



Revista Iberoamericana de Ciencia, Tecnología y Sociedad
- CTS

ISSN: 1668-0030

ISSN: 1850-0013

secretaria@revistacts.net

Consejo Nacional de Investigaciones Científicas y Técnicas
Argentina

Terrones Rodríguez, Antonio Luis

Inteligencia artificial, responsabilidad y compromiso cívico y democrático

Revista Iberoamericana de Ciencia, Tecnología y Sociedad
- CTS, vol. 15, núm. 44, 2020, -Junio, pp. 253-276

Consejo Nacional de Investigaciones Científicas y Técnicas
Argentina

Disponible en: <https://www.redalyc.org/articulo.oa?id=92463902013>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UAEH  redalyc.org

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

Inteligencia artificial, responsabilidad y compromiso cívico y democrático *

Inteligência artificial, responsabilidade e compromisso cívico e democrático

Artificial Intelligence, Responsibility and Civic and Democratic Commitment

Antonio Luis Terrones Rodríguez **

Este artículo parte del reconocimiento de los profundos cambios que la inteligencia artificial está provocando en diversas esferas de la vida humana e invita a una reflexión sobre la necesidad de incorporar criterios éticos de responsabilidad en la práctica tecnológica. Para ello se toma como punto de partida el terreno político, por entender que es un terreno fértil desde el que formular el concepto de inteligencia artificial responsable. Este concepto de responsabilidad tecnológica se expresa desde la preocupación por los asuntos políticos humanos relativos a los derechos humanos y los objetivos de desarrollo sostenible y concibe que los avances en materia de inteligencia artificial deben incorporar a sus proyectos criterios de responsabilidad que sepan reconocer el valor y la necesidad de fortalecer las habilidades cívicas y la democracia desde la tecnología.

253

Palabras clave: inteligencia artificial; responsabilidad; democracia; derechos humanos; objetivos de desarrollo sostenible

* Recepción del artículo: 15/02/2019. Entrega de la evaluación final: 23/06/2019.

** Profesor titular en la Escuela de Filosofía de la Pontificia Universidad Católica del Ecuador. Correo electrónico: antonioluis.terrones@gmail.com. Este artículo fue elaborado en el marco del proyecto de investigación "Ética normativa de la existencia (primera fase)", desarrollado por la Escuela de Filosofía de la Pontificia Universidad Católica del Ecuador, bajo la dirección de Dennis Schutjser.

Este artigo parte do reconhecimento das profundas mudanças que a inteligência artificial está causando em diversas esferas da vida humana e convida à reflexão sobre a necessidade de incorporar critérios éticos de responsabilidade na prática tecnológica. Para isso, o terreno político é tomado como ponto de partida, pois é um terreno fértil para formular o conceito de inteligência artificial responsável que representa o elemento central deste texto. Esse conceito de responsabilidade tecnológica é expresso a partir da preocupação pelos assuntos políticos humanos relacionados aos direitos humanos e aos objetivos de desenvolvimento sustentável, e concebe que os avanços na inteligência artificial devem incorporar em seus projetos critérios de responsabilidade que saibam reconhecer o valor e a necessidade de fortalecer as habilidades cívicas e a democracia a partir da tecnologia.

Palavras-chave: inteligência artificial; responsabilidade; democracia; direitos humanos; objetivos de desenvolvimento sustentável

This paper starts by acknowledging the profound changes that artificial intelligence is causing in the different spheres of human life and reflects on the need for adopting ethical criteria regarding responsibility within technological practice. To this end, the political arena is taken as a starting point, as it is seen as a fertile ground on which to establish the concept of responsible artificial intelligence. This concept of technological responsibility is expressed from the concern for the human political affairs related to human rights and the sustainable development goals, and conceives that the advancements in artificial intelligence must include responsibility criteria that can recognize the value and need for strengthening civic capabilities and democracy through technology.

Keywords: artificial intelligence; responsibility; democracy; human rights; sustainable development goals

Introducción

Los asombrosos avances que está experimentando el campo de la inteligencia artificial (IA) en los últimos años nos adentran en un nuevo escenario que presenta cambios revolucionarios en la vida humana. Dichos avances están impregnando las diversas esferas que conforman la vida humana —profesional, política, social, económica, etc.— y nos sitúan en una etapa de transición a un nuevo tiempo que depara asombrosos desafíos. Esta nueva época que marcará la historia de la humanidad no se encuentra exenta de un cierto grado de inquietud y a la vez de esperanza. Provoca el planteo de diversos cuestionamientos que son fruto de la incertidumbre tecnológica.

Frente a esa incertidumbre, que es a su vez fruto del poder de la tecnología avanzada, surge la necesidad de crear un discurso en el que la filosofía juegue un papel esencial. La tarea de la filosofía, entre otras, puede ser la de invitar a la reflexión en el campo de la IA, pues en ocasiones la lógica tecnológica de mercado ocasiona que muchas cuestiones relevantes y decisivas para nuestras vidas no sean objeto de preocupación. En ese sentido, la responsabilidad entra en escena como un principio ético que debería ser ineludible en estos tiempos en los que la acción tecnológica ha demostrado tener un gran poder generador de muchas problemáticas. Así, la formulación del concepto de inteligencia artificial responsable (IAR) se exhibe como un deseo de fundamentar desde la filosofía el campo tecnológico, para de ese modo enriquecer la tarea reflexiva y dotar a las acciones de un mayor compromiso con los desafíos del futuro.

Como señala Parselis, el fenómeno tecnológico no puede abordarse con la exclusividad de una disciplina particular, sino que debe ser pensado escapando de los discursos totalizadores (2018: 13), motivo por el que se requiere un marco deliberativo amplio que sirva para construir puentes de conocimiento. El modelo de innovación abierta y responsable (MIAR) proporciona esos puentes de diálogo entre diversas perspectivas e invita a pensar desde una nueva cultura de generación de conocimiento que sea consciente de la necesidad de apertura y responsabilidad. Además, la sociedad civil y la participación política simbolizan dos esferas desde las que iniciar un ciclo de generación de conocimiento mucho más amplio, enriquecedor y sensible con multiplicidad de entornos.

En torno al principio ético de la responsabilidad serán presentadas algunas propuestas valiosas e interesantes; sin embargo, estas propuestas presentan una carencia fundamentadora en lo que a filosofía se refiere. Es importante reconocer el valor de estos planteos, aunque presentan debilidad; por ello se torna necesario plantear un concepto de IAR formulado desde el MIAR, donde se pueden afrontar los desafíos tecnológicos del futuro tomando como un punto de referencia insalvable el compromiso con los derechos humanos (DDHH) y los objetivos de desarrollo sostenible (ODS). Una IA al servicio de fortalecimiento de las habilidades cívicas y la democracia, preocupada con el bienestar de la humanidad y el respeto a la biosfera. La responsabilidad es planteada en el terreno tecnológico desde una mirada de preocupación política e impulsada desde la filosofía como una matriz de conocimiento desde la que construir puentes de diálogo entre el universo de disciplinas involucradas

en la acción tecnológica, y con el deseo que enfrentar los desafíos con altura de miras y compromiso con la humanidad y la democracia.

1. La tecnología y la sugerencia de responsabilidad

En los últimos años los importantes avances que están siendo experimentados en el campo de la IA son presentados por los medios de comunicación como un fenómeno de esperanza para la resolución de muchas problemáticas que enfrenta el ser humano. Los medios contribuyen a la generación de un sentimiento de confianza en la tecnología que puede acarrear un disimulo de ciertas amenazas y una dificultad para emprender una vía crítica, que no tecnófoba. En cambio, la historia reciente pone de relieve que nuestra acción técnica no tiene exclusivamente impactos positivos, sino también negativos, como demuestra el cambio climático y la tecnología empleada para la guerra (Benjamin, 2014; Jordán Enamorado y Baques Quesada, 2014). Por lo tanto, este gran conocimiento que la humanidad ha creado, la tecnología, es decir, la técnica basada en la ciencia pasa a tener implicaciones que comprometen la vida y el futuro, y exigen ser enfrentadas desde una actitud crítica y responsable con el fin de que el sueño de la máquina inteligente no se convierta en una pesadilla que escape a nuestras manos.

