

Proyecto Ética IA

“Riesgos, Dilemas Éticos y Buenas prácticas en Inteligencia Artificial”

Directora: Dra. Bibiana Luz Clara; Co Directora: Mg. Lucía Rosario Malbernat. Investigadores: Alexis Antoniucci

Período: 2021-2023.

Fecha de Finalización: 31/03/2023

Materias en las que impactó: Aspectos Profesionales para Sistemas; Aspectos Profesionales para Ingeniería; Inteligencia de Negocios

Informe de cierre

Introducción

Si bien existen normas jurídicas que tipifican los delitos informáticos, protegen datos o algoritmos y múltiples regulaciones locales, es necesario conocer la normativa que regula el desarrollo y utilización de actividades propias de la IA así como también las normas técnicas que configuran estándares para el mercado.

Las normas ISO sobre inteligencia artificial la define de forma un poco vaga, con la finalidad de no limitarse, como: “la capacidad de una unidad funcional para realizar funciones que generalmente están asociadas con la inteligencia humana, como el razonamiento y el aprendizaje.” En el ámbito de la Unión Europea, se propone el cumplimiento de estos objetivos a través de la aplicación de normas internacionales, de la creación de un código de Conducta, de la revisión de las reglas de responsabilidad vigentes y de la evaluación del impacto social que la IA pueda llegar a tener. Se propone, como una de las primeras actividades del Proyecto, relevar las normativas vigentes en la región, que se vinculen con cualquier etapa de la producción y uso de aplicaciones.

Dentro de las ciencias de la computación, la IA constituye una categoría cuya finalidad es diseñar sistemas inteligentes y autónomos, basados en el conocimiento humano. El concepto de IA se aplica a los sistemas que manifiestan un comportamiento inteligente, capaces de analizar su entorno y pasar a la acción con cierto grado de autonomía, con el fin de alcanzar objetivos específicos. Se propone como otro de los objetivos del proyecto y primeras actividades, conocer el panorama de desarrollo y aplicación de IA en las organizaciones de la región, identificando quienes llevan a cabo el desarrollo, con qué fines y contraparte, quienes son los usuarios externos o internos, etc.

La incorporación de Inteligencia Artificial a los procesos organizacionales impulsa a dotar de valores y principios a la aplicación tecnológica y a comprender las implicaciones morales y éticas que conllevan. En este sentido, se plantea, también, que es necesario, además, relevar las preocupaciones de las organizaciones sobre principios éticos aplicables a la Inteligencia Artificial.

Los riesgos e impactos que pueden derivar de la utilización de sistemas provistos de IA llevan a preguntarse por la responsabilidad civil contractual y extracontractual que debe considerarse dentro de diversos ámbitos y, en ese sentido, se estableció otro objetivo con el fin de conocer los resguardos que toman las empresas para evitar que la IA sea utilizada con fines ilícitos o

contra la ética en todas las etapas de los procesos vinculados a su utilización y de tipificar tales riesgos.

En particular, los riesgos que aparecen ante la falta de explicabilidad o transparencia ya que la IA trae consigo riesgos que pueden derivarse de defectos, errores o sesgos en la programación que atenten contra la privacidad, intimidad, dignidad, salud o patrimonio de las personas y la responsabilidad eventualmente podría estar distribuida entre distintos actores haciendo más complejo el problema. En este aspecto surge otro objetivo tendiente a averiguar si las organizaciones que producen las aplicaciones y programan los algoritmos, toman, también, recaudos, ya sean las mismas u otras distintas a las que las aplican.

Implementación de la IA

La Inteligencia Artificial es una disciplina relacionada con la teoría de la computación cuyo objetivo es emular algunas facultades humanas en sistemas artificiales; sin embargo, hoy en día el principal objetivo de la inteligencia es el tratamiento y análisis de datos, siendo las aplicaciones más frecuentes, la robótica, el análisis de imágenes y el tratamiento automático de textos. (Benitez et al, 2013).

Para Russell y Norvig (2004), a lo largo de la historia de la IA se han seguido cuatro enfoques, dos ellos se refieren a procesos mentales y al razonamiento, mientras que los otros 2 aluden a la conducta. Los primeros son sistemas que piensan como humanos y sistemas que piensan racionalmente y los dos restantes, sistemas que actúan como humanos y sistemas que actúan racionalmente.

