

Promesas y riesgos de la cuarta revolución industrial

Explorando el lado oscuro de los algoritmos inteligentes y las máquinas cognitivas

Por David Valle Cruz y José Ramón Gil García

La inteligencia artificial ha llamado la atención en todo el mundo y ha generado disrupción en importantes sectores, incluso se ha ubicado como una de las ciencias interdisciplinarias más importantes y con gran impacto en muchas áreas del conocimiento. La cuarta revolución industrial (inteligencia artificial) tiene el potencial de explorar de forma creativa, heurística y automatizada los problemas complejos del mundo, tales como el cambio climático, la pobreza y la corrupción. Los algoritmos inteligentes y las máquinas cognitivas —como núcleo del funcionamiento de la inteligencia artificial— son cada vez más omnipresentes en actividades que se consideran exclusivamente humanas, como la creatividad y la especialización profesional.

Aunque todavía falta mucho por hacer y entender en torno a la inteligencia artificial, actualmente tenemos ejemplos claros de su aplicación en diferentes sectores. Por mencionar algunos, existen “startups” que realizan desarrollos, con procesamiento de lenguaje natural para la revisión de contratos legales, previsión de litigios e investigación jurídica (Toews, 2019). También, científicos de la NASA desarrollan algoritmos de aprendizaje automático para identificar asteroides cercanos y descubrir vida en el universo (Shekhtman, 2019). Asimismo, la pandemia por COVID-19 fue un detonante para el diseño emergente de inteligencia artificial que contribuyera a disminuir el contagio y ayudara a los médicos en el diagnóstico; para lograr lo anterior, se construyeron modelos basados en aprendizaje automático para predecir la propagación del virus, así como para el reconocimiento de patrones en la detección de síntomas de COVID-19 y del virus que lo provoca (SARS-CoV-2) en radiografías de tórax (Latif et al., 2020).

La inteligencia artificial ha llegado a límites hasta ahora no imaginados. En este sentido, por ejemplo, la creatividad computacional estudia el desarrollo de software que presenta un comportamiento que sería considerado creativo en seres humanos. También existen

experimentos que buscan la forma de inventar teorías matemáticas, escribir poesía, pintar cuadros y componer música. Además, la creatividad computacional también permite comprender cómo funciona la creatividad humana (López de Mántaras, 2016). Todo esto puede sonar a temas solamente vistos en películas de ciencia ficción. Sin embargo, ha sido tanto el avance de la inteligencia artificial en los últimos años que existen problemas inherentes en su funcionamiento relacionados con la falta de transparencia, explicación y entendimiento.

A lo largo de la historia de la inteligencia artificial el debate científico sobre los algoritmos inteligentes y las máquinas cognitivas ha estado dominado por un enfoque casi puramente tecnológico, pero en el siglo XXI han comenzado a ser de interés también aspectos legales, éticos y morales. Debido a los efectos potenciales de la inteligencia artificial, es importante comprender mejor sus retos, beneficios y consecuencias no deseadas tanto desde una perspectiva académica como práctica. El lado oscuro de la inteligencia artificial se encuentra en el sesgo resultante de los algoritmos inte-



Tecnología

Fotomontajes: Gustavo Contreras



ligentes y del comportamiento inesperado de las máquinas cognitivas (Liu *et al.*, 2019), ya que el entrenamiento de los sistemas basados en inteligencia artificial depende de la calidad de los datos que analiza y que podrían contener sesgos implícitos de raza, género o ideología, entre otros. De tal forma que, en los patrones encontrados por los algoritmos de inteligencia artificial se corre el riesgo de presentar normas e interpretaciones sociales dominantes, más que resultados puramente técnicos y neutrales (Meijer *et al.*, 2021). Asimismo, esta cara poco favorecedora de la inteligencia artificial surge de la falta de comprensión de su funcionamiento en lo colectivo, pues los que la comprenden podrían dominar a los que no lo hacen, además existe la posibilidad de que tengan ventajas sobre los sectores más vulnerables, por esta razón existen riesgos potenciales en el crecimiento de las desigualdades, la injusticia, las brechas de inequidad en la sociedad y los fallos del mercado (Williams *et al.*, 2020).

