia ficción es respetuosa con la ciencia e intenta proyectar los hitos científicos hacia el futuro. Parece que la ciencia empieza ahora a respetar también esta literatura visionaria y a encontrar en ella cierto objeto de inspiración. La confluencia entre los dos ámbitos podría ser sumamente productiva y es una muy buena noticia, dado que abre perspectivas interesantes para el futuro inmediato.

RECONOCIMIENTOS

Este trabajo ha sido parcialmente posible gracias al apoyo del Consejo de Investigación Europeo, bajo el programa de investigación e innovación Horizonte 2020 de la Unión Europea, a través del proyecto CLOTHILDE - CLOTH manipulation Learning from DEMonstrations (acuerdo de concesión núm. 741930), y la Agencia de Investigación Española, a través del Sello de Excelencia María de Maeztu concedido al IRI (MDM-2016 hasta 0656).

REFERENCIAS

- 1 — Stephenson, N. (2011) «Innovation starvation». *World Policy Journal*, 28(3), 11-16. Disponible [online](#).
- 2 — Grosz, B.J.; Grant, D.G.; Vredenburg, K.; Behrends, J.; Hu, L.; Simmons, A. y Waldo, J. (2019) «Embedded EthiCS: integrating ethics across CS education». *Communications of the ACM*, 62(8), 54-61.
- 3 — Stephenson, N. (2011) «Innovation starvation». *World Policy Journal*, 28(3), 11-16. Disponible [online](#).
- 4 — Rushkoff, D.; Hammond, R.; Thomas, S. i Markus, H. (2012) *The Tomorrow Project*. Els millors escriptors descriuen la vida quotidiana en el món del futur. Intel: Santa Clara.
- 5 — Everett III, H. (2007) «Many Worlds» *Nature*, 448(7149): 1-104.
- 6 — Sullins, J.P. (2015) «Applied professional ethics for the reluctant roboticist». Dentro de *The Emerging Policy and Ethics of Human-Robot Interaction*, editado por L.D. Riek, W. Hartzog, D. Howard, A. Moon y R. Calo. Taller de trabajo dentro del X Congreso Internacional del ACM/IEEE sobre la Interacción Humanos-Robots, Portland.
- 7 — Wallach W. y Allen C. (2008) *Moral machines: Teaching robots right from wrong*. Oxford: Oxford University Press.
- 8 — Stephenson, N. (2011) «Innovation starvation». *World Policy Journal*, 28(3), 11-16. Disponible [online](#).
- 9 — Burton, E.; Goldsmith, J. y Mattei, N. (2018) «How to teach computer ethics through science fiction». *Communications of the ACM*, 61(8), 54-64.

- 10 —
- Torras C. (2018) «Social networks and robot companions: Technology, ethics and science fiction». *Metode Science Studies Journal*, 99: 47-53.
 - Torras C. y López de Mántaras, R. (ed.) (2019) «Interlinked: machines and humans facing the 10101 century». *Metode Science Studies Journal - Annual Review*, 9.
- 11 — Teach with Movies (2012) «Robot Ethics Using Clips from *Robot and Frank*». Disponible [online](#).
- 12 — Murphy, R. (ed.) (2018) *Robotics Through Science Fiction: Artificial Intelligence Explained Through Six Classic Robot Short Stories*. MIT Press: Cambridge, Massachusetts.
- 13 — Nourbakhsh, I.R. (2013) *Robot Futures*. MIT Press: Cambridge, Massachusetts.
- 14 — Torras C. (2018) *The Vestigial Heart. A Novel of the Robot Age*. MIT Press: Cambridge, Massachusetts (libro del profesor y presentación de 100 diapositivas como material de soporte para el docente (en inglés), disponible [online](#)).
- 15 —
- Torras C. (2016) «Service robots for citizens of the future». *European Review*, 24(1), 17-30.
 - Torras C. (2019) «Assistive robotics: Research challenges and ethics education initiatives». *DILEMATA: International Journal of Applied Ethics*, 30: 63-77.
- 16 — Solomon, R.C. (1977) *The Passions*. Nueva York: Anchor Press / Doubleday.



Carme Torras

Carme Torras es novelista y profesora investigadora del Instituto de Robótica e Informática Industrial, un centro de investigación conjunto del Consejo Superior de Investigaciones Científicas (CSIC) y la Universidad Politécnica de Catalunya (UPC). Es licenciada en Matemáticas e Informática por la Universidad de Barcelona y la Universidad de Massachusetts, respectivamente, y tiene un doctorado en Informática por la UPC. Es IEEE Fellow, Eura Fellow, integrante de la Academia Europea y miembro de la Real Academia de Ciencias y Artes de Barcelona. Compagina la dirección de un grupo de investigación en robótica asistencial con su dedicación a la escritura literaria y a la promoción de la ética en la aplicación de las nuevas tecnologías. Ha escrito varios libros y artículos y también es autora de novelas de ciencia-ficción relacionadas con la robótica y la inteligencia artificial.