Los sistemas que piensan como humanos son capaces de tomar decisiones, de aprender y de resolver problemas imitando el comportamiento del sistema nervioso de un humano gracias a las Redes Neuronales Artificiales.

Los sistemas que piensan racionalmente son los sistemas expertos, tecnología capaz de razonar en situaciones como lo haría un ser humano. Los sistemas que actúan como humanos son sistemas artificiales capaces de repetir tareas emulando al hombre, como es el caso de los Robots, tecnología que actúa como persona.

Por último, los sistemas que actúan racionalmente son los que integran agentes inteligentes o racionales capaces de analizar patrones. Se trata de tecnología artificial que permite actuar como lo haría una persona.

Para reflexionar sobre el impacto que la Inteligencia Artificial tendrá en la sociedad (Rouhiainen, 2018) puede observarse que la IA ha comenzado a desarrollar la capacidad de ver (visión artificial), de oír (reconocimiento de voz) y de comprender (procesamiento natural del lenguaje) y que el aprendizaje automático o machine learning es uno de sus principales enfoques por el que las computadoras adquieren la capacidad de aprender sin estar programados para ello, utilizando algoritmos para aprender de los patrones de datos a través de un aprendizaje supervisado (basado en tareas), un aprendizaje no supervisado (basado en datos) o a través de un aprendizaje de refuerzo, es decir, aprendiendo a reaccionar al entorno o en base a la experiencia.

Un caso particular de machine learning, es el Deep Learning, que usa una red neuronal que se compone de un número de niveles jerárquicos.

En el nivel inicial de la jerarquía la red aprende algo y lo envía al siguiente nivel donde se combina con alguna información más compleja y así sucesivamente. Deep Learning lleva a cabo un proceso de aprendizaje que puede aplicarse a grandes volúmenes de datos para el

descubrimiento y aplicación de conocimientos así como a la realización de predicciones a partir de él.

En ese contexto, la implementación de la IA y otras tecnologías digitales, debe considerar:

- i) que esta sea alimentada con datos confiables y accesibles;
- ii) fomentar el diálogo y la cooperación entre gobierno- industria-comercio y ciencia para favorecer la transferencia tecnológica;
- iii) facilitar la medición y evaluación de la consecución de los objetivos de desarrollo sostenible (ODS) propuestos;
- iv) planes de acción coordinados y adaptados a cada región, ya que cada una tiene sus características propias, y no todas se encuentran en las mismas condiciones para adaptarse a la tecnología digital que les permitan afrontar proyectos para lograr los ODS.

Sostenibilidad de la IA

La sostenibilidad está caracterizada por tres dimensiones que deben ser tenidas en cuenta: la dimensión económica; la dimensión social; y dimensión medioambiental. Los grandes datos y de calidad, unificados, accesibles y abiertos, sobre cada una de ellas puede aportar mucho a una transformación digital que sea adecuada para la preservación de los recursos naturales y el desarrollo en las distintas áreas. De este modo se podrá avanzar en proyectos basados en tecnología que respeten los objetivos de desarrollo sostenible (ODS) propuestos, dentro de un marco ético.

Para ello, empresas y gobiernos deberán aportar y compartir datos que lleven al diseño de soluciones acordes a las metas propuestas. La IA y las tecnologías digitales asociadas a la conectividad digital, 5G, cloud computing, IoT, Smart cities, Smart regions, big data, gemelo digital, Edge computing, son herramientas necesarias hoy para lograr el progreso en la sociedad digital, pero debe encararse con alta moral, responsabilidad y ética.

Para hacer frente a las nuevas problemáticas éticas y de seguridad se viene trabajando en la definición de estrategias, planes de desarrollo y buenas prácticas basadas en la capacitación de las personas y en la definición de normas con el fin de enfrentar a los peligros inminentes de vigilancia masiva, de amenazas a la privacidad y dignidad de las personas, al aumento de prejuicios de género o étnicos, ciberviolencia y ciberacoso, etc.

Normas, estándares y buenas prácticas en IA

La Unesco, en este sentido, ha propuesto recomendaciones enfocándose en las ventajas que aporta la Inteligencia Artificial a la sociedad y en la reducción de sus riesgos enfocándose en la necesidad de garantizar la protección de datos personales, la prohibición de marcadores sociales y de vigilancia masiva y la reducción del impacto medioambiental de los sistemas e infraestructuras de IA, lo cual no dista mucho de la propuesta de la Unión Europea que se orienta hacia la construcción de sistemas de inteligencia artificial fiables a lo largo de todo el ciclo de vida del sistema.