256

Radnitzky define el cientificismo como “la creencia dogmática de que el modo de conocer llamado ciencias es el único que merece el título de conocimiento; y su forma vulgar, la creencia de que la ciencia resolverá eventualmente todos nuestros problemas o, cuando menos, todos nuestros problemas más significativos” (1973: 254-255). Kurzweil (2013 y 2017), personalidad del mundo de la IA, se sitúa en la estela del tecnicismo, una forma que adopta el cientificismo en el campo técnico, por entender que la tecnología en el tiempo de la singularidad será capaz de encontrar remedio a muchos de los males que enfrenta la humanidad, como las enfermedades o el cambio climático. Este presupuesto dogmático comienza a gestarse a partir de la Modernidad y se consolida en la Ilustración, dando lugar a la idea del progreso ilimitado en el conocimiento y el dominio sobre la naturaleza. En ese sentido, la mejor actitud, y más optimista, consiste en la no interferencia en los asuntos técnicos, pues ya se encargarán de corregir los errores y solucionar los males los expertos en el campo. Los expertos, en este caso tecnólogos y científicos, deberían desarrollar su actividad con total libertad y sin limitaciones, ya que se entiende que son los propios conocimientos de sus campos de estudio los que ya les presentan limitaciones.

Además de la versión optimista, existe una versión pesimista. Tras la degradación medioambiental, la sustitución de trabajadores por máquinas, el consumismo tecnológico y el gran poder destructor de la tecnología de guerra, hay quienes han elaborado un discurso muy crítico que pone en tela de juicio el desarrollo tecnológico. Entre esos pensadores nos podemos encontrar a Martin Heidegger (1997), Lewis Mumford (1998 y 2017) y Jacques Ellul (2003), quienes hacen una crítica sobre el impacto que ha tenido la tecnología para el ser humano y cómo se han transformado negativamente los valores de la humanidad. Atribuyen muchos de los males de la humanidad al hecho de que la tecnología haya fagocitado los sistemas humanos, lo que ha venido provocando una falta de control.

El pesimismo tecnológico suscita una actitud pasiva y desesperanzadora, algo que es contraproducente para los retos que depara el futuro. La estrategia del desaliento y el desánimo conduce a un callejón sin salida en la búsqueda de soluciones a las problemáticas. No obstante, el optimismo exacerbado tampoco es una solución, pues la esperanza ciega en el progreso tecnológico tampoco ha traído buenos resultados. La actitud que se origina en el dogmatismo cientificista no parece representar una buena acción a la hora de llevar a cabo una operación crítica. En ese sentido, el optimismo exacerbado representa un envite demasiado arriesgado y el pesimismo es demasiado inmovilizador. Elegir la estrategia que deja en manos de científicos y tecnólogos la solución a los problemas sería un craso error, ya que los desafíos de este tiempo exigen nuevas dinámicas de búsqueda de soluciones que involucren a más agentes como el modelo de innovación abierta y responsable (MIAR), que será discutido más adelante. Diéguez también se sitúa en esta línea:

“No es recomendable, pues, dejar por entero en manos de científicos y técnicos la solución de los problemas mencionados ni desesperar de toda solución. Por el contrario, deberíamos buscar lo que todos los miembros de la sociedad podemos hacer y hacerlo con urgencia. De hecho, hay razones para la esperanza (...) Ahora bien, sea lo que sea lo que podamos hacer y sean lo profundos que hayan de ser los cambios a realizar, éstos no deberían venir impuestos desde arriba por una estructura de poder central y autoritaria. Aun cuando resulte mucho más complicado, para ser efectivos y duraderos, los cambios deberían ser establecidos democráticamente” (1993: 192-193).

257

En medio del huracán del progreso tecnológico se disimulan ciertas exigencias bajo el valor de lo novedoso. La novedad cuenta con ventaja, y más aún si está fundamentada en la ciencia. Tanto es así que cualquier exigencia de crítica suele permanecer ausente. La apertura de posibilidades y nuevos horizontes que nos brinda el progreso tecnológico es lo exclusivamente importante; al margen quedan los cuestionamientos, las críticas y las dudas. Esto provoca que la tarea de búsqueda de responsabilidades sea muy compleja, ya que los supuestos responsables de una cadena de decisiones siempre dirigen la carga a otra instancia, ya sea superior o inferior. El progreso tecnológico nos ha llevado a un escenario en el que la exigencia de responsabilidades se diluye en medio del huracán, como sostiene Diéguez (1993: 194). A su vez, este pensador señala cuatro posibles causas a la hora de identificar con dificultad la responsabilidad en el ámbito tecnológico (1993: 194-195).

Los sistemas técnicos están constituidos por un complejo de redes en el que la toma de decisiones se somete a infinidad de procesos donde en ocasiones es difícil aplicar el principio de causalidad. La toma de decisiones se encuentra incrustada en procesos que dificultan la exigencia de responsabilidad a las partes. En ese sentido, las acciones llevadas a cabo son fruto del trabajo de redes de grupos de trabajo colectivo que cada vez más van ampliando su espectro; por ejemplo, grupos de investigación, empresa financiadora, departamento de investigación de una institución académica y organismo de gobierno, pueden ser algunas partes de la colección que intervengan en el diseño y puesta en marcha de un determinado proyecto. Esta ampliación del

campo de trabajo obstaculiza la exigencia de responsabilidad, pues a veces es difícil identificar en qué grupo recae la carga.

En último lugar, otra de las causas que impide el reconocimiento de responsabilidad en el campo de la tecnología es la tecnocracia, es decir, ese momento en el que los fines se imponen por la propia técnica y cualquier cuestionamiento por el fundamento pierde todo sentido y razón de ser. El sistema técnico y sus tentáculos son el medio de despliegue de la vida, y las otras esferas, como la política, ejercen una tarea auxiliar de mera gestión. Siguiendo a Diéguez, puede entenderse que esta tecnocracia se levanta sobre el soporte de un presupuesto dogmático inherente en el campo tecnológico de nuestro tiempo que viene gestándose durante toda la Modernidad —a saber, el cientificismo— y que ya fue definido anteriormente.

Estos son los puntos que Diéguez identifica como principales causas de dificultad para introducir parámetros de responsabilidad en el campo tecnológico. Sin embargo, existe otra causa. La racionalidad instrumental representa un modo de pensar que disimula y oculta los presupuestos dogmáticos de la tecnología y eso obstaculiza la asunción de responsabilidad. Horkheimer y Adorno exponen su crítica sobre esta racionalidad en la Modernidad en *Dialéctica de la Ilustración* (2016), y posteriormente Horkheimer en *Crítica de la razón instrumental* (2010).

Más allá de los obstáculos que recientemente han sido expuestos para incorporar criterios de responsabilidad en el ámbito tecnológico, existen motivos más que razonables para pensar que, frente a los desafíos a los que conduce la tecnología, no se debería permanecer en la indiferencia. Existe un imperativo ético que empuja a la transformación del obrar humano para asumir responsabilidad. Los desafíos exigen una reorientación de la actividad tecnológica hacia fines que sean estrictamente éticos y armoniosos con el enriquecimiento de la condición humana y el respeto a la biosfera.

258

2. La tecnología y su responsabilidad democrática

La IA puede ser impulsada para diversos propósitos, pues su dominio de acción es muy amplio. Existe un abanico de posibilidades dentro del campo de la IA y, por tanto, infinidad de usos y finalidades. Cuando las tecnologías son cooptadas por ciertos actores que promueven acciones incompatibles con la democracia y los DDHH, es fundamental la reflexión acerca de las implicaciones que pueden tener determinados mecanismos, pensando en la necesidad de asumir responsabilidad y por lo tanto en la incorporación de principios éticos que sirvan para orientar la acción.

El campo de la IA ha despertado un gran interés por parte de algunos expertos en el campo de la neuroética como Goering y Yuste (2016); de las tecnologías cognitivas (TC), como Dascal y Dror (2005) o Walker y Herrmann (2004); y también de la ética de la computación y la información como Luciano Floridi (2001; 2002a; 2002b; 2003; 2005; 2006a; 2006b; 2007a; 2007b). Estos pensadores, junto a otros, reconocen que, cuanto más se amplía el poder de estas tecnologías, mayores son las implicaciones éticas que existen en el terreno social y cognitivo. El desafío de las tecnologías avanzadas reside en que sepan desplegarse de forma compatible y

respetuosa con los sistemas democráticos y todos los principios éticos que en ellos se encuentran integrados. En ese sentido, es oportuno observar el fenómeno de la IA desde una óptica democrática, con el objetivo de poner de relieve qué controversias éticas existen y qué premisas deben ser tenidas en cuenta para el planteamiento de una IAR.

