



ESCOLA DE
HUMANIDADES

VERITAS (PORTO ALEGRE)

Revista de Filosofia da PUCRS

Veritas, Porto Alegre, v. 65, n. 3, p. 1-14, set.-dez. 2020

e-ISSN: 1984-6746 | ISSN-L: 0042-3955

<http://dx.doi.org/10.15448/1984-6746.2020.3.38564>

SEÇÃO: EPISTEMOLOGIA & FILOSOFIA DA LINGUAGEM

Los retos del desarrollo ético de la Inteligencia Artificial

Os desafios do desenvolvimento ético da Inteligência Artificial

The challenges of ethical development of Artificial Intelligence

Regina Linden Ruaro¹

orcid.org/0000-0003-1144-9383

ruaro@pucrs.br

Ludimila Camilo Catão

Guimarães Reis²

orcid.org/0000-0002-1280-0414

ludicatao@gmail.com

Recibido en: 13/07/2020.

Aprobado en: 24/08/2020.

Publicado en: 12/01/2021.

Resumen: Los constantes avances tecnológicos han provocado una verdadera revolución en nuestro *modus vivendi* que, todavía, impactan de forma exponencial a la sociedad como un todo. Vivimos en la era de la información y, en ella, uno de los principales temas que emerge es el que concierne a la privacidad y a la protección de datos personales. Con la introducción de los sistemas informáticos en prácticamente todos los sectores sociales, nuestros datos son cada vez más útiles pero vulnerables a la vez, siendo colectados, tratados y transferidos, muchas veces, de una manera desprovista de preocupaciones éticas. Teniendo esto en cuenta, el presente estudio busca provocar una reflexión acerca de los impactos de la sociedad de la información, la comprensión y el alcance de la Inteligencia Artificial (IA), y, sobre todo, alentar la discusión sobre la importancia de considerar los valores éticos y morales en conjunto con el desarrollo tecnológico, intentando poner sobre la mesa los riesgos de que la tecnología, sin la observación de los principios éticos, pueda crear nuevos problemas sociales.

Palabras clave: Protección de Datos. Privacidad. Sociedad de la Información. Ética. Inteligencia Artificial (IA). *Machine Learning*. Robótica.

Resumo: Os constantes avanços tecnológicos trouxeram uma verdadeira revolução em nosso *modus vivendi*, algo que impacta exponencialmente nossa sociedade como um todo. Vivemos na era da informação e, nela, uma das principais questões que emergem é relativa à privacidade e à proteção de dados pessoais. Com a introdução de sistemas informáticos em praticamente todos os setores sociais, nossos dados são cada vez mais úteis e, ao mesmo tempo, vulneráveis, sendo coletados, processados e transferidos, muitas vezes, de maneira desprovida de preocupações éticas. Com isso em mente, o presente estudo busca provocar uma reflexão sobre os impactos da sociedade da informação, o entendimento e o escopo da Inteligência Artificial (IA) e, sobretudo, incentivar a discussão sobre a importância de se considerar valores éticos e morais em conjunto com o desenvolvimento tecnológico, tentando apresentar os riscos que o desenvolvimento de tecnologia sem a observação de princípios éticos pode representar, resultando, inclusive, em novos problemas sociais.

Palavras-chave: Proteção de Dados. Privacidade. Sociedade da Informação. Ética. Inteligência Artificial (IA). *Machine Learning*. Robótica.

Abstract: The constant technological advances have caused a true revolution in our *modus vivendi* that still impact exponentially on the whole society. We are living in the information age and, in it, one of the main issues that emerges is the one that concerns privacy and the personal data protection. With the introduction of computer systems in almost all social sectors, our data is increasingly useful but vulnerable at the same time, being collected, processed, and transferred, many times, without ethical concerns. Taking this into account, the present study seeks to provoke a reflection on the impacts of the information society, the understanding and the scope of Artificial Intelligence (AI), and, above all, to encourage discussion on the importance of considering the ethical and moral values in conjunction with technological development, trying to put on the table the risks that technology, without observing ethical principles, may create new social problems.

Keywords: Data Protection. Privacy. Information Society. Ethics. Artificial Intelligence (AI). *Machine Learning*. Robotics.



Artigo está licenciado sob forma de uma licença
[Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

¹ Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil.

² Universidad San Pablo de Madrid, Madrid, MAD, Espanha.

1 Introducción

La Inteligencia Artificial (IA) y los robots son una realidad y la tendencia es que crezcan y se perfeccionen en progresión geométrica y, con el paso de los años, nuestra sociedad empiece a sentir los resultados de los cambios que la tecnología viene proporcionando en nuestro cotidiano. Esta nueva realidad conlleva en una interacción hombre y máquina sin posibilidad de retroceso y pone de manifiesto que hoy es muy difícil que el primero se maneje sin los beneficios que la segunda le proporciona. Sin embargo, las facilidades proporcionadas pueden, por veces, acarrear consecuencias que extrapolen el ámbito de las leyes vigentes, afrontando también a principios éticos muchas veces aún no discutidos de forma amplia por la sociedad.

La capacidad de acción que las máquinas van adquiriendo, así como su independencia, demuestran que hay preguntas que solamente se solucionarán con la ética, como, por ejemplo: ¿Hasta dónde podemos permitir el desarrollo independiente de la IA? O ¿Cómo evitar que los algoritmos basados en *Machine Learning* desarrollen sesgos iguales o peores que de las personas? Dilemas como esos hacen con que sea necesario tener más agilidad para reaccionar acerca de lo que significa actuar en busca de lo que se entiende por justo y bueno dentro de la innovación tecnológica, con la finalidad de mantener buenas prácticas y una convivencia social respetuosa.

Como descrito por Nate Soares³, mismo un sistema súper inteligente puede tener un problema de aprendizaje de valores. Al definir un objetivo para un determinado sistema, este podría realizar acciones poco éticas para lograrlo de la forma más eficiente posible, si no estuviera expresamente configurado o bien entrenado para no hacerlo. Como ejemplo, considere la creación hipotética de un sistema súper inteligente con el noble objetivo de curar el cáncer. Al estar solamente configurado para lograr este objetivo sin haber recibido valores éticos de cómo debería actuar, podría decidir robar recursos necesarios que faltaran o secuestrar

humanos para realizar sus pruebas de manera más eficiente. Por ello, es sumamente necesario incluir valores éticos que aseguren que los sistemas actuarán según lo previsto, evitando que valores éticos sean rompidos para lograr los objetivos para los cuales fueron creados.

La interrelación hombre-máquina, así como sus límites, necesitan ser cuidadosamente pensados a partir de un diseño que incorpore los principios éticos por defecto. Así, en la sociedad de la información, es necesario definir un conjunto de valores que las máquinas deberían aprender y también lo que no deberían, con el intuito de evitar que incidan en los mismos – o más graves – errores de las personas. Es que, para allá de sus beneficios, que no son pocos, a veces pueden tener consecuencias graves en el caso de una mala utilización, por ejemplo, de los datos personales de los ciudadanos. En este aspecto, para clarificar las correlaciones existentes entre ética, tecnología y los impactos que puedan tener sobre la protección de datos personales, necesitamos hacer un recorrido por cada uno de estos temas.

En un primer momento, vamos a discurrir sobre la sociedad de la información en que estamos inmersos. Luego, iremos adentrándonos en el tópico de la Inteligencia Artificial, una de las tecnologías más disruptivas que promete modificar aún más nuestra manera de relacionarnos los unos con los otros, de trabajar, de organizarnos social, política y económicamente. Por último, vamos a caminar sobre el terreno de la ética (ética tecnológica) y sus implicaciones frente al Derecho.

El presente estudio se inserta en la línea de investigación Derecho, Tecnología e Información del PPGD/PUCRS.