Debido a que los responsables del diseño y administración de la inteligencia artificial, así como los programadores pueden manipular el comportamiento de los algoritmos inteligentes y las máquinas cog-

nitivas, resulta fundamental establecer marcos legales y éticos que brinden transparencia y rendición de cuentas del funcionamiento de la inteligencia artificial, protección de los datos personales críticos, así como la explicación sobre la forma en que la inteligencia artificial genera resultados para la toma de decisiones (algunos países a lo largo del mundo han comenzado a trabajar en esto). De este modo pueden evitarse, o al menos reducirse significativamente, los abusos y las injusticias sistemáticas al utilizar inteligencia artificial. Los involucrados en el diseño, implementación y uso estratégico de la inteligencia artificial deben ser conscientes de los beneficios potenciales,



pero también del lado oscuro para evitar las posibles consecuencias negativas. La inteligencia artificial puede proporcionar herramientas sofisticadas, precisas y altamente automatizadas para mejorar la toma de decisiones, pero rara vez deberá sustituir a la persona responsable de esa toma de decisiones, quien se enfrenta a un problema complejo que afectará a miles o incluso millones de personas.

En resumen, la mayor parte de los riesgos relacionados con el lado oscuro de la inteligencia artificial son heredados de valores, y sesgos humanos, no de las técnicas de inteligencia artificial en sí mismas. Por tanto, se requiere hacer un análisis minucioso de los datos y ser transparentes en cuanto a cómo funcionan los algoritmos que se van a usar, particularmente en el sector público. Además, por el momento, para la mayoría de las circunstancias, la decisión inicial y final —a pesar de los algoritmos inteligentes y las máquinas cognitivas— debe ser realizada por personas que tomen en cuenta las circunstancias específicas y tengan un conocimiento amplio sobre el tipo de sesgos que se tienen en los datos históricos. Sin embargo, no estamos seguros de lo que nos depara el futuro, que podría ser muy diferente. 🤖

Referencias

- Latif, Siddique. *et al.* (2020). "Leveraging data science to combat covid-19: A comprehensive review", en *IEEE Transactions on Artificial Intelligence*. *TechRxiv*. Preprint. <<https://doi.org/10.36227/techrxiv.12212516.v2>>.
- Liu, Han-Wei., Ching-Fu Lin y Yu-Jie Chen (2019). "Beyond State v Loomis: artificial intelligence, government algorithmization and accountability", en *International Journal of Law and Information Technology*, vol. 27, núm. 2, pp. 122–141.
- López de Mántaras, Ramón (2016). "La inteligencia artificial y las artes. Hacia una creatividad computacional", en *El próximo paso: la vida exponencial*. Turner.
- Meijer, Albert, Lukas Lorenz y Martijn Wessels (2021). "Algorithmization of Bureaucratic Organizations: Using a Practice Lens to Study How Context Shapes Predictive Policing Systems", en *Public Administration Review*.
- Shekhtman, Lonnie (2019). "NASA Takes a Cue from Silicon Valley to Hatch Artificial Intelligence Technologies", en *NASA Tv*. <https://www.nasa.gov/feature/god_dard/2019/nasa-takes-a-cue-from-silicon-valley-to-hatch-artificial-intelligence-technologies>.
- Toews, Rob (2019). "AI Will Transform The Field Of Law", en *Forbes*. <<https://www.forbes.com/sites/robtoews/2019/12/19/ai-will-transform-the-field-of-law/?sh=bb5a6467f01e>>.
- Williams, Arthur, Petr Suler y Jaromir Vrbka (2020). "Business Process Optimization, Cognitive Decision-Making Algorithms, and Artificial Intelligence Data-driven Internet of Things Systems in Sustainable Smart Manufacturing", en *Journal of Self-Governance and Management Economics*, vol. 8, núm. 4, pp. 39-48.



David Valle Cruz es profesor en la Universidad Autónoma del Estado de México e integrante del Sistema Nacional de Investigadores. Es ingeniero en Computación, maestro en Informática y doctor en Ciencias Económico-Administrativas. Su investigación ha sido publicada en importantes revistas internacionales, como *Government Information Quarterly*, *Cognitive Computation*, *First Monday*, *Information Polity* e *International Journal of Public Sector Management* (entre otras). Sus intereses de investigación están relacionados con la inteligencia artificial aplicada, las redes sociales digitales y las tecnologías emergentes en el sector público.



José Ramon Gil García es profesor asociado de Administración y Política Pública y director de investigación del Centro de Tecnología en el Gobierno de la Universidad de Albany, Universidad Estatal de Nueva York (SUNY). Miembro de la Academia Mexicana de Ciencias y del Sistema Nacional de Investigadores Nivel III. En el 2009, fue considerado el autor más prolífico en el campo de la investigación del gobierno digital en el mundo y en el 2013 fue seleccionado para el Premio a la Investigación, la más alta distinción que otorga anualmente la Academia Mexicana de Ciencias a jóvenes investigadores destacados. Fue denominado una de las 100 personas más influyentes del mundo en el gobierno digital en 2018 y 2019 por Apolitical, organización sin fines de lucro con sede en Londres, Reino Unido.