La IA presenta algunas implicaciones que generan cierta controversia como la singularidad, el mejoramiento de la especie, la automatización del campo laboral, el uso político y militar, etc. Esta controversia surge al reconocer el doble uso que puede hacerse de la tecnología, diferente al propósito para el que fue creada. Forge aborda esta cuestión y señala que la determinación del doble uso de un artefacto depende del contexto en el que está inmerso:

“La clasificación de algo como de uso dual no debe consistir simplemente en que el elemento podría tener un mal uso. Si esto fuera así, entonces la categoría de doble uso sería increíblemente grande y los objetivos finales de control y regulación inalcanzables. La pregunta se centra en cómo definir el uso dual para restringir su membresía. Esto debe hacerse teniendo en cuenta la historia o los factores contextuales, el tiempo y el lugar” (2010: 116).

Cuando hay sistemas de IA que involucran propósitos para los que no fueron concebidos nos situamos frente a un desafío ético, ya que estamos hablando de fines éticamente cuestionables. Este cuestionamiento se plantea al considerar que la IA utilizada para un uso diferente del que fue concebido, puede presentar riesgos que tienen impactos de diversa índole para los que no existen protocolos. No obstante, hay que reconocer que existen múltiples ejemplos que demuestran que algunas tecnologías que fueron promovidas para uso militar han sido empleadas con éxito en el ámbito civil.

259

Es más que evidente que el progreso de la IA está provocando una revolución política que tendrá un fuerte impacto en el seno de las sociedades democráticas a corto y medio plazo. Por ello es necesario contextualizar la democracia en este nuevo escenario incorporando la tecnología avanzada como un medio de fortalecimiento democrático. En 2006 se organizó un taller en la Universidad Estatal de Arizona, con el propósito de abordar la capacidad de cambio sociocultural que tiene la tecnología más avanzada impulsada desde la IA. En dicho taller se reconoció que la IA tiene un gran potencial de influencia sobre la inteligencia y las capacidades cognitivas del ser humano, potencial que podría tener efectos desestabilizadores. El resultado de ese encuentro fue la publicación de un documento por parte de los Sandia National Laboratories, titulado *Policy Implications of Technologies for Cognitive Enhancement* (Sarewitz y Karas). En ese documento se identifican cuatro perspectivas sobre las tecnologías del mejoramiento y su impacto político: *laissez-faire*, optimismo tecnológico, escepticismo tecnológico y esencialismo humano.

- *Laissez-faire*: esta perspectiva se sostiene sobre la defensa de la libertad individual, alejada de los controles gubernamentales y de cualquier otra instancia que pudiera ejercer control. La libertad tecnológica es producto de la expresión individual. El mercado es la única entidad capaz de regular las expresiones tecnológicas por medio de la competencia. Los gobiernos tienen la responsabilidad de promover la libertad creativa para la producción tecnológica.
- Optimismo tecnológico: esta perspectiva se basa en el ideal ilustrado de la innovación técnica y científica como clave del progreso humano. Si bien esta visión es consciente del gran aporte de la tecnología, también lo es de las problemáticas que se presentarían. Los entes gubernamentales tienen la responsabilidad de asegurar las condiciones para la innovación tecnológica, aunque es en el sector privado donde recae la actividad productiva. Además, el gobierno debe favorecer un discurso abierto e inclusivo, especialmente destinado a aquellos sectores más desfavorecidos tecnológicamente hablando.
- Escepticismo tecnológico: desde esta perspectiva se reconoce la aportación que la tecnología tiene para la sociedad y la democracia, pero se duda de que la tecnología sea inherentemente beneficiosa. En ese sentido, se afirma que los desequilibrios generados no pueden ser solucionados por la propia lógica tecnológica, sino a través de la acción política. Los escépticos se sitúan del lado de la regulación y la moderación del poder para minimizar los riesgos. Esta es una visión regulacionista que entiende que las entidades de control gubernamental deben proporcionar los mecanismos políticos necesarios para estudiar el impacto y la investigación en el campo tecnológico.
- Esencialismo humano: esta es otra de las perspectivas que tratan de valorar el fenómeno de las TC. Esta perspectiva se sostiene sobre los conceptos de dignidad y esencia humana. La interferencia de las tecnologías de mejora sobre los seres humanos representaría una amenaza para ambos conceptos. El ser humano está limitado por su naturaleza y no es aceptado ningún intento de mejora procedente de la tecnología, salvo en el caso de los discapacitados, que será regulada. El contexto de mejora de los individuos tiene lugar en la familia y la comunidad. Dios ha proporcionado unos atributos naturales que no deben ser modificados.

Más allá de estas cuatro posiciones, podría existir una quinta vía para dar forma al proyecto de una IAR dentro del contexto de la democracia. Entre las cuatro perspectivas, la que podría provocar una mayor aceptación es el escepticismo tecnológico, sobre todo por hacer hincapié en la necesidad de las regulaciones gubernamentales; sin embargo, una excesiva regulación podría erosionar los principios democráticos más básicos. En ese sentido, es necesaria una quinta perspectiva que encuentre su fundamento principalmente en el humanismo tecnológico, en el principio de responsabilidad y en otros postulados filosóficos que sirvan de guía para la política democrática en este tiempo tecnológico.

Ya ha sido mencionada la necesidad de contextualizar la tecnología en un escenario democrático y para ello deben tenerse en cuenta nuevos aspectos. A colación de dicha contextualización, Quintanilla diferencia entre una democracia tecnológica mínima y una democracia tecnológica plena, dentro de las cuales identifica varios

puntos cardinales (2002: 640-647). En el interior de la democracia tecnológica mínima pueden ser identificados los siguientes puntos:

- Derecho de todos los ciudadanos a participar en las decisiones sobre el uso de las posibilidades tecnológicas en asuntos de interés público.
- Derecho de todos los ciudadanos de acceder al conocimiento técnico y de contar con el juicio de los expertos como elemento fundamental para conformar la opinión pública y participar en las decisiones políticas sobre asuntos técnicamente complejos.
- No convertir en problemas políticos aquellos asuntos para los que existen soluciones técnicas solventes y contrastadas.

En cuanto a la democracia tecnológica plena destaca principalmente el siguiente aspecto: el derecho de todos los ciudadanos a acceder a todo el conocimiento tecnológico relevante para la toma de decisiones en asuntos de interés público y a participar en el diseño, evaluación y control del desarrollo tecnológico.

Las posibilidades que plantea el desarrollo de la IA nos sitúan en un escenario político que demanda un nuevo contrato social en el que la tecnología promueva el bienestar social y se generen nuevas dinámicas participativas en el seno de las sociedades. La participación de los sectores implicados es presentada como la punta de lanza en este nuevo escenario, donde lo político es entendido como el espacio de participación de la toma de decisiones más cruciales sobre el desarrollo tecnocientífico. El modelo basado en la experticia, donde la toma de decisiones reside exclusivamente en los científicos, tecnólogos, y en los monopolios tecnológicos, debería pasar a la historia, pues no se entiende un futuro de profundo desarrollo de la IA con importantes implicaciones sin un carácter cívico y democrático. La acción democrática tendrá que ser fruto de un trabajo participativo y de una orientación basada en principios éticos que sean fundamentados en la contextualización del nuevo escenario tecnológico.

261

3. Inteligencia artificial responsable

Es bastante la bibliografía sobre las implicaciones éticas de la IA, pero casi todo lo escrito se ha centrado en casos particulares como la robótica, por ejemplo en su impacto en el ámbito profesional, la biotecnología en el ámbito sanitario o el de los drones en el terreno militar. Aunque estas aportaciones centradas en aspectos particulares son muy valiosas, es necesario aportar un marco de reflexión mucho más amplio y sistemático. Esto podría abrir un nuevo horizonte de posibilidades para pensar la IA, las oportunidades que de ella se derivan y las consecuencias. En ese sentido, la propuesta de una IAR serviría para sentar las bases de un mayor conocimiento de las necesidades de la sociedad y también para una mayor sensatez a la hora de considerar los impactos tecnológicos.