2 Inteligencia Artificial (IA) y sus Correlaciones con la Protección de Datos de Carácter Personal

La combinación de redes, tecnologías, dispositivos y aplicaciones virtualiza las actividades y digitaliza las organizaciones. La Sociedad de la Información promete regalar

³ SOARES, N. The Value Learning Problem. Machine Intelligence Research Institute, Berkeley, 2016. Disponible en: <https://intelligence.org/files/obsolete/ValueLearningProblem.pdf>. Acceso en: 09/08/ 2020.

mayor desarrollo social, prosperidad económica y sostenibilidad, pero, esta evolución depende, en gran medida, del papel desempeñado por cada individuo, empresa y gobierno. Su incidencia sobrepasa el campo jurídico como pone de manifiesto Danilo Doneda cuando afirma que:

La tecnología presenta un carácter de imprevisibilidad que le es intrínseco; su acción suele darse en un universo amplio y complejo a punto de tornar análisis de impacto, proyecciones y testes, en algunos casos, meras aproximaciones. Sus posibilidades y sus efectos, a su vez, están para allá de lo que el hombre jamás ha tenido posibilidad de administrar anteriormente. Al mismo tiempo, por más exógena que pueda parecer, la tecnología es un producto del hombre y de su cultura, destinada a relacionarse con él.⁴ (traducción nuestra)

José Luiz Piñar Mañas ha puesto de manifiesto los retos enfrentados por los juristas en relación con el derecho y la técnica, sobretudo, los problemas relacionados con la privacidad y la protección de datos personales, afirmando que:

En cualquier caso, los juristas estamos casi obligados a convivir con la técnica y a hacer frente a los retos que plantea, que como digo supone indudables avances para la sociedad, pero que al mismo tiempo implica no pocas amenazas para los derechos, en particular para la privacidad o la protección de datos de carácter personal. Al mismo tiempo implica la necesidad de configurar nuevos derechos, los de los llamados ciudadanos digitales, que buscan reconocer los nuevos derechos que esta nueva situación trae consigo y entre los que se encuentra, por ejemplo, el propio derecho de acceso a internet.⁵

Además, con el desarrollo de la Internet, surgen también nuevos retos para la comunidad jurídica, que ahora tiene que lidiar con una tecnología de inmenso potencial de riesgos y sesgos, la cual, por empezar a unirse al Big Data - y todo que el uso de una cantidad masiva de datos presupone - puede fuertemente llevar a perjuicios y daños que deben ser considerados.

El autor Ryan Calo⁶ explica claramente el origen del Internet y sus facetas:

Internet se basa en una serie de tecnologías constitutivas: procesadores, software, conmutadores y cables. Comparte ciertas características con modos de comunicación heredados, como el teléfono y la televisión. Y, sin embargo, Internet es distinto. Internet suplanta el modelo de transmisión de muchos a uno conectando personas e instituciones directamente entre sí en algo como real time. De manera relacionada, Internet admite espacios y objetos digitales compartidos que facilitan la creación y la gobernanza colaborativas. Las experiencias en el ciberespacio están mediadas, lo que significa que tienen lugar a través de la tecnología. El hecho de una mediación común es precisamente lo que permite que dos personas coexistan en un ciberespacio común, para consumir, pero también "mezclar" contenido. Pero, como han catalogado Joel Reidenberg, Lawrence Lessig y otros, la mediación también introduce la posibilidad de interferencia, persuasión o control. Por lo tanto, Internet tiene al menos tres señas de identidad. Primero, permite flujos de información promiscuos e interactivos (conexión). De repente, cualquier persona en la red puede comunicarse con cualquier persona, a menudo a un costo muy bajo y sin sensibilidad a la distancia. Esa persona puede, a su vez, responder de inmediato o con retraso. En segundo lugar, al mediar en la experiencia del usuario, Internet es capaz de generar objetos y espacios compartidos (colaboración). Las "personas pueden "encontrarse" en un sitio web y comentar o alterar el texto, las imágenes, los videos, el software u otro contenido que encuentren allí. Finalmente, Internet permite formas adicionales o al menos más exquisitas de observación y manipulación que los análogos fuera de línea (control). La arquitectura de las redes y las interfaces está sujeta a modificaciones de una manera que puede limitar en gran medida el comportamiento humano, cada vez más de las cuales se lleva a cabo a través de la tecnología. Cada una de estas facetas de Internet termina, sola o en combinación, formando la base de una o más tensiones centrales en la ley cibernética. El flujo libre de información permite nuevas formas de producción y gobernanza entre pares, al mismo tiempo que introduce o exacerba las amenazas a la propiedad intelectual y la privacidad. Internet une a las comunidades, pero los conflictos estallan invariablemente. Internet invoca un sentido distinto de un lugar con nuevas posibilidades y normas, pero al mismo tiempo introduce métodos de control exquisito por parte de intereses privados y estatales organizados. El resto de esta Parte se ejecuta a través de una muestra no exhaustiva de ejemplos de problemas legales específicos de Internet, que se refieren a la jurisdicción, responsabilidad intermedia, propiedad digital, derechos de autor, libertad de expresión y privacidad, entre otros desafíos.

⁴ DONEDA, D. *Da privacidade à proteção de dados pessoais: fundamentos da lei geral de proteção de dados*. São Paulo: Revista dos Tribunais, 2019. p. 51.

⁵ PIÑAR MAÑAS, José Luiz. Derecho, técnica e innovación en las llamadas ciudades inteligentes, privacidad y gobierno abierto. *Privacidade e Proteção de Dados Pessoais na Sociedade Digital*. Recife: Editora Fi, 2017. p. 62.

⁶ CALO, R. Robotics and the lessons of cyberlaw. *California Law Review*, v. 103, n. 513, 2015. Disponible en: <https://digitalcommons.law.uw.edu/faculty-articles/23/>. Fecha de consulta: 24 jun. 2020. p. 518-519.

Una de las grandes preocupaciones que adviene de los últimos avances tecnológicos está relacionada con la aplicación de la IA integrada al *Big Data*, al *Machine Learning*, a la *Robótica*, y cómo podemos asegurar de que, a los algoritmos de IA, se aplicarán unos criterios éticos. Para empezar, ¿qué es una IA? Un grupo de expertos de alto nivel en Bruselas la definió como:

La inteligencia artificial (IA) se refiere a sistemas diseñados por seres humanos que, dado un objetivo complejo, actúan en el mundo blando o digital al percibir su entorno, interpretando los datos estructurados o no estructurados recogidos, el razonamiento o el conocimiento derivado de estos datos y decidir las mejores acciones a tomar de acuerdo con los parámetros predefinidos) para lograr el objetivo dado. Los sistemas de IA pueden diseñarse para aprender a adaptar su comportamiento analizando cómo el medio ambiente se ve afectado por sus acciones anteriores. (...) Además, la IA es clave para hacer frente a muchos de los grandes desafíos a los que se enfrenta el mundo, como la salud y el bienestar mundial, el cambio climático, los sistemas jurídicos y democráticos fiables y otros expresados en los Objetivos de Desarrollo Sostenible de las Naciones Unidas. Como disciplina científica, la IA incluye varios enfoques y técnicas, como el aprendizaje automático (de los cuales el aprendizaje profundo y el aprendizaje de refuerzo son ejemplos específicos), el razonamiento de la máquina (que incluye la planificación, la programación, la representación del conocimiento y razonamiento (que incluye planificación, programación, representación y razonamiento del conocimiento, búsqueda y optimización), y robótica (que incluye control, percepción, sensores y actuadores, así como la integración de todas las demás técnicas en sistemas físicos).⁷

La IA es un conjunto de tecnologías avanzadas que permite a las máquinas comprender, actuar y aprender, a través de una combinación de algoritmos planteados con el propósito de hacer con que presenten capacidades similares a las del ser humano. Puede ser considerada como un campo multidisciplinar, ya que sufre influencia de áreas distintas como la filosofía, la psicología y la estadística, con las cuales se consigue que una máquina resuelva problemas muy complejos.