La Unión Europea, por su parte, ha redactado Directrices Éticas para una Inteligencia Artificial Fiable, el Banco Interamericano de Desarrollo (BID) se ha enfocado en el entendimiento generalizado de la inteligencia artificial, sus oportunidades, aplicaciones y riesgos, el Instituto de Ingenieros Eléctricos y Electrónicos, IEEE, tiene abierto un capítulo sobre inteligencia computacional (Computational Intelligence) al que redirecciona disertaciones, congresos y recursos, la Organización para la Cooperación y el Desarrollo

Económico, OCDE, en colaboración con el CAF, Banco de Desarrollo de América Latina, preparó un informe para entender las acciones estratégicas y capacidades para la inteligencia artificial en el sector público, identificar enfoques y acciones específicas para construir gobiernos eficientes, eficaces y receptivos e impulsar la colaboración en la búsqueda de una visión regional para la IA en el sector público, en tanto que la Red Iberoamericana de Protección de Datos elaboró documentos orientadores para el uso adecuado de datos personales en el diseño e implementación de Inteligencia Artificial.

De este modo, en torno a estas tecnologías, comienzan a aparecer las primeras comunidades y consejos asesores, a esbozarse los primeros proyectos de ley, planes estratégicos y de acción, políticas nacionales y regulaciones con distintos niveles de alcance que, entre otras cuestiones hacen hincapié en que los sistemas que utilizan inteligencia artificial deben ser transparentes y explicables. Para ello, por un lado, la lógica de sus algoritmos debe ser confiable, robusta y sin sesgos ideológicos, de género, étnicos o religiosos mientras que, por otro, las organizaciones que emplean tecnologías inteligentes deben garantizar la protección de los datos, identificar quienes prueban y usan sus sistemas, que datos se utilizan en las pruebas y de qué manera se realiza el entrenamiento de modelos. Es decir, el impacto alcanza personas, procedimientos, aplicaciones y datos.

Esta problemática es la que demanda el desarrollo de normas, la generación de buenas prácticas y la capacitación de todos los involucrados, independientemente del lugar que ocupen en la pirámide organizacional y del rol que ejerzan las cuales tendrán distintos niveles de alcance pero impactarán de lleno en el futuro del empleo.

Productos

Malbernat, LR (2023) *Impacto de los Aspectos Éticos de la Inteligencia Artificial en el Trabajo del Futuro*. En Global Overview Magazine. Marzo 2023. ISSN: 2618-1916

Luz Clara, B (2022) *Tecnologías Cuánticas en los Servicios de Justicia* XI Fórum de Expertos y Jóvenes Investigadores en Derecho y Nuevas Tecnologías, Fodertics. Universidad de Salamanca.

Malbernat, LR (2022) *Aspectos Éticos de la Inteligencia Artificial con Impacto en las Organizaciones Modernas. Nuevas Competencias*. En Global Overview Magazine. Diciembre 2022. ISSN: 2618-1916

Luz Clara, BB; Malbernat, LR (2021) *Riesgos, dilemas éticos y buenas prácticas en inteligencia artificial*. De XXIII Workshop de Investigadores en Ciencias de la Computación, WICC 2021, Chilecito, La Rioja. Red de Universidades con Carreras en Informática, REDUNCI, ISBN: 978-987-24611-3-3; 978-987-24611-4-0.

Luz Clara, BB; Malbernat, LR, Antoniucci, AG (2021) *Ética, Objetivos de Desarrollo Sostenible e Inteligencia Artificial*. En Memorias II Jornada de Ética en la Ingeniería Argentina. UCASAL, Salta, Argentina. ISBN 978-950-623-244-3.

Malbernat, LR. Luz Clara, B, Antoniucci, AG (2021). *Protocolos y Normas en Inteligencia Artificial*. Ciudad Autónoma de Buenos Aires: Universidad CAECE. De Investigar en Tiempos de Incertidumbre: Libro de Actas de las VII Jornadas de Investigación de UCAEC, Malbernat LR, Sanchez Negrete MG, Copelli, S. Compilación de Sanchez Negrete MG [et al.]. - 1a ed -. (2021). ISBN 978-987-46267-8-3