En torno al criterio de responsabilidad, existen dos planteamientos representados por Dignum (2013, 2017a, 2017b y 2018) y la Declaración de Montreal (2018), que

significan un primer paso desde el que plantear la necesidad de responsabilidad en el contexto de la IA, aunque carecen de una solidez filosófica en el plano de la fundamentación y también de un marco deliberativo concreto. Además, tampoco hacen el suficiente hincapié en cuestiones de suma relevancia en la actualidad como los DDHH o los ODS. Por lo tanto, en este apartado se expondrán los principales pilares de fundamentación desde el que la IAR será planteada y también cómo el MIAR proporciona un soporte metodológico desde un espacio deliberativo en el que se pueden generar nuevas dinámicas para la producción de conocimiento. El MIAR es el soporte fundamental desde el que plantear una IAR fuerte, alejada de la debilidad y la superficialidad de otros planteamientos, que sepa asumir la complejidad de los retos de nuestro tiempo reconociendo la importancia de la participación de la sociedad civil y del ejercicio deliberativo y fronético. La IAR es presentada como una alternativa de fortaleza frente a aquellos planteamientos que destacan por su carencia filosófica y su superficialidad.

3.1. Modelo de innovación abierta y responsable

El proceso de innovación convencional de carácter cerrado ha sido durante mucho tiempo el enfoque predominante y más exitoso dentro de las empresas, como señala Nedon (2015: 7). Sin embargo, este modelo de innovación estrictamente cerrado y de triple hélice ha significado con el paso del tiempo un aislamiento respecto a otros conocimientos que van generándose en el mundo. El modelo de innovación de triple hélice se erige sobre la idea de que exclusivamente un grupo es capaz de poseer el monopolio de conocimiento en un determinado campo, algo que es muy difícil de sostener en un mundo donde todo el conocimiento está conectado por redes de diversa índole. Este aislamiento sostenido sobre la idea monopolística ejerce resistencia frente a otros conocimientos y frente a una realidad que requiere continua adaptación. Resulta paradójico creer que los miembros de un grupo de trabajo puedan concentrar todas las ideas posibles y que no necesiten otras fuentes de conocimiento para la innovación. La globalización y el desarrollo de las tecnologías han favorecido el intercambio de conocimiento y la movilidad, creando de ese modo puentes de intercambio de ideas entre diferentes esferas del saber. Urra Canales (2018: 184) resume con claridad las posturas críticas que giran en torno al modelo de innovación de triple hélice:

- Defensa romántica de la autonomía académica y demonización de las empresas
- Minusvaloración de la cooperación internacional y transdisciplinar
- Ciencia básica amenazada por la ciencia aplicada
- Escrúpulos metodológicos
- Desdén hacia los problemas sociales, las ONG, la sociedad civil y los países en vías de desarrollo y los problemas sociales
- Incredulidad en las capacidades de hibridación de las esferas académicas, gubernamental y empresarial

El modelo de innovación abierta, tratado principalmente entre 2011 y 2016 en La Haya a partir de la iniciativa recogida en el Programa de Investigación Maatschappelijk

Verantwoord Innoveren del Consejo de Investigación holandés, comienza a jugar un papel cada vez más relevante en numerosos ámbitos. Representa un nuevo alcance de la ética aplicada que pasa por reconocer la importancia práctica de la filosofía moral en diversidad de esferas. En el contexto científico y tecnológico, la ética comenzó a preguntarse sobre la posibilidad de contribuir a la solución de los problemas y la mejora de las condiciones de vida de la sociedad. Además, la participación de los afectados por la actividad tecnológica se convierte en una exigencia muy relevante. Más allá de los muros de los laboratorios y las academias se encuentran infinidad de ciudadanos que reivindican una participación en aquellos proyectos que comprometen sus vidas. A partir de esta consideración comienza a concebirse la tecnología como un sistema sociotécnico que se encuentra insertado en contextos sociales.

Dentro del MIAR la innovación social juega un papel se suma relevancia. La innovación nutre al modelo conocimientos enriquecedores para dar respuesta a las demandas sociales. En este sentido, la Comisión Europea afirma lo siguiente:

“La innovación no es sólo un mecanismo económico o un proceso técnico. Es sobre todo un fenómeno social. A través de ella, los individuos y las sociedades expresan su creatividad, necesidades y deseos (...) La innovación puede y debe ofrecer una respuesta a los problemas cruciales de la actualidad. Esto hace posible una mejora en las condiciones de vida (los nuevos medios de diagnóstico y tratamiento de las enfermedades, la seguridad en el transporte, más fáciles de comunicación, un medio ambiente más limpio, etc.” (Comisión Europea, 1995: 11).

263

Esta visión de la innovación social introduce en el MIAR una nueva concepción de la innovación tradicional en la empresa y en la economía. En lo que respecta a sus procesos y productos, no se limita al emprendimiento social, sino que va más allá de él, no limita su actividad a la aplicación de la tecnología en los ámbitos de exclusión social y tampoco simplifica su actividad reduciéndola a una simple metodología de participación y creatividad. Además le da mucha importancia a la responsabilidad social, promueve la disolución de la fronteras entre el diálogo y la cooperación que subyacen en las relaciones entre los sectores público, privado y otras organizaciones sin fines de lucro, involucra a los afectos y beneficiarios, nutriéndose de ese modo de sus experiencias, se desarrolla bajo un enfoque integral y holista, incorporando diversas problemáticas de la complejidad social e implica el fortalecimiento de las alianzas entre las esferas que se encuentran presentes en la cuádruple hélice, entre otros aspectos importantes (Bureau of European Policy Advisers, 2011; Comisión Europea, 2010; Morales, 2008, 2009a y 2009b).

Este modelo también se caracteriza por la introducción de la alerta temprana como condición esencial, lo que Grunwald (2014: 16) denomina como “compromiso temprano”. Esta alerta se presenta como sinónimo de un compromiso anticipado y prudente en el despliegue histórico frente a los impactos futuribles. Es algo similar a lo que Jonas mencionaba cuando abordaba el tema de la heurística del temor de la siguiente manera: “Resulta, pues, necesario elaborar una ciencia de la predicción

hipotética, una ‘futurología comparada’” (1995: 64). Ejemplos como la investigación que se refleja en *The Malicious Use of Artificial Intelligence; Forecasting, Prevention, and Mitigation* son un claro ejemplo de la necesidad de llevar a cabo reflexiones que planteen escenarios futuros y previsibles.

En el ámbito de la IA el modelo de innovación es presentado como una novedad en lo que respecta a la incorporación de criterios de responsabilidad. Puede aportar los mecanismos suficientes para ir configurando una IAR. Sirve para orientar los diagnósticos y los planes de acción de forma responsable. Además, el trabajo participativo promueve la aportación de diversas perspectivas que proceden de esferas diferenciadas y a la vez permite poner atención a nuevos factores y variables que normalmente no serían tenidos en cuenta desde una única perspectiva. El modelo es planteado desde un encuentro de perspectivas plurales.

La responsabilidad que motiva el MIAR nace de la toma de conciencia frente a la concepción que predominó hasta la década de los noventa y que consideraba la tecnología desde una neutralidad axiológica. No existe ninguna tecnología exenta de consideración moral, ajena el mundo, pues recordemos que tiene un carácter sociotécnico. En ese sentido, para Grunwald (2014: 23) la responsabilidad tiene tres dimensiones en el contexto tecnológico:

- *Dimensión sociopolítica*: el impacto está dirigido a la sociedad y su carácter político.
- *Dimensión moral*: los criterios y códigos morales forman parte del marco normativo que sirve para juzgar las acciones tecnológicas.
- *Dimensión epistémica*: centra su atención en la calidad de los conocimientos de los que se dispone para la evaluación.

264

3.2. Deliberación, participación y sociedad civil

No existe una teoría ética perfecta y absoluta, por eso es fundamental el ejercicio deliberativo. La deliberación se convierte en una importante herramienta para la construcción de un proyecto de ética aplicada a la IA de forma participativa, donde tengan cabida todos aquellos sectores que estén afectados por la actividad: políticos, académicos, psicólogos, tecnólogos, etc. Una ética de este tipo nace de la posibilidad de reconocer perspectivas, pues el ámbito de la ética es contingente, y como señala Moratalla: “La deliberación es el método de análisis ético una vez que reconocemos la contingencia de los asuntos humanos y comprendemos que nuestro ámbito de acción y decisión es el de la incertidumbre y la complejidad” (2017: 41).