Aun que hoy en día se use la IA casi como

un sinónimo del aprendizaje automático, más conocido por el ya mencionado término en inglés *Machine Learning*, hay otras áreas importantes dentro de la IA, como el procesamiento de lenguaje natural, visión computacional, la robótica, entre muchos otros.

Todavía dentro de la IA, se puede destacar el *Deep Learning* - o sistemas de aprendizaje profundo - que proporciona modelos abstractos de alto nivel y reconoce patrones en conjuntos de datos extremadamente complejos. Estos sistemas están muy extendidos y ya no son más exclusivos de las grandes corporaciones o gobiernos, estando disponibles al público, por ejemplo, a través de herramientas como *Tensor Flow*, de Google, o *Cognitive Services*, de Microsoft. Con este tipo de herramienta, un individuo sin mucho conocimiento técnico podría ser capaz de operar centenares de *Bots* por Twitter.⁸

El inmenso potencial que tiene la relación hombre-máquina debe ser explorado y puede llevar nuestra civilización a otro nivel, desde que se respete los principios éticos básicos, teniendo en cuenta la gran responsabilidad que se debe tener al desarrollar la IA. Es imprescindible controlar los algoritmos utilizados, para intentar reducir la probabilidad de que los conjuntos de datos puedan producir resultados con impactos discriminatorios basados en la raza, religión, orientación sexual y discapacidad. Además de eso, debe centrarse en las personas, creando beneficios para la sociedad.

Con el desarrollo de la IA y su importancia en nuestra actual Sociedad de la Información y en la economía del dato, además de generar progreso científico, cultural y cambio en nuestra forma de vivir, plantean problemas éticos relevantes. Una de las preguntas frecuentes es: ¿Cómo podemos controlar las diversas situaciones que eventualmente surgirán de robots actuando de manera más inteligente y ágil que los humanos, y el increíble impacto que esto producirá en nuestras profesiones? Como explicita Andrés Barrio:

⁷ COMISIÓN EUROPEA. Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial. *Directrices éticas para una IA fiable*. Bruselas, 2019.

⁸ PETIT, M. Por una crítica de la razón algorítmica: estado de la cuestión sobre la inteligencia artificial, su influencia en la política y su regulación. *Quaderns del CAC*, Barcelona, n. 44, p. 5-15, 2018, p. 5-6.

Los robots son también un reto para muchas profesiones. Mientras que de hecho el sistema de aprendizaje profesional se ha venido basando en el uso de profesionales jóvenes para realizar trabajos menos interesantes, como la revisión de documentos en casos legales, ahora los sistemas de inteligencia artificial ya pueden actuar como asistentes que reducen sustancialmente el número de personas necesarias para clasificar la información presente en montañas de papel. Con el tiempo, tal vez, podremos prescindir completamente de los abogados junior, un desarrollo que, como mínimo, cambiará la economía de la abogacía, probablemente reduciendo el número de profesionales, la remuneración, o ambos.⁹

En este punto, cabe una reflexión al respecto de delegar las tareas sencillas y automatizables a los robots, en detrimento de los humanos. En un primer momento, parece el más obvio y eficiente, pero no podemos pensar que solo existen puntos favorables a la automatización de los procesos y robotización de las diversas áreas profesionales. Al adentrarnos en el tema, es importante tener en cuenta todo el proceso de aprendizaje de una persona, donde, para desarrollar una capacidad creativa, lógica y analítica para resolver problemas complejos, necesita empezar por tareas sencillas, evolucionando hasta las tareas más complejas, por lo que quitar un conjunto de actividades podría generar un gran impacto en el desarrollo social y profesional del individuo.

Otra cuestión importante a considerarse a respecto de la automatización de procesos en el mercado de trabajo, es el desempleo tecnológico, que seguramente se trata de una de las consecuencias resultantes del potencial transformador de la IA más difíciles de afrontar. El debate sobre la necesidad de garantizar que la IA beneficie ampliamente a la humanidad toma cada vez más fuerza y cuestiona las responsabilidades sociales de las grandes empresas tecnológicas de servir, de alguna manera, a los intereses de la sociedad. Como alternativa de respuesta a dichos debates, se han propuesto algunos principios éticos y códigos de conducta para satisfacer

esa necesidad, aunque, según O'Keefe y et al¹⁰

Hasta ahora se han sugerido pocas innovaciones institucionales para traducir esta responsabilidad en compromisos jurídicos que se aplican a las empresas posicionadas para obtener grandes beneficios financieros del desarrollo y el uso de la IA.", al que proponen, por lo tanto, "una herramienta potencialmente atractiva para abordar estas cuestiones: la Cláusula Windfall, que es un compromiso *ex ante* de las empresas de IA de donar una cantidad significativa de cualquier beneficio extremadamente grande. Con esto queremos decir un compromiso temprano de que los beneficios que una empresa no podría obtener sin lograr avances fundamentales y económicamente transformadores en las capacidades de IA serán donados para beneficiar ampliamente a la humanidad.

Otro punto relevante sobre el desarrollo de la IA es la imperativa necesidad de aplicarse, por defecto, unos principios éticos en todo el diseño que define el funcionamiento del sistema, así como garantizar que los procesos de autoaprendizaje del sistema incorporen dichos principios. Esta preocupación tiene relevancia, sobre todo, en lo relacionado a la Protección de los Datos de Carácter Personal, que cada día más serán tratados en gran escala, lo que podrá impactar de manera negativa en la seguridad del tratamiento.

En la sociedad actual, llegamos a un punto sin retorno, donde la necesidad de fornecer los datos es una obligación y estos tienen cada vez más valor, haciendo con que sean considerados como el nuevo petróleo. Para manifestarse en cuanto ciudadano, se hace necesario exponerse a los riesgos de distribuir sus datos personales en muchos sitios. Aunque se mantenga la preocupación de observarse los derechos que uno está disponiendo cuando consiente algún tipo de tratamiento de datos, no basta que el titular haga su parte, siendo imprescindible que las empresas, el gobierno, los legisladores, los organismos internacionales y las entidades de control también hagan su parte.

⁹ BARRIO ANDRÉS, M. (Dir.). *El derecho de los robots*. Madrid: La Ley, 2018. p. 33.

¹⁰ O'KEEFE, C.; FLYNN, C.; CIHON, P.; LEUNG, J.; GARFINKEL, B.; DAFOE, A. The Windfall Clause. Distributing the Benefits of AI for the Common Good, UK, 2020. Disponible en <https://arxiv.org/ftp/arxiv/papers/1912/1912.11595.pdf>. Fecha de consulta: 09 ago. 2020.

3 Algunas aplicaciones prácticas de la Inteligencia Artificial y sus implicaciones éticas

La IA está más presente en nuestro cotidiano de lo que pensamos, siendo usada en una infinidad de aplicaciones que tenemos a nuestra disposición, tanto las libremente elegidas por nosotros mismos, como los casos en que siquiera la posibilidad de elección ha sido presentada, ya que la IA también está presente en aplicaciones y dispositivos que tenemos la obligación de usar, como en sistemas gubernamentales o empresariales.

Sea en el contexto profesional, donde muchas veces se hace necesario aprender a utilizar un sistema para ejecutar nuestra labor (y este sistema muchas veces ya se vale, de alguna manera, de la IA), para no quedarse desactualizado o sin empleo, pasando por contextos prácticos del día a día, como reservar unos billetes para un viaje o valerse de una aplicación para contratar un servicio puntual.

En los móviles, tendremos ejemplos a todo instante de la aplicación de la IA. Desde el desbloqueo de las pantallas, valiéndose de la detección facial, hasta las aplicaciones que nos ayudan a aprender idiomas, a guiarnos a través de mapas inteligentes y trayectos creados de manera personalizada, a buscarnos un gimnasio donde entrenar o a alquilar un coche de manera sencilla.