La actitud deliberativa implica la necesidad de alejarse de determinados “vicios” propios del modo de pensar lineal deductivo, donde todo se reduce a aspectos cuantificables y dilemas. Según Gracia, si queremos emprender la vía deliberativa, es necesario huir de los dilemas, pues “más que dilemas, hay problemas, es decir casos con múltiples cursos de acción posibles que será preciso tener en cuenta a la hora de tomar una decisión razonable o prudente” (2016: 13). Esto quiere decir que el reconocimiento de los problemas permite conocer que existen diversidad de

perspectivas desde las que afrontar un caso. Sin embargo, los dilemas responden a una lógica binaria donde todo está bien o mal, es verdadero o falso, etc. En el proceso de deliberación no pueden convertirse los problemas en dilemas, pues el horizonte de posibilidades de discusión se reduce considerablemente y eso conduce a un empobrecimiento de la posibilidad de enriquecer una ética aplicada a la IA. No obstante, la tarea deliberativa no es fácil, puesto que surge desde una formación vinculada al respeto activo, tal como sostiene Cortina, y que consiste “no solo en soportar estoicamente que otros piensen de forma distinta, tengan ideales de vida feliz diferentes a los míos, sino en el interés positivo por comprender sus proyectos, por ayudarles a llevarlos adelante, siempre que representen un punto de vista moral respetable” (1998: 240).

Es importante destacar en esta parte que la idea que actualmente tenemos de la deliberación ha sido producto de la confluencia de diversas tradiciones que han ido desarrollándose a lo largo de la historia y de las que Vega Reñón (2016) destaca tres: en primer lugar, su origen antiguo; en segundo lugar, la contribución de la Modernidad; y en tercer lugar, el momento actual en el que la reflexión deliberativa se da dentro de un marco socio-institucional del discurso público. Puesto que el conocimiento científico es considerado en este trabajo como un recurso público para la potenciación de las habilidades cívicas y las prácticas democráticas, nos centraremos en el tercer momento por hacer hincapié en el aspecto público.

Walton (2004 y 2006) es mencionado por Vega Reñón (2016: 219-220) como un claro ejemplo de caracterización de la deliberación como un tipo de diálogo y esquema argumentativo que tiene como objetivo la resolución de problemas prácticas a partir de la inferencia práctica medios-fin o actuación-riegos/consecuencias. Esta propuesta tiene un profundo carácter reflexivo al partir de la preocupación por los planes de acción, los objetivos y su pertinencia, las consecuencias y riesgos, la plausibilidad, etc. Esta propuesta considera de suma importancia la labor crítica de la deliberación prudencial. Pero Vega Reñón va más allá y se sitúa en el terreno de aquellos problemas que suscitan una dimensión pública y un alcance colectivo. En este sentido, el MIAR proporciona a la IAR un punto de partida desde el que considerar al conocimiento científico como un recurso público y por lo tanto preocupado por las cuestiones de alcance colectivo, las habilidades cívicas y el enriquecimiento de la democracia en términos generales. Vega Reñón caracteriza su propuesta de la siguiente manera:

“(i) el reconocimiento de una cuestión de interés y de dominio públicos, donde lo público se opone a lo privado y a lo privativo; (ii) el empleo sustancial de propuestas; (iii) las estimaciones y preferencias fundadas en razones pluridimensionales que remiten a consideraciones plausibles, criterios de ponderación y supuestos de congruencia práctica; (iv) el propósito de inducir al logro consensuado y razonablemente motivado de resultados de interés general -no siempre conseguido” (2016: 220).

La superación del modelo de innovación cerrada o de triple hélice, como sostiene Urra Canales (2017: 193), ha suscitado en la sociedad del conocimiento la aparición

de una nueva esfera más allá del Estado, el mercado y la universidad, a saber: la comunidad, que es presentada como un conjunto de personas que comparten un espacio geográfico, intereses y características (2017: 194). El ser humano es un ser social y vive en comunidad desde su lejano origen. El concepto de comunidad ha sido objeto de sendas críticas a partir de aspectos que tienen que ver con el autoritarismo o el conservadurismo, entre otros. No obstante, estas críticas han ido perdiendo su valor cuando el concepto de comunidad se ha relacionado con el de sociedad civil. El término “sociedad civil” encuentra su origen en el latín en *societas civilis* y hace referencia a un conjunto de ciudadanos que son miembros de una comunidad y que poseen ciertos derechos que les permiten participar en la vida pública. El tema de la sociedad civil y la comunidad ha sido objeto de reflexión en la filosofía a lo largo de su historia, comenzando con la Política de Aristóteles, los contractualistas clásicos Hobbes, Locke, Rousseau, y otros como Kant, Heguel, Marx o Gramsci. Salvando las diferencias que existen entre los diversos postulados de las figuras que han sido mencionadas, y también otras, existe un punto de encuentro que permite dar cuenta de que el ser humano vive en comunidad por diversos motivos.

El contexto social de desarrollo de la IA es el de la sociedad del conocimiento, que es también el escenario en el que se despliega en ejercicio la sociedad civil. En *Developing democracy. Towards consolidation*, Larry Diamond define a la sociedad civil a partir de cinco aspectos (1999: 223-227):

266

1. La sociedad civil centra sus esfuerzos en los fines públicos y no en aquellos que son estrictamente privados.
2. La sociedad civil establece una relación con el Estado, pero no pretende reemplazarlo en materia de control, no buscando de esa manera gobernar por sí misma, sino más bien lograr influencia en los poderes políticos institucionales.
3. La sociedad civil se construye a partir del respeto a la pluralidad y la diversidad. La sociedad civil no promueve una perspectiva holística, sino que incorpora los intereses que emanan de los diversos grupos, no representando ningún interés en concreto, ya se deriven de una persona o una comunidad.
4. El concepto de sociedad civil se diferencia del de “sociedad cívica”, ya que el segundo hace referencia a la cooperación y la reciprocidad voluntaria entre individuos sin la necesidad de trascender socialmente.

Anteriormente se mencionó la importancia de que el proyecto de una IAR sirviera como punta de lanza para promover los ODS, por lo que es fundamental llevar a cabo una mención especial a la participación ciudadana como un mecanismo desde el que reflexionar acerca de la naturaleza de la democracia y las habilidades cívicas. La participación adquiere un importante valor dentro del contexto del MIAR, pues es presentada como una posibilidad transformadora desde la que abordar las problemáticas y gestionar los asuntos públicos de un modo innovador.

Gaventa (2006) plantea seis desafíos que sirven para valorar la importancia que tiene la inclusión de la ciudadanía en los asuntos públicos y sus posibilidades transformadoras. Dentro de la propuesta que aquí se está esbozando sobre el valor

que la IAR puede tener para el enriquecimiento del civismo y la democracia, se destacan los siguientes:

- Relacionar a la gente con las instituciones
- Repensar las relaciones entre la sociedad civil y las instituciones políticas
- Reflexionar sobre el valor de la participación ciudadana
- Contar con mayor conocimiento para entender las relaciones de poder

Siguiendo a García Inda (2003), si en la teoría política contemporánea ha sido el conjunto de derechos y deberes lo que ha configurado los límites del ciudadano, la paulatina degradación del sistema de democracia representativa liberal y la creciente conflictividad social, incapaz de satisfacer y hacer efectivos el sistema de derechos, ha empujado a los científicos sociales a reformular el sentido de ciudadanía. En esta línea es la participación del sujeto social y político lo que vendría a profundizar y a hacer efectivo el conjunto de derechos y deberes ciudadanos. Se trata de participar en la construcción de las reglas de juego (participación política), así como en la producción, la distribución y el control de los bienes de la comunidad política, económicos, sociales, políticos y culturales. Es decir, el sujeto ciudadano-participativo pasaría de ser un mero titular pasivo de enunciados jurídicos con más o menos posibilidades de materialización, a construir y ampliar activamente su implementación. Se superan de esta manera los tradicionales derechos civiles que otorgan, en una democracia representativa, el estatus de ciudadano identificado con los derechos de asociación y derecho de representación activa y pasiva a través del voto.

267

La participación ciudadana dibuja un nuevo sujeto político, el ciudadano, que participa activamente, que decide, que adquiere compromiso y responsabilidad. Este sujeto político se dibuja en un mundo globalizado de la comunicación tecnológica, de las crisis económicas y ambientales, de la tecnocracia, de las corrupciones favorecidas por la opacidad del sistema político, en medio de la deslegitimación del sistema político de la democracia representativa liberal, de las convulsiones y los movimientos sociales.