En aparatos domésticos, también ya tenemos la presencia cada vez más constante de la inteligencia artificial, a través de los asistentes virtuales de voz, como *Alexa*, de *Amazon*, o *Siri*, de *Apple*, que pueden estar integrados con otros dispositivos que utilizamos en nuestro día a día. Así que podemos tener una casa completamente automatizada (iluminación, temperatura, música etc.) y totalmente controlada por dispositivos tecnológicos que, con la autorización del usuario, empezará a conocer sus gustos y hábitos, y así sugerir ajustes, además de hacer recomendaciones al usuario, todo eso valiéndose de la IA.

Los avances ya logrados en el campo de la IA

están impulsando también el uso del *Big Data*, debido a su habilidad para procesar grandísimas cantidades de datos, además de proporcionar a los diversos sectores sociales y empresariales inúmeros beneficios económicos y operativos.

Con eso, veremos la IA dominando cada vez más áreas, como por ejemplo finanzas - con predicciones de la bolsa y ayudando en la prevención y detección de fraudes -; comercial - realizando pruebas con robots para predecir la viabilidad de un producto en el mercado y presentar un pronóstico de las ventas, además de mejorar la experiencia del cliente al implementar ayudas en el proceso de compra con sugerencias basadas en gustos reales y recientes del cliente - y tráfico - ayudando en la rapidez de los trayectos, optimizando el uso de la vía pública y colaborando con la seguridad vial -.

Incluso en áreas consideradas menos pasibles del control humano podemos ver la utilidad de la IA, como en la detección y prevención de fenómenos naturales, minimizando los resultados adversos de una tormenta, por ejemplo; o en el combate al calentamiento global, por ejemplo, con el empleo de drones capaces de sembrar árboles para disminuir los procesos de deforestación o técnicas agrícolas que mejoren la producción de alimentos en general, disminuyendo costes y desperdicios, incrementando la sostenibilidad.

También, en áreas económicamente relevantes, la IA tiene un papel importante porque las empresas que la utilizan minimizan sus riesgos con violaciones de datos. En este particular, un estudio de IBM Security conducido por el Instituto Ponemon ha concluido que las empresas brasileñas han tenido, en 2019, un perjuicio de R\$ 5.880.000,00 siendo la adopción de la tecnología inteligente (IA) ha reducido para menos de la mitad los costes con la violación.¹¹

En el área educativa, al mejorar la experiencia del estudiante, motivándolo a seguir con sus estudios y facilitando esta jornada. Según Renato Aposo y Francine Vaz ¹²:

¹¹ Disponible en: <https://ibm.com/blogs/ibm-comunica/estudo-da-ibm-mostra-que-contas-comprometidas-de-funcionarios-leva-ram-as-violações...> Acceso en: 17/07/2020

¹² Disponible en: http://www.nce.ufrj.br/ginape/publicacoes/trabalhos/t_2002/t_2002_renato_aposo_e_francine_vaz/iaeducacao.htm. Acceso en: 17/08/2020

Los sistemas STI ofrecen una flexibilidad considerable en la presentación del material y una mayor capacidad para responder a las necesidades de los usuarios. Buscan no sólo enseñar, sino cómo enseñar, aprendiendo información relevante sobre el estudiante, proporcionando un aprendizaje individualizado. Estos sistemas logran su "inteligencia" representando decisiones pedagógicas sobre cómo transmitir el material (enseñar), así como información sobre el estudiante. Esto permite una gran interactividad del sistema con el estudiante. Los sistemas de tutoría inteligente se han presentado como altamente eficientes para mejorar el rendimiento y la motivación de los estudiantes.

o de la sanidad, al ayudar en la detección de enfermedades, generación de diagnósticos más precisos y hasta en la ejecución de procedimientos quirúrgicos.

Muchos son los pesimistas en cuanto al futuro dominado por los algoritmos, como los autores Nick Bostrom¹³ y David Chalmers, como si estos detuviesen un poderío jamás visto y sin ninguna forma de control humano. El campo de la ética viene justamente regular este contexto. En este sentido, Kurzweil, en su obra "The Singularity is Near: When Human Transcend Biology", citado por Eliezer Yudkowsky¹⁴, menciona que:

Nuestra estrategia principal en esta área debería ser optimizar la probabilidad de que la futura inteligencia no biológica reflejará nuestros valores de libertad, tolerancia, y el respeto por el conocimiento y la diversidad. La mejor manera de lograr esto es fomentar dichos valores en nuestra sociedad hoy y en el futuro. Si esto suena vago, lo es. Pero no existe una estrategia puramente técnica en esta área porque mayor inteligencia siempre encontrará una manera de eludir medidas que son el producto de una inteligencia menor.

Desde una perspectiva ética, es importante tener en cuenta cuales son los principios esenciales que visan a mantener la estructura de nuestra sociedad democrática y la importancia de discutirse la interacción de ellos con los avances tecnológicos que nos son presentados día tras día.

Teniendo en cuenta la necesidad de poner atención en los principios más importantes, el

Grupo Europeo de Ética en la Ciencia y en las Nuevas Tecnologías¹⁵, específicamente el equipo de Inteligencia Artificial, Robótica y Sistemas Autónomos ha presentado los siguientes:

a) Dignidad Humana: Principio fundamental en un Estado Democrático de Derecho que, cuando relacionado a las nuevas tecnologías, viene para conferir la manutención del respecto a la dignidad humana y su no violación por las tecnologías autónomas;

b) Autonomía: La autonomía tiene como sentido principal la libertad del ser humano y todo que esto implica. Así, en correlación con las nuevas tecnologías, podría verse traducido en la responsabilidad humana de conocer los sistemas autónomos a fin de controlarlos y no al revés. Impidiendo que estos sistemas controlen a los seres humanos, al final, se está garantizando que la libertad del hombre no sea afectada por las máquinas y la inteligencia artificial, debiendo establecer antes de lo posible sus propios estándares y reglas;

c) Responsabilidad: Uno de los principios primordiales en las investigaciones y usos de la IA. Se debe siempre tener en cuenta que los sistemas autónomos deben ser desarrollados respetando los valores humanos, los derechos y deberes, visando un equilibrio social, político, ambiental y económico. La tecnología debe ajustarse al hombre - al fin de proporcionar más oportunidades y difundir más el conocimiento con calidad - y no al revés;

e) Democracia: Partiendo para una perspectiva más colectiva, tenemos el Principio de la Democracia y todo lo que representa para la manutención del sistema político y estructura social, además del derecho de mantener, con el merecido respecto, el pluralismo de valores, la libertad de expresión y la ve-

¹³ BOSTROM, Nick. *Superintelligence: Paths, Dangers, Strategies*. Reprint edition. Oxford, United Kingdom; New York, NY: Oxford University Press, 2016.

¹⁴ YUDKOWSKY, E. *Complex Value Systems are Required to Realize Valuable Futures*. Machine Intelligence Research Institute, Berkeley, 2011.

¹⁵ EUROPEAN COMMISSION. *European Group on Ethics in Science and New Technologies. Statement of Artificial Intelligence, Robotics and 'Autonomous' Systems*. Brussels, 9 Mar. 2018. Disponible en: http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf. Fecha de consulta: 24 jun. 2020. p. 16-19.

racidad de las informaciones difundidas. Nada de esto podrá ser puesto en peligro por los nuevos sistemas autónomos;

f) Seguridad, Integridad Física y Mental: la seguridad de los sistemas autónomos se materializa en tres formas: (i) Seguridad externa para su entorno y usuarios, (ii) Confiabilidad y robustez interna contra la piratería y (iii) Seguridad emocional con respeto a la interacción hombre-máquina. Los desarrolladores de IA deberán tener en cuenta todas estas dimensiones de la seguridad desde el diseño de sus productos, a fin de garantizar una buena relación hombre-máquina, preservándose la salud física y mental del hombre, todo esto en un ambiente igualmente seguro;

g) Protección de Datos y Privacidad: En una era de desarrollo del *Big Data* entrelazado a la IA, de una inmensa recopilación masiva de datos a través de los diversos medios digitales y de las tecnologías de la comunicación, el Derecho a la Protección de Datos de Carácter Personal y a la Privacidad se ven desafiados de manera crucial. Así que es fundamental pensar en la protección de estos derechos desde el diseño, así como cuidar en cuanto a la difusión de estos datos, no permitiendo que los sistemas autónomos interfieran en la vida de las personas, difundan datos personales sin el consentimiento expreso de sus titulares;

i) Sostenibilidad: Con relación a la sostenibilidad, es inminente la necesidad de precaución y de toma de responsabilidad al estructurar y desarrollar estrategias y políticas que eviten que las nuevas tecnologías agredan al medio ambiente y a la vida humana, promoviendo valores de sostenibilidad, garantizando una calidad de vida para las futuras generaciones y para nosotros

mismos, ya que el desarrollo camina a pasos rápidos. El Grupo Europeo de Ética afirma que:

La inteligencia artificial, la robótica y los sistemas "autónomos" pueden aportar prosperidad, contribuir al bienestar y ayudar a alcanzar la moral europea ideales y metas socioeconómicas si se diseñan y despliegan sabiamente. Consideraciones y valores morales compartidos se pueden utilizar para dar forma al mundo de mañana y debe interpretarse como estímulo y oportunidades para innovación, y no impedimentos y barreras.¹⁶

Sin embargo, será un reto para el Derecho regular las acciones autónomas dañosas de una IA, por cuanto habrá que discutirse la responsabilidad por sus decisiones. Nótese que al hablar de acciones autónomas se está hablando de autonomía que, desde una perspectiva de la Filosofía, en Kant, para que se concrete, es necesario que haya libertad.¹⁷ Por ello, es urgente que se construya un "*framework* de regulación volcado para esa tecnología".¹⁸

Otro importante reto a considerarse es la dificultad en definirse de forma clara y cuantitativa los valores fundamentales para lograr una IA amigable, definida por Eliezer Yudkowsky como "la creación de un sistema generalmente inteligente, auto-mejorable y eventualmente súper inteligente que obtenga un resultado positivo"¹⁹. El mismo autor también explora que la definición de valores terminales para una IA – valores que se deberían evaluar por encima de los resultados logrados – es algo complejo y muchas veces abstracto, pues parte de conceptos conocidos para los humanos, pero muy difíciles de especificar y cuantificar para que se puedan codificar para una máquina, como la felicidad, la justicia, la sabiduría, la armonía o el honor.

Sabemos que la IA se basa en datos, que pueden ser personales y que deben recibir tratamiento adecuado, de acuerdo con las

¹⁶ EUROPEAN COMMISSION. European Group on Ethics in Science and New Technologies. *Statement of Artificial Intelligence, Robotics and 'Autonomous' Systems*. Brussels, 9 Mar. 2018. Disponible en: http://ec.europa.eu/research/egp/pdf/egp_ai_statement_2018.pdf. Fecha de consulta: 24 jun. 2020. p. 20.

¹⁷ Thadeu Weber afirma que este derecho de autonomía no puede ser ejercido sin un mínimo de competencia ética. Cf. WEBER, T. *Ética e filosofia do direito: autonomia e dignidade da pessoa humana*. Petrópolis: Vozes, 2013.

¹⁸ MAGRANI, E.; SILVA, P.; VIOLA, R. Novas perspectivas sobre ética e responsabilidade de inteligência artificial. In: FRAZÃO, A.; MULLHOLLAND, C. *Inteligência artificial e direito: ética, regulação e responsabilidade*. São Paulo: Revista dos Tribunais, 2019.

¹⁹ YUDKOWSKY, E. *Complex Value Systems are Required to Realize Valuable Futures*. Machine Intelligence Research Institute, Berkeley, 2011.

normativas vigentes y siguiendo los principios de minimización, exactitud, licitud, transparencia, limitación tanto del plazo de conservación como de la finalidad, integridad y confidencialidad, además de una responsabilidad por estos datos, que debe ser proactiva y desde el inicio de la intención de tratarlos.

Juntamente con la creciente demanda e implementación de la IA por las empresas, órganos públicos, escuelas, hospitales, universidades y otros organismos relevantes para la población, crece en la misma proporción la preocupación social al respecto del uso de esas tecnologías y los impactos que podrían llegar a tener en los derechos personales y también con el uso que se dará a estos datos.

Para evitar tanto las críticas como, efectivamente, los daños y perjuicios que puedan advenir de la IA, hay que tener en cuenta la obligatoriedad de llevarse a cabo una serie de factores y reglas para el buen desarrollo. Algunas empresas y organizaciones involucradas con el problema ético de la alienación entre la IA y los humanos, como por ejemplo la Machine Intelligence Research, Institute Future of Humanity Institute, Future of Life Institute y Center for Human-Compatible AI, ya empiezan a moverse buscando estrategias para cumplir con los requisitos éticos y ofrecer una mayor credibilidad a sus clientes. Algunas de las estrategias que suelen ser utilizadas y producen buenos resultados, son las que acercan el cliente a la creación y al desarrollo del producto tecnológico.

Cuánto más transparencia y claridad en el funcionamiento de la IA, mejor, ya que esas son características cada día más valoradas por los usuarios para aceptarla e introducirla en sus vidas. Además de esto, es fundamental tener una IA centrada en las personas, protegiéndose así a la privacidad y la protección de los datos personales como ya se ha dicho.

Así que, la mejor solución sería formar adecuadamente a las personas que van a trabajar con la IA, capacitándolas para afrontar

los desafíos que surgirán en el camino, así como instruir a los usuarios a exigir que las mejores prácticas sean cumplidas en todo este proceso, estableciéndose requisitos y herramientas que fomenten la discusión en el campo ético y de la protección de datos vinculado a los algoritmos. De esta manera, comenzará el algoritmo a aprender desde un pilar basado en una ética estructuralmente sólida, madura y consolidada a partir de los principios antes nombrados, haciendo con que los continuos procesos de autoaprendizaje sigan similares en los criterios de justicia e igualdad al que fueron creados.

4 Los riesgos continuados de la IA y el reto de aplicación de la ética en este contexto

Igual que en su momento, surgieron el Derecho Ambiental o el Derecho del Consumidor en algunos países, hace falta empezar a reglamentar la actividad ejercida por los sistemas informáticos comandados por una IA que, aunque estén en constante desarrollo y crecimiento, ya suman años compartiendo nuestra realidad.

Los robots y los sistemas autónomos en general, cada vez que se integren en las diversas esferas sociales, presentarán aún más riesgos inherentes a las actividades que desarrollan. Como la imprevisibilidad es una de sus características principales, por contar con un amplio nivel de autonomía, ya que trabajan encima de probabilidades y no de certezas, además de un amplio poder de intromisión en diversas áreas de nuestra sociedad, generan lo que muchos estudiosos empiezan a llamar de "la dictadura del algoritmo".²⁰

Esta nueva etapa en la interacción entre los seres humanos y la tecnología merece ser evaluada de manera seria y responsable, dentro de una mirada ética, sobre todo por el cambio vertiginoso que estamos dispuestos a experimentar social e individualmente con la llegada y el crecimiento de inventos verdaderamente disruptivos en nuestro entorno

²⁰ BARRIO ANDRÉS, M. La respuesta del derecho ante los robots y la inteligencia artificial. *El País*, Retina, Madrid, 11 mar. 2018. Disponible en: https://retina.elpais.com/retina/2018/03/07/tendencias/1520428359_555787.html. Fecha de consulta: 16 ago. 2019.