3.3. Compromiso con los derechos humanos y los objetivos de desarrollo sostenible

En la era de la información y las comunicaciones, de los retos que representan los ODS, entre otros desafíos, la humanidad se enfrenta a un importante escenario que puede ser abordado desde los mecanismos más avanzados que nos ofrece la tecnología. En la línea del planteamiento de la IAR que se está llevando a cabo en estas páginas, es fundamental tener en cuenta dos espacios desde los que diseñar la tecnología más avanzada, a saber: los DDHH y los ODS. En ese sentido, se llevará a cabo un acercamiento a estos dos espacios que en ciertos momentos presentaran elementos comunes como el compromiso por la igualdad de género o la reducción de la pobreza.

El profundo avance de la IA y su impacto en la vida también tiene una importante implicación en materia de DDHH. Access Now (2018) ha realizado un estudio preliminar en el que analiza la gama de problemas en materia de DDHH que suscita la introducción disruptiva de la IA en nuestras vidas. Muchos de los problemas en esta materia no son nuevos; sin embargo, se someten a una profundización o adquieren nuevas formas debido a la influencia de la IA. A diferencia de las tecnologías tradicionales, los sistemas artificiales nos sitúan en nuevo escenario debido a su potencial. Así pues, el derecho internacional y las instituciones dedicadas a los DDHH pueden servir como un impulso organizado desde el que abordar el fenómeno de la IA desde una óptica de responsabilidad y compromiso.

3.3.1. *Respeto a los derechos humanos*

Cada vez son más las empresas, gobiernos y otras instituciones las que ven la IA como una herramienta para fortalecer muchos campos de desempeño humano. Sin embargo, este importante aprovechamiento de la tecnología está despertando la legítima inquietud en grupos de la sociedad civil y en otros actores políticos que plantean la necesidad de formular preguntas sobre las implicaciones que la IA tiene para los DDHH en materia de sesgos injustos, violaciones de la privacidad y la libertad. Es importante abordar esta cuestión desde una perspectiva de responsabilidad enmarcada al interior de un MIAR para valorar dichas implicaciones, priorizando de ese modo la dignidad de los seres humanos y protegiendo los derechos más básicos, con el objetivo de generar confianza en la sociedad civil y promover un fortalecimiento de las habilidades cívicas y democráticas por medio de la tecnología más avanzada.

268

La IA actual suele fundamentar el despliegue técnico de su acción en el *machine learning* con un propósito definido. Los sistemas artificiales son sometidos a un entrenamiento con datos pertinentes para el cumplimiento de sus propósitos que llevan a cabo mediante procesamiento matemáticos para la detección de patrones en la información que es suministrada y el posterior desarrollo de un modelo que permita ejercer predicciones o recomendaciones sobre nuevos datos. Por ejemplo, en algunos departamentos de recursos humanos ya se están empleando intelectos sintéticos para la selección de candidatos profesionales a través de un entrenamiento en el que se utilizan datos existentes sobre los trabajadores.

Satya Nadella, director ejecutivo de Microsoft, escribió un artículo en el que reflexionaba sobre la alianza entre los seres humanos y la inteligencia artificial, en el que se refleja la necesidad de enriquecer el debate para que sea más productivo y orientarlo hacia aquellos valores que son inculcados a las personas y las instituciones que promueven la IA. Este texto reflexiona sobre los seis principios y objetivos que son expuestos y analizados en *The Future Computed: Artificial Intelligence and Its Role in Society* y que la industria y la sociedad deben someter a un profundo debate: equidad, fiabilidad, seguridad, privacidad y seguridad, inclusividad, transparencia y responsabilidad. Todos estos principios y objetivos son claramente pertinentes para la garantía y protección de los DDHH. El planteamiento de la IAR juega un papel fundamental en la promoción de una tecnología respetuosa y garante con los derechos de la humanidad, pues reconoce el gran potencial que la IA tiene y a la vez invita a pensar sobre la necesidad de orientar este potencial en beneficio de la humanidad.

Cuando el propósito de una IA es definido, son las instituciones humanas las que deciden qué aspectos del contexto son los más relevantes en la consideración. Recuperando el ejemplo planteado anteriormente, cuando se diseña un sistema artificial para la selección de personal de una empresa, los seres humanos dirigen la actividad de diseño de estos sistemas y determinan qué criterios incorporar; por ejemplo, pueden decidir centrarse exclusivamente en la maximización de la rentabilidad económica o bien incorporar otros aspectos más integrales y holísticos que también son generalmente importantes en el ámbito humano.

Propública, una agencia de noticias independiente de Manhattan, publicó en 2016 un informe en el que se examinaba la validez de una herramienta de evaluación de riesgos que Northpointe utilizaba para tomar decisiones sobre la libertad condicional de aproximadamente 35.000 convictos federales. Los investigadores Jennifer Skeem y Christopher T. Lowenkamp encontraron que los negros obtuvieron un puntaje promedio más alto que los blancos:

Tabla 2. Porcentajes de predicción de incidencia en la población negra y blanca

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Fuente: Skeem y Lowenkamp (2016: 7)

269

Este informe demuestra que el incidente a “pequeña” escala ocasionado por este *software* de predicción puede replicarse en el futuro y tener un impacto mucho mayor y serias implicaciones en otros ámbitos de la vida. Además, pone de relieve que aquellos sesgos sistémicos por los que las personas se han pasado décadas impulsando campañas de educación y legislación pueden caer en el olvido con motivo de un impulso tecnológico.

Otro sesgo que se encuentra presente en el ámbito de los intelectos sintéticos es el sexista. Amazon prescindió del uso de una herramienta de IA que se centraba en la contratación de personas porque discriminaba a las mujeres. La compañía de Jeff Bezos comenzó en 2014 a diseñar sistemas informáticos para la revisión de currículums de los aspirantes y el objetivo se centraba en facilitar la selección de los mejores talentos. En 2015 se observó que su nuevo sistema de selección de personal estaba llevando a cabo un ejercicio sesgado en materia de género. Esto se debe principalmente a la introducción de una serie de patrones en los sistemas artificiales que habían sido fruto de la recopilación de datos durante una década, y que la mayoría de ellos pertenecían a hombres, por lo que la IA fue evidentemente entrenada bajo un dominio masculino.

Los casos de Northpointe y Amazon ponen de relieve la necesidad de reflexionar acerca de la calidad de los datos que son utilizados en el entrenamiento para

el aprendizaje automático de los intelectos sintéticos, ya que ese es un factor fundamental y decisivo del carácter ético del modelo resultante. La propuesta de una IAR en el contexto de la discriminación no puede girar en torno a la exclusión de datos delicados como el color de la piel o el género, ya que en ocasiones son factores que condicionan una discriminación positiva que garantiza el respeto a los DDHH. Por el contrario, la IAR promueve un entrenamiento de la IA contando con todas las variables necesarias, incluso aquellas que son problemáticas, ya que es un aspecto necesario para la identificación de los problemas y su posterior resolución tras dicho entrenamiento.

En mayo de 2018 organizaciones como Amnistía Internacional, Access Now y otras organizaciones asociadas presentaron un documento bajo el título *Toronto's Declaration on Machine Learning*, un texto en el que se promueve la protección del derecho a la igualdad y la no discriminación que en ocasiones puede derivarse de la IA. Esta declaración persigue la aplicación de las normas internacionales de DDHH al contexto de la IA y las actividades que de ella se derivan por medio del aprendizaje automático. Además, dicha declaración tiene como objeto la proposición de discusiones sobre los principios y documentos donde se analizan los daños que han surgido a partir de esta tecnología. También se señala con especial interés lo siguiente: en relación a la máquina de sistemas de aprendizaje e inteligencia artificial en general, los Estados deberían promover el derecho positivo para el disfrute de los avances en ciencia y tecnología como afirmación de los derechos económicos, sociales y culturales (2018: 2). La Declaración de Toronto centra su preocupación en los algoritmos que respetan los DDHH. En ese sentido, la propuesta de una IAR contextualizada en el marco de los DDHH va destinada al diseño de sistemas artificiales respetuosos con los DDHH y por lo tanto a la elaboración de aquellos mecanismos que se consideren necesarios para el fortalecimiento de la democracia y el civismo en provecho de la humanidad.