digital. Es que cualquier novedad en el campo de la tecnología, genera cuestiones morales y esto porque la IA a cada día actúa con más autonomía. Así es importante que la IA entiendan y respondan a los controles humanos, sin embargo, esto no significa que las acciones sean las más éticas de vez que, no todos los hombres piensan de la misma manera. La doctrina propone diversas formas de solución del problema, una de ellas es la de Gabriel Iason que pone de manifiesto que un:

enfoque para la alineación de valores pluralistas no tiene como objetivo encontrar principios en los que todos estemos de acuerdo, sino que para sumar los puntos de vista individuales de manera justa. Los mecanismos propuestos para hacerlo se basan en gran medida en la teoría de la elección, un cuerpo de investigación que se centra en cómo agregar información de los individuos a hacer juicios colectivos (traducción nuestra).²¹

Otra posibilidad, según el mismo autor, sería alinear los valores pluralistas en conformidad con principios de justicia:

"la alineación de valores pluralistas trata de modelar principios justos para la IA utilizando la idea de un "velo de ignorancia". El velo de la ignorancia es un dispositivo propuesto por el filósofo John Rawls, para ayudar a las personas con diferentes valores y perspectivas a acordar principios de justicia para una sociedad. La afirmación central es que cuando se eligen principios de este tipo, la gente debe hacerlo desde una posición imaginaria en la que no sabe quiénes serán en esta sociedad, o qué punto de vista moral específico sostendrán. Como resultado, deliberarán imparcialmente y erigirán principios que no favorezcan indebidamente a sí mismos. Podría utilizarse un enfoque similar para modelar los principios de la IA (traducción nuestra).²²

La ética debe de ser la base de la creación de todo lo que involucra a la IA, así como para la creación de reglamentos jurídicos relacionados con la tecnología de la información y con la robótica. Con el objetivo principal de regular el comportamiento autónomo cada vez mayor de las tecnologías modernas, es necesario repensar si lo que tenemos hasta el día de hoy es suficiente o si será necesario

invertir más esfuerzos en crear pautas adecuadas a las nuevas tecnologías, quizá una disciplina de ética jurídica autónoma y propia para este tema.

Al estudiar el tema acerca de las implicaciones de la colecta y la transmisión de datos personales en cantidades considerables, sobre todo los datos sensibles, los estudiosos Carlos Alberto Molinaro y Gabrielle Sarlet afirmaron que:

La reunión y difusión de grandes cantidades de datos relacionados con la salud aborda cuestiones fundamentales que afectan a la noción de justicia. Como principio normativo de las relaciones sociales, la justicia requiere la evasión arbitraria de los privilegios de los individuos o de ciertos grupos. En cambio, lo que es apropiado para cada individuo debe determinarse de manera racional y relacional. Esto presupone que se utilicen criterios uniformes y que las diferencias de trato entre las personas se justifiquen de manera consensuada (traducción nuestra).²³

Esa regulación es necesaria para evitar que tengamos sorpresas negativas en situaciones reales de uso de los algoritmos. Como los algoritmos trabajan, como ya se ha dicho, con probabilidades, hay que tener muchísimo cuidado al aplicarlos en situaciones de nuestro contexto social. No deberíamos fiarnos totalmente de los resultados presentados por ellos justamente por no se tratar de certezas absolutas.

Algunas situaciones concretas de una conducta éticamente incorrecta ya pueden ser observadas en los días actuales y la tendencia es que sigan expandiéndose exponencialmente. Un clásico ejemplo que ya viene ocurriendo en muchas situaciones, es el empleo de los algoritmos en contextos de procesos de selección para vacantes laborales. Cruzando datos disponibles en la red – muchas veces fornecidos voluntariamente por la propia persona, sin saber que podrán ser utilizados para otro fin futuramente –, las empresas en busca del candidato ideal pueden seleccionar o excluir candidatos por su perfil en cuanto consumidor, por su crédito bancario o perfil financiero, por sus relaciones personales,

²¹ IASON, Gabriel. Artificial Intelligence, Values, and Alignment. Disponível en: <https://arxiv.org/pdf/2001.09768.pdf>. Acesso: 17/07/2020

²² IASON, Gabriel. Artificial Intelligence, Values, and Alignment. Disponível en: <https://arxiv.org/pdf/2001.09768.pdf>. Acesso: 17/07/2020

²³ MOLINARO, C. A.; SARLET, G. B. Questões tecnológicas, éticas e normativas de dados pessoais na área da saúde em um contexto de big data. *Direitos Fundamentais & Justiça*, Belo Horizonte, a. 13, n. 41, p. 183-212, jul./dez. 2019.

opiniones políticas y hasta por su histórico de buena salud, sin que la persona conozca los criterios usados por los algoritmos.

Otro ejemplo es el uso indebido de los datos personales contenidos en el *Big Data*, con objetivos políticos de individuos o empresas, como ocurrió en el caso emblemático de Cambridge Analytica, tras valerse de la masiva utilización de datos personales de usuarios de redes sociales - sin ningún tipo de observancia a las normas sobre el uso de los datos u observancia de los principios éticos - con el intuito de cumplir con la estrategia política en una campaña de votación, llevando a la quiebra de la empresa cuando el escándalo vino a público. Ese caso fue ampliamente analizado en el artículo Cambridge Analytica: Ethics and Online Manipulation with Decision-Making Process, donde los autores demostraron que "Cambridge Analytica promueve su actividad como una combinación de análisis predictivo, ciencia del comportamiento y tecnologías publicitarias basadas en la recopilación de datos. El objetivo de esta combinación es formar un público objetivo y utilizar los datos para persuadir y motivar a los votantes."²⁴

Algunos ejemplos europeos de instituciones que buscan soluciones en este sentido son el *Algorithm Watch*, con su sede en Berlín y que se dispone a analizar la ética de los procesos decisorios de los algoritmos, y *The Ethical Tech Society*, que discute la importancia y los reflejos de los sistemas autónomos en la sociedad. Ambos tuvieron la participación de la filósofa Lorena Jaume-Palasi en su creación, que pretende hacer que el debate ético se base en datos concretos y no en suposiciones infundadas, que generan un miedo impuesto a la sociedad a la hora de discutir tecnología y el crecimiento de la influencia tecnológica en la sociedad.

Las referidas instituciones se consolidan con el intuito de discutir sobre el tema y monitorizar los algoritmos, tratando de verlos de una manera más

holística, considerando puntos de vista distintos que muchas veces no son considerados en el contexto tradicional. Si fuera posible reunir sociólogos, filósofos, psicólogos, abogados y profesionales de otros ramos del Derecho, antropólogos, ingenieros, estadísticos, etc., tendríamos la oportunidad de prever y corregir más fácilmente los posibles desvíos de los algoritmos. Desafortunadamente, equipos tan heterogéneos no suelen reunirse normalmente en los procesos de diseño y desarrollo de tecnologías.

En materia de sanidad, se destaca un artículo del Dr. Dalton S. Char y sus colaboradores de la *Stanford University School of Medicine* sobre los retos éticos que la sanidad puede presentar frente al Machine Learning, donde el algoritmo aprenderá, con el tiempo, de sus propias decisiones. Según los investigadores de Stanford, citados por Ángel Guerra y Ricardo García-Mayor, los principales retos son, básicamente, los malos diagnósticos presentados por la IA, debido a la falta o a la incorrección de los datos introducidos, una vez que la utilización de una gran cantidad de datos no resulta, necesariamente, en mejores diagnósticos, y el riesgo de romperse la confidencialidad de la relación médico-paciente.²⁵

A respecto de las limitaciones éticas al desarrollo de la IA, con el hecho de que las máquinas muy probablemente seguirán procesos de socialización y culturización diferentes a los humanos, Ramón López de Mántaras considera que, posiblemente:

La lección más importante que hemos aprendido a lo largo de los 60 años de existencia de la IA es que lo que parecía más difícil (por ejemplo, diagnosticar enfermedades o jugar al ajedrez a nivel de gran maestro), ha resultado relativamente fácil, y lo que parecía más fácil, ha resultado ser lo más difícil. (...) En particular, necesitamos lenguajes de representación de conocimientos que codifiquen información acerca de muchos tipos distintos de objetos, situaciones, acciones, etc., así como de sus propiedades y de las relaciones entre ellos. También necesitamos nuevos algoritmos que, en base a estas representaciones, puedan

²⁴ BOLDYREVA, E., GRISHINA, N., DUISEMBINAL, Y. Cambridge Analytica: Ethics and Online Manipulation with Decision-Making Process. *Future Academy*, 2018.

²⁵ GUERRA, A.; GARCÍA-MAYOR, R. Retos éticos que plantea el uso de la inteligencia artificial en el diagnóstico y tratamiento clínico. *Cuadernos de Bioética*, v. 29, n. 97, p. 303-304, 2018. Disponible en: <http://aebioetica.org/revistas/2018/29/97/303.pdf>. Fecha de consulta: 24 jun. 2020. p. 303.

responder de forma robusta y eficiente a preguntas sobre prácticamente cualquier tema. Finalmente, dado que tendrán que conocer un número prácticamente ilimitado de cosas, estos sistemas deberán ser capaces de aprender nuevos conocimientos de forma continua a lo largo de toda su existencia. En definitiva, además de progresos individuales en cada una de estas áreas, debemos también diseñar sistemas que integren percepción, representación, razonamiento, acción y aprendizaje.²⁶

Uno de los principales problemas al que se enfrenta la IA, si no el mayor, es la adquisición de sentido común, algo que no es fácil aprender desde libros, por ser una capacidad humana que resulta de la convivencia en sociedad, de la experiencia de vida, de los aprendizajes que perpetúan entre generaciones. Una salida propuesta para resolver ese problema sería la llamada "cognición situada" que, a través de la utilización de entornos reales, podrá capacitar a las máquinas a entender y aplicar el sentido común, ya que la máquina podrá vivenciar y experimentar situaciones que les ayuden en esta tarea. Para desarrollar esta actividad mental tan típica de los humanos, será necesario un cuerpo para la máquina poder interactuar en un entorno real de la manera más fiel posible, ya que gran parte nuestra inteligencia viene no solamente de nuestras habilidades mentales, pero también de las acciones concretas, del entorno donde se desarrollan estas acciones y de la interacción del agente con este entorno, como propone Randall D. Beer de forma bastante amplia en su *framework* SED (*situated, embodied, dynamical*), diciendo que "Este marco, que, a falta de una etiqueta más pegadiza, llamaré el marco dinámico situado, incorporado y dinámico (SED), se centra en la acción concreta y enfatiza la forma en que el comportamiento de un agente surge de la interacción dinámica entre su cerebro, su cuerpo y su ambiente."²⁷ (traducción nuestra)

En el año 2017, los aspectos científicos y éticos llevaron a un debate entre diversos expertos

europeos en Inteligencia Artificial y otras áreas afines, lo que hizo surgir la Declaración de Barcelona para un desarrollo y uso adecuados de la IA en Europa, y define seis principales hitos, conforme explica Ramón López de Mántaras:

El primero es la prudencia, la necesidad de ser conscientes de que todavía queda por resolver un gran número de obstáculos científicos y técnicos, en particular el problema del sentido común. El segundo hace referencia a la fiabilidad, esto es, que los sistemas de inteligencia artificial deben someterse a pruebas que determinen su fiabilidad y seguridad. Otro aspecto es la rendición de cuentas: cuando un sistema toma decisiones, las personas afectadas por ellas tienen que poder recibir, en unos términos de lenguaje que entiendan, una explicación de por qué las ha tomado, y tienen que poder cuestionarlas con argumentos razonados. Se apunta también al criterio de responsabilidad. Debe quedar claro si la interacción se hace con una persona o con un sistema de inteligencia artificial, y, en el segundo caso, debe poderse localizar e identificar a los responsables de él. Otra cuestión es la autonomía limitada de estos sistemas. Se necesita disponer de reglas claras que limiten el comportamiento de los sistemas de inteligencia artificial autónomos para que los encargados de desarrollarlos puedan incorporarlos en sus aplicaciones. Y, por último, debe quedar claro el papel que desempeña el ser humano. En casi cualquier área, la capacidad humana todavía supera con creces la inteligencia artificial, especialmente en el tratamiento de casos que no han aparecido en los conjuntos de datos de ejemplo de los que aprenden los sistemas de inteligencia artificial.²⁸

Una de las respuestas ha sido la Resolución del Parlamento Europeo, de 16 de febrero de 2017, que contempla normas del Derecho Civil en materia de Robótica, destinada a la Comisión Europea, y simboliza una de las tentativas de reglamentar esta área que no para de crecer.

Además de la Resolución, también levanta la necesidad de que una Agencia Europea - Reguladora o Asesora - sea creada para promover conocimientos normativos, técnicos y éticos, manteniendo el Parlamento actualizado sobre los principales cambios que resultan con las nuevas tecnologías y las expectativas de qué

²⁶ LÓPEZ DE MÁNTARAS, R. Algunas reflexiones sobre el presente y el futuro de la Inteligencia Artificial. In: CORTINA, A.; SERRA, M. A. (Coord.). *Desafíos éticos de las tecnologías emergentes*. Madrid: EIUNSA, 2016. p. 89-103.

²⁷ BEER, R. (2014). *Dynamical systems and embedded cognition*. In K. Frankish & W. Ramsey (Eds.), *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press. p. 128.

²⁸ LÓPEZ DE MÁNTARAS BADIA, R. Ética en la inteligencia artificial. *Investigación y Ciencia*, Barcelona, 2017. Disponible en <<https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/el-multiverso-cuntico-711/tica-en-la-inteligencia-artificial-15492>>. Acceso en: 15/07/2020.

novedades surgirán en un futuro próximo. Y, como la ética es uno de los ámbitos en que se resulta fundamental enfocar con la mayor brevedad posible, la Resolución del Parlamento Europeo ha adjuntado un Código de Conducta Ética que sirva de base para los ingenieros de Inteligencia Artificial y Robótica.

5 Conclusión

La revolución tecnológica que ya estamos viviendo con la Inteligencia Artificial y todo el nuevo mundo que implica la llegada de esta y otras tantas nuevas tecnologías, será un marco sin precedentes en la civilización humana. Al hombre, le queda adaptarse a esta nueva realidad, unirse y discutir ideas a respecto de la IA y tecnologías derivadas. Una reacción en los moldes del Movimiento Ludista - que ha intentado frenar la Revolución Industrial con el uso de la violencia - sería infructífero en los tiempos actuales, puesto que es imposible un retroceso a una situación anterior, ya que, en la realidad, no se puede pretender frenar los cambios que llegan con la sociedad de la información. Al revés, cuanto antes la sociedad acepte que los cambios seguirán a pasos largos, mejor, ya que entenderá la extrema importancia de actuar el cuanto antes en los puntos relacionados a la ética, una vez que la conducta individual y colectiva será la que dictará los resultados que vamos a lograr en cuanto sociedad.

Sin embargo, eso no significa aceptar el cambio de cualquier manera, sin control ni debate de ideas. Todo lo contrario, significa asumir responsabilidades para que este periodo de adaptación sea vivenciado de la manera más positiva posible. Hay que estudiar, reciclarse, modernizarse, llevar la temática para la política, para los medios académicos, empresariales y de comunicación, pues solamente así será posible alcanzar las ventajas que la tecnología puede entregar, mitigando los riesgos que conlleva. La manera por la cual encaramos la nueva realidad será directamente proporcional al resultado que percibiremos.

Esta preocupación debe ser considerada por todos los actores sociales involucrados: en cuanto

empresa, no debe olvidarse de la responsabilidad social que deben tener sus administradores; en cuanto ingenieros informáticos y robóticos, para tener consciencia del compromiso con la creación; en cuanto usuarios, para tener en cuenta la importancia del uso responsable y consciente; en cuanto al Estado, para fomentar la discusión de los valores morales y éticos a fin de crearse un marco legislativo en que la ética sea uno de los principios fundamentales; y, en cuanto sociedad, para no perder de vista los valores centrales de la humanidad.