270

3.3.2. Impulso a los objetivos de desarrollo sostenible

Como ya se ha mencionado, la IA puede servir como un motor de impulso para el cumplimiento de los ODS. La IAR podría situar como una de sus tareas centrales la promoción de la innovación en el contexto de los ODS. La aplicación de ideas innovadoras en materia tecnológica puede ayudar a los países para avanzar más rápidamente hacia el logro de estos objetivos. Los avances en el campo de la IA pueden orientarse hacia el empoderamiento de gobiernos, comunidades y organizaciones que requieren la puesta en marcha de soluciones efectivas para las problemáticas.

El Programa de las Naciones Unidas para el Desarrollo (PNUD) ha llevado a cabo un análisis sobre el valor de los sistemas artificiales para la generación de progreso y de mecanismos transformadores en materia de desarrollo:

- En el campo de la medicina, la IA ha proporcionado predicciones muy valiosas para la predicción de ataques cardíacos. En ese sentido, esta tecnología avanzada podría contribuir con los diagnósticos para salvar millones de vidas.
- En el campo agrícola, proporciona datos muy valiosos para un mayor conocimiento de los patrones meteorológicos y la producción. Entre las herramientas más usadas podemos encontrar LettuceBot, un sistema que permite la identificación y eliminación

de las malas yerbas mediante una base de datos de más de un millón de imágenes que permite identificar las plantas. Gracias a estas tecnologías, la producción puede ser más eficiente y eso podría prevenir los riesgos del sector agrario y favorecer la producción de bienes más estables para las poblaciones del ámbito rural.

- En el contexto lingüístico, los intelectos sintéticos también están contribuyendo a la traducción de infinidad de lenguas por medios de los *chatbots*, un mecanismo que, por ejemplo, permite entender las necesidades de los refugiados. Además, facilita la tarea de los abogados y funcionarios públicos en materia de traducción.

Es evidente que la IA no es la solución para todos los males; sin embargo, el planteamiento de una IAR podría contribuir en muchas de las exigencias que se encuentran presentes en los 17 ODS. El monitoreo sobre las funciones de la IA es presentado como una necesidad ineludible en el marco de las problemáticas actuales, ya que la tecnología puede proporcionar interesantes soluciones institucionales que en este momento podrían pasar desapercibidas. Además, en este contexto de los ODS es fundamental identificar de modo pertinente cuáles son los principales puntos desde los que plantear alternativas tecnológicas y posibles soluciones.

En la línea de esta visión acerca de las tecnologías facilitadoras de la consecución de los ODS se encuentra Open Data en Europa y Asia Central (ODECA), una plataforma basada en un programa de datos abiertos para el desarrollo que apoya a los representantes gubernamentales, la sociedad civil y otras organizaciones para que trabajen en el intercambio de datos para la generación de conocimientos innovadores. En su web, ODECA se define de la siguiente manera:

271

“La red cubre 18 países de la región y tienes como objetivo estimular la innovación, el intercambio de conocimientos y el aprendizaje entre profesionales y aficionados de los datos abiertos a nivel regional y mundial. Nuestro objetivo es utilizar el potencial de los datos abiertos para transformar las sociedades al empoderar a los ciudadanos y ayudar a los gobiernos a cumplir los Objetivos de Desarrollo Sostenible de la ONU. Si bien todavía estamos explorando todas las formas en que los datos contribuirían a los ODS, es innegable que jugará un papel importante para alcanzarlos y medirlos” (ODECA, s/f).

La IAR que proporciona una preocupación sistemática por el fortalecimiento de las habilidades cívicas y la democracia a partir del reconocimiento del conocimiento científico como un recurso público, se encuentra estrechamente vinculada con los ODS. Un claro ejemplo de este estrecho vínculo lo representa Citibeats, una plataforma de IA especializada en la escucha activa de la ciudadanía que ha sido creada para trabajar en el marco de los ODS. Como reza en su web, Citibeats “estructura, analiza y sintetiza las opiniones de las personas a partir de grandes cantidades de datos en cualquier región, en cualquier idioma y de cualquier fuente de datos de textos, para que sean fáciles de usar”. Esta plataforma tiene un importante compromiso en la potenciación del compromiso cívico y la mejora de la participación ciudadana. Sirve como soporte informativo para las ciudades a través del Observatorio de Impacto

Local, que permite obtener un mayor conocimiento sobre las problemáticas locales que existen entre las diversas organizaciones y los entes de gobierno. Son numerosas las ciudades que ya están trabajando con Citibeats: Barcelona, Madrid, Londres, Nueva York, Tokio y urbes de China. Además, proyectos de este tipo pueden apostar por mecanismos de transparencia y de ese modo enriquecer los sistemas democráticos

Finalmente, y recuperando el espíritu de una nueva ética que nos planteaba Jonas, alejada de un antropocentrismo exacerbado, puede sugerirse una agenda de acciones de la IAR a través de una preocupación ineludible y comprometida por el medioambiente. La importancia de la naturaleza para los seres humanos es fundamental, ya que los ecosistemas nos brindan una cantidad de bienes y servicios que determinan nuestra calidad de vida y hacen posible la producción económica y por lo tanto un desarrollo saludable a las sociedades. En materia medioambiental la IA puede llevar a cabo enriquecedoras aportaciones a partir del estudio de los comportamientos meteorológicos para la predicción de fenómenos climáticos. El monitoreo exhaustivo permite dar recomendaciones simples y muy determinantes al mismo tiempo, por ejemplo al señalar a las autoridades gubernamentales cuáles son las zonas que presentan mayores niveles de sequía. Además, en las ciudades inteligentes la IA proporciona un mejor conocimiento sobre aquellas variables que influyen en el consumo de energía como el clima, la ubicación geográfica o los eventos que suceden en los alrededores, por lo que podría contribuir considerablemente en una reducción de la energía que se utiliza en los edificios. El enriquecimiento que está brindando la asesoría de la IA en el terreno medioambiental es indudable como evidencia el proyecto Artificial Intelligence for Ecosystem Services (ARIES) financiado inicialmente por la National Science Foundation en Estados Unidos y que presenta una interesante herramienta que facilita el trabajo para el mapeo de servicios medioambientales. Mediante la utilización de una base de datos muy diversa y pertinente al ecosistema de estudio, ARIES elabora un mapa donde se presentan las interacciones entre factores económicos y ambientales. Así pues, es más fácil visualizar cuáles son los servicios ambientales, sus beneficiarios y los flujos que se dan en una zona.

Sin duda todos estos compromisos con los ODS por parte de la IAR no podrían llevarse a cabo sin una nueva concepción de la economía que incorporara criterios éticos para el ejercicio de su actividad. En relación al estrecho vínculo entre la economía y la ética de la responsabilidad en el contexto de una IAR es importante destacar las siguientes palabras de Conill:

“Algunos han creído que la ética de la responsabilidad es cosa de los políticos, dado que tal fue el contexto en el que Max Weber lo planteó en su momento, pero esto sería excesivamente unilateral. En realidad, son las condiciones propias de la vida moderna, es decir, de las sociedades crecientemente complejas y diferenciadas, las que exigen el enfoque de la responsabilidad en todos los órdenes de la vida, por parte de aquéllos que quieran que sus principios y sus convicciones sean operativos en los diversos ámbitos que configuran nuestra vida real. Por tanto, no sólo en la política (en sentido restringido), sino también en las actividades económicas se exige un ejercicio de la razón pública en forma de responsabilidad pluridimensional” (2013: 15).

Conclusiones

Escribir un trabajo en el que se intenten dar suficientes argumentos sobre la necesidad de introducir criterios éticos de responsabilidad, y por lo tanto hacerlo creando un momento de fundamentación filosófica de la tecnología, es algo sumamente difícil. Sin embargo, este artículo busca contribuir a un deseo que se viene gestando en las últimas décadas y que justificará su razón de ser en los próximos años. Los desafíos que nos depara el futuro de la IA son sumamente valiosos para enriquecer el pensamiento tecnológico por medio de la deliberación con otros campos del saber, como en este caso es la filosofía.

La generación de nuevos conocimientos que sepan afrontar los retos éticos del futuro de la IA requiere de un cambio de matriz cultural en el seno de centros de saber y poder como la academia, el Estado, el sector privado y la sociedad civil. Estos actores se encuentran precisados a admitir esta exigencia para los nuevos tiempos y el MIAR puede representar un importante terreno fértil desde el que pensar innovadoras ideas que ayuden a que la IA adquiera un compromiso con los DDHH y los ODS y con el fortalecimiento y enriquecimiento de las habilidades cívicas y la democracia.