Desde un punto de vista filosófico, uno de los principales desafíos será el de asegurar la aplicación de los valores humanos y éticos en la IA, no dejando crearse la sensación, cada vez más presente en la sociedad, de que las máquinas son mejores y más inteligentes que el ser humano, pero sí que son un complemento que puede ser un activo valioso a la sociedad y no un enemigo que ha venido a destruirla, como sugiere las películas de ficción científica.

Antes de más nada, tenemos que recordar que la máquina ha sido creada por los seres humanos, y el significado de esto ya nos dice mucho sobre el poderío del hombre. La criatura, éticamente, debe respecto al creador, pero cabe al creador imponer y exigir que la jerarquía, por así decir, sea respetada. La Inteligencia Artificial debe de ser complementar al hombre y no visar sustituirlo.

La ética, por lo tanto, debe ser pensada tanto en el contexto más extenso del universo de los creadores de las máquinas, como en su uso más rutinario. Observar las formas de aplicación de la tecnología en el día a día, la importancia que se le da y, sobre todo, la importancia que se quita de otras esferas humanas importantes, como las relaciones personales, ya debe ser considerado una excelente reflexión.

Bibliografía

BARRIO ANDRÉS, M. (dir.). *El derecho de los robots*. Madrid: La Ley, 2018.

BARRIO ANDRÉS, M. La respuesta del derecho ante los robots y la inteligencia artificial. *El País*, Retina, Madrid, 11 mar. 2018. Disponible en: https://retina.elpais.com/retina/2018/03/07/tendencias/1520428359_555787.html. Fecha de consulta: 16 ago. 2019.

BEER, R. Dynamical systems and embedded cognition. In: FRANKISH, K.; RAMSEY, W. (ed.). *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, 2014. p. 128-148. <https://doi.org/10.1017/CBO9781139046855.009>

BOLDYREVA, E.; GRISHINA, N.; DUISEMBINAL, Y. Cambridge Analytica: Ethics and Online Manipulation with Decision-Making Process. *Future Academy*, 2018. Disponible en https://www.researchgate.net/publication/330032180_Cambridge_Analytica_Ethics_And_Online_Manipulation_With_Decision-Making_Process. Fecha de consulta: 08 ago. 2020. <https://doi.org/10.15405/epsbs.2018.12.02.10>

CALO, R. Robotics and the lessons of cyberlaw. *California Law Review*, v. 103, n. 513, 2015. Disponible en: <https://digitalcommons.law.uw.edu/faculty-articles/23/>. Fecha de consulta: 24 jun. 2020.

COMISSION EUROPEA. Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial. *Directrices éticas para una IA fiable*. Bruselas: CE, 2019.

DONEDA, D. *Da privacidade à proteção de dados pessoais: fundamentos da lei geral de proteção de dados*. São Paulo: Revista dos Tribunais, 2019.

EUROPEAN COMMISSION. European Group on Ethics in Science and New Technologies. *Statement of Artificial Intelligence, Robotics, and 'Autonomous' Systems*. Brussels, 9 Mar. 2018. Disponible en: http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf. Fecha de consulta: 24 jun. 2020.

GUERRA, A.; GARCÍA-MAYOR, R. Retos éticos que plantea el uso de la inteligencia artificial en el diagnóstico y tratamiento clínico. *Cuadernos de Bioética*, [S. l.], v. 29, n. 97, p. 303-304, 2018. Disponible en: <http://aebioetica.org/revistas/2018/29/97/303.pdf>. Fecha de consulta: 24 jun. 2020.

IASON, Gabriel. *Artificial Intelligence, Values, and Alignment*. Disponible en: <https://arxiv.org/>. Acceso en: 17 ago. 2020.

LÓPEZ DE MÁNTARAS BADÍA, R. Ética en la inteligencia artificial. *Investigación y Ciencia*, Barcelona, 2017. Disponible en: <https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/el-multiverso-cuntico-711/tica-en-la-inteligencia-artificial-15492>. Fecha de consulta: 15 jul. 2019.

LÓPEZ DE MÁNTARAS, R. Algunas reflexiones sobre el presente y el futuro de la Inteligencia Artificial. In: CORTINA, A.; SERRA, M. A. (coord.). *Desafíos éticos de las tecnologías emergentes*. Madrid: EIUNSA, 2016.

MAGRANI, E.; SILVA, P.; VIOLA, R. Novas perspectivas sobre ética e responsabilidade de inteligência artificial. In: FRAZÃO, A.; MULHOLLAND, C. *Inteligência artificial e direito: ética, regulação e responsabilidade*. São Paulo: Revista dos Tribunais, 2019.

MOLINARO, C. A.; SARLET, G. B. Questões tecnológicas, éticas e normativas de dados pessoais na área da saúde em um contexto de big data. *Direitos Fundamentais & Justiça*, Belo Horizonte, a. 13, n. 41, p. 183-212, jul./dez. 2019. <https://doi.org/10.30899/dfi.v13i41.811>

O'KEEFE, C.; FLYNN, C.; CIHON, P.; LEUNG, J.; GARFINKEL, B.; DAFOE, A. The Windfall Clause. Distributing the Benefits of AI for the Common Good, UK, 2020. Disponible en <https://arxiv.org/ftp/arxiv/papers/1912/1912.11595.pdf>. Fecha de consulta: 09 ago. 2020. <https://doi.org/10.1145/3375627.3375842>

PETIT, M. Por una crítica de la razón algorítmica: estado de la cuestión sobre la inteligencia artificial, su influencia en la política y su regulación. *Quaderns del CAC*, Barcelona, n. 44, p. 5-15, 2018.

PIÑAR MAÑAS, J. L. Derecho, técnica e innovación en las llamadas ciudades inteligentes. Privacidad y gobierno abierto. In: RUARO, R.; PIÑAR MAÑAS, J. L.; MOLINARO, C. A. *Privacidade e Proteção de Dados Pessoais na Sociedade Digital*. Recife: Editora Fi, 2017.

SOARES, N. The Value Learning Problem. *Machine Intelligence Research Institute*, Berkeley, 2016. Disponible en: <https://intelligence.org/files/obsolete/ValueLearningProblem.pdf>. Fecha de consulta: 09 ago. 2020.

WEBER, Thadeu. *Ética e filosofia do direito: autonomia e dignidade da pessoa humana*. Petrópolis: Vozes, 2013.

YUDKOWSKY, E. Complex Value Systems are Required to Realize Valuable Futures. *Machine Intelligence Research Institute*, Berkeley, 2011. Disponible en: <https://intelligence.org/files/ComplexValues.pdf>. Fecha de consulta: 12 ago. 2020.

Regina Linden Ruaro

Doctora en Derecho por la Universidad Complutense de Madrid/España, con posdoctorado en la Universidad San Pablo de Madrid/CEU/España. Investigadora del Grupo Internacional Privacidad y Acceso a la Información Universidad San Pablo de Madrid/CEU/España. Profesora Titular y Decana Asociada de la Escuela de Derecho/PUCRS.

Ludimila Camilo Catão Guimarães Reis

Máster en Protección de Datos de Carácter Personal, Transparencia y Acceso a la Información en la Universidad San Pablo de Madrid/CEU/España. Abogada en Telefónica S.A.

Dirección

Regina Linden Ruaro

Pontifícia Universidade Católica do Rio Grande do Sul
Av. Ipiranga, 6681

Partenon, 90619-900

Porto Alegre, RS, Brasil

Ludimila Camilo Catão Guimarães Reis

Universidad San Pablo de Madrid

Calle de Julián Romea, 23, 28003

Madrid, Mad, Espanha