La IAR es presentada como la respuesta a una necesidad frente a los desafíos del futuro y también como una oportunidad desde la que enriquecer el campo tecnológico desde la filosofía, construyendo de ese modo puentes de conocimiento entre saberes diferenciados. Una oportunidad que deben saber apreciar los agentes que promueven la toma de decisiones en nuestras sociedades, y que sin duda deben estar a la altura de los tiempos.

273

Bibliografía

BENJAMIN, M. (2014): *La guerra de los drones: matar por control remoto*, Barcelona, Anagrama.

BUREAU OF EUROPEAN POLICY ADVISERS (2001): *Empowering people, driving change. Social Innovation in the European Unión*, Luxemburgo, Oficina de Publicaciones de la Unión Europea.

COMISIÓN EUROPEA (2010): *EUROPA 2020 Una Estrategia para un crecimiento inteligente, sostenible e integrador*. Disponible en: http://ec.europa.eu/archives/commission_2010/2014/president/news/documents/pdf/20100303_1_es.pdf

CONILL SANCHO, J. (2013): *Horizontes de economía ética: Aristóteles*, Adam Smith, Amartya Sen, Madrid, Tecnos.

CORTINA, A. (1996): "El estatuto de la ética aplicada. Hermenéutica crítica de las actividades humanas", *Isegoría*, nº 13, pp. 119-127.

DASCAL, M. y DROR, I. E. DROR (2013): "The impact of cognitive technologies: Towards a pragmatic approach", *Pragmatic & Cognition*, n° 13.

DIAMOND, L. (1999): *Developing democracy. Towards consolidation*, Baltimore, The John Hopkins University Press.

DIÉGUEZ, A. (2003): "Tecnología y responsabilidad", *Revista de Filosofía*, n° 9, pp. 189-200.

DIGNUM, V. (2017a): "Responsible Autonomy", *International Joint Conference on Artificial Intelligence*.

DIGNUM, V. (2017b): "Responsible artificial intelligence: designing AI for human values", *ITU Journal: ICT Discoveries*, n° 1.

DIGNUM, V. (2018): "Ethics in artificial intelligence: introduction to the special issues", *Ethics and Information Technology*, n° 20, pp. 1-3.

DIGNUM, V. y PADGET, J. A. (2013): "Intelligent Robotics and Autonomous Agents", *Multiagents systems*, pp. 51-98.

DOMINGO MORATALLA T. (2017): "De la narración fílmica a la deliberación ética. La fenomenología hermenéutica como mediación (J. Marías y P. Ricoeur)", *SCIO. Revista de Filosofía*, n° 13, pp. 25-55.

ELLUL, J. (2003): *La edad de la técnica*, Barcelona, Octaedro.

FLORIDI, L. (2001): *The philosophy of information*, Oxford, Oxford University

FLORIDI, L. (2002): "Information ethics: An environmental approach to the digital divide", *Philosophy in the Contemporary World*, n° 9, pp. 39-45.

FLORIDI, L. (2002b): "On the intrinsic value of information objects and the infosphere", *Ethics and Information Technology*, vol. 4, n° 4, pp. 287-304.

FLORIDI, L. (2005) "The ontological interpretation of informational privacy", *Ethics Information Technology*, vol. 7, n° 4, pp. 185-200.

FLORIDI, L. (2006a): "Four challenges for a theory of informational privacy", *Ethics and Information Technology*, vol. 8, n° 3, pp. 109-119.

FLORIDI, L. (2006b): "Information technologies and the tragedy of the good will", *Ethics and Information Technology*, vol. 8, n° 4, pp. 253-262.

FLORIDI, L. (2007a): "Global information ethics: The importance of being environmentally Earnest", *International Journal of Technology and Human Interaction*, vol. 3, n° 3, pp. 1-11.

FLORIDI, L. (2007b): "Understanding information ethics", *American Psychology Association Newsletters*, vol. 8, n° 2, pp. 3-12.

FORGE, J. (2010): "A note on the definition of "dual use"", *Science and Engineering Ethics*, vol. 16, n° 1, pp. 111-118.

GOERING, S. y YUSTE R. (2016): "On the Necessity of Ethical Guidelines for Novel Neurotechnologies", *Cell*, vol. 167, n° 3, pp. 882-885.

GRACIA, D. (2016): "Problemas con la deliberación", *Folia Humanística. Revista de Salud, Ciencias Sociales y Humanidades*, n° 3.

GRUNWALD, A. (2014): "Technology Assessment for Responsible Innovation", en VV.AA.: *Responsible Innovation 1: Innovate Solutions for Global Issues*, Springer, Nueva York, pp. 15-32.

HEIDEGGER, M. (1997): *Filosofía, ciencia y técnica*, Santiago de Chile, Editorial Universitaria.

HORKHEIMER, M. (2010): *Crítica de la razón instrumental*, Madrid, Trotta.

HORKHEIMER, M. y ADORNO, T. (2016): *Dialéctica de la Ilustración: fragmentos filosóficos*, Madrid, Trotta.

KURZWEIL, R. (2013): *Cómo crear una mente: el secreto del pensamiento humano*, Berlín, Lola Books.

KURZWEIL, R. (2017): *La singularidad está cerca: cuando los humanos trascendamos la biología*, Berlín, Lola Books.

JONAS, H. (2008): *El principio de responsabilidad: ensayo de una ética para la civilización tecnológica*, Barcelona, Herder.

JORDÁN, J. y BAQUÉS, J. (2014): *Guerra de drones: política, tecnología, y cambio social en los nuevos conflictos*, Madrid, Biblioteca Nueva.

NEDON, V. (2018): *Open Innovation in R&D Departments*, Hamburgo, Springer.

UNIVERSITÉ DE MONTRÉAL (2018): *Montreal Declaration for a responsible development of artificial intelligence*. Disponible en: https://docs.wixstatic.com/ugd/ebc3a3_c5c1c196fc164756afb92466c081d7ae.pdf.

MORALES, A. C. (2008): "Innovación social: un proceso emergente en las dinámicas de desarrollo", *Revista de Fomento Social*, n° 63, pp. 411-444.

MORALES, A. C. (2009a): "Innovación social: un ámbito de interés para los servicios sociales", *Zerbitzuan: Revista de servicios sociales*, n° 45, pp. 151-175.

MORALES, A. C. (2009b): “Innovación ‘abierta’ en el tercer sector: el modelo organizativo 2.0”, *Revista Española del Tercer Sector*, n° 13, pp. 17-37.

MUMFORD, L. (1998): *Técnica y civilización*, Madrid, Alianza Editorial.

MUMFORD, L. (2017): *El mito de la máquina*, Logroño, Pepitas de calabaza.

QUINTANILLA, M. A. (2002): “La democracia tecnológica”, *Arbor*, n° 683-684, pp. 637-652.

PARSELIS, M. (2018): *Dar sentido a la técnica ¿pueden ser honestas las tecnologías?*, Madrid, Los libros de la Catarata.

RADNITZKY, G. (1973): “Hacia una teoría de la investigación que no es ni reconstrucción lógica ni psicología o sociología de la ciencia”, *Teorema*, n° 3, pp. 197-264.

SAREWITZ, D. y KARAS, T. H. (2007): “Policy Implications of Technologies for Cognitive Enhancement”, Sandia National Laboratories.

SKEEM, J. L. y LOWENKAMP, C. (2016): “Risk, Race & Recidivism: Predictive Bias and Disparate Impact”, SSRN.

URRA CANALES, M. (2017): *Estado, mercado, academia... y comunidad. Una cuádruple hélice para el desarrollo integral y la innovación*, tesis doctoral, Madrid, Universidad Pontificia Comillas.

VEGA REÑÓN, L. (2016): “Variaciones sobre la deliberación”, *Dilemata*, n° 22, pp. 203-230.

VV.AA. (2018): *The Malicious Use of Artificial Intelligence; Forecasting, Prevention, and Mitigation*. Disponible en: <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf>.

VV.AA. (2018): *Toronto's Declaration on Machine Learning*. Disponible en: https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf.

WALKER, W. R. y HERRMANN, D. J. (2005): *Cognitive Technology: Essays On The Transformation Of Thought And Society*, McFarland & Company.

Cómo citar este artículo

TERRONES RODRÍGUEZ, A. L. (2020): “Inteligencia artificial, responsabilidad y compromiso cívico y democrático”, *Revista Iberoamericana de Ciencia, Tecnología y Sociedad —CTS*, vol. 15, n° 44, pp. 253-276.