La inteligencia artificial es omnipresente. Hay que ser cuidadosos con que se aplique cuando y donde sea necesaria y teniendo en cuenta requisitos, condiciones y términos absolutamente imprescindibles para evitar consecuencias no deseadas.

Sin que se anunciara, sin que se pidiera, casi sin que se necesitara. Con sigilo, pero sobre todo de forma tan irreversible como imparable, la inteligencia artificial (IA) ha llegado a nuestras vidas y en la actualidad nuestros electrodomésticos, sistemas de navegación, automoción, recomendación o planificación personal llevan incorporadas tecnologías basadas en IA que poco a poco nos conducen a un nuevo modelo de sociedad, que presumimos tan dinámico como desconocido, que por un lado genera esperanza y por otro desconfianza. Esperanza, por una parte, porque la IA nos sirve para resolver muchos y muy importantes problemas, ya sean domésticos, profesionales y sociales, entre otros, pero también desconfianza porque tememos que la IA, la tecnología que se basa en ella, no sepa aprovechar la riqueza de la normal comunicación verbal entre personas, ni el conocimiento, estructurado o no estructurado, basado en la experiencia que poseemos, ni poner coto a la intimidad de las personas.

Siendo ambos ejes de reflexión válidos, hay un tercero que no conviene menospreciar: el realista, porque la IA ha venido para quedarse y es necesario conocer cuanto más se pueda sus realidades, sus problemas, sus progresos y sus límites. En definitiva, saber que puede y que no puede hacer. Se trata de un desafío muy importante en el que debemos ir de la mano de la ciencia y no de creencias infundadas, tan frecuentes a veces en algunos medios, que recuerdan a los místicos: «Para venir a donde no sabes has de ir por donde no sabes», pero que no aportan demostración alguna de sus previsiones ni de sus supuestos.

Partamos de que la IA es ante todo racional, consiguiendo esa racionalidad mediante la percepción del entorno en el que se esté aplicando a través de datos y sensores, razonando sobre lo percibido, decidiendo cuál es la mejor acción a tomar y finalmente actuando en consecuencia. Por tanto, ni es necesariamente aleatoria, aunque a veces pueda incorporar elementos de dicha naturaleza, ni cuál pócima mágica se comporta como el bálsamo de Fierabrás[1]. Lo mismo que no podemos entender la naturaleza sin las matemáticas, no podemos analizar la IA sin recurrir a ellas, entre otras cosas porque son las matemáticas las que nos proporciona las herramientas necesarias para poder explicar el funcionamiento de los sistemas basados en IA y, consiguientemente, emitir juicios científicamente justificados sobre la validez, corrección y viabilidad de sus resultados.

La IA es ante todo racional

Sobre esa base teórica, científica y bien fundada, para el aprovechamiento de todo lo que puede aportar a la sociedad la IA, se exige un proceso previo de digitalización y uno, quizás más específico, de transformación digital de cada ecosistema concreto. El primero se refiere a crear o mejorar nuestras vidas sacando ventajas de las tecnologías digitales y los datos previamente digitalizados. El segundo requiere realizar cambios en la organización del entorno social considerado, ya sea judicial, sanitario, industrial o cualquier otro, rentabilizando al máximo las posibilidades de la tecnología.

Una vez alcanzado cierto nivel de madurez de la transición digital de un ecosistema, la incorporación de tecnologías basadas en IA al normal funcionamiento del mismo ayudará a superar muchos de los desafíos que haya planteados en dicho contexto, a pesar de posibles incrementos iniciales de los costos, de la escasez de partida de profesionales capacitados en esas nuevas tecnologías, de una más que probable destrucción de

empleo a corto plazo o de las intangibles, pero reales expectativas no satisfechas de manera inmediata. En cualquier caso, el desarrollo de estrategias de IA en cada posible ecosistema social plantea grandes desafíos económicos, técnicos, éticos, políticos, regulatorios y de recursos humanos a los que hay que sumar el no tener garantías acerca del buen desempeño del sistema que estemos considerando, lo que a veces se produce por falta de capacitación de los usuarios finales y otras por fallos de diseño.

De hecho, del mismo modo que se abusa del término IA cuando en realidad está ausente, también se abusa de ella cuando para resolver un problema recurrimos a sistemas informáticos convencionales, o dotados de IA pobremente fundamentada, asumiendo que por realizar cierta tarea de manera digital se está aplicando IA de modo riguroso. Mientras que el primer mal uso podemos entenderlo en cierto sentido como una forma de fraude, el segundo lo podemos explicar, aunque no justificar, desde el punto de vista del desconocimiento de quién así actúa que, sin excusas, supone mala intención.

Como consecuencia, no teniendo una idea clara de lo que es IA por diversas razones que sería prolijo enumerar, el usuario final asume confiado que el sistema que está manejando funciona con IA, cuando en realidad no es así, dando lugar a problemas que con razón generan el rechazo a las soluciones que se ofrecen. Ese es el caso de la existencia de sesgos en el desarrollo de aplicaciones de la IA. A partir de errores en el proceso de construcción de algoritmos que pueden dar lugar a prácticas discriminatorias que se imputan a la IA, cuando la responsabilidad tiene que recaer en quien promueve un mal uso de ella.

Podemos identificar tres maneras en las que la IA se encuentra relacionada con sesgos algorítmicos y discriminatorios según que los sesgos:

1. Estén en los conjuntos de datos. Estos sesgos son prácticamente indetectables de antemano y para evitarlos no queda más remedio que recurrir a pruebas previas a la puesta en práctica del sistema que estemos considerando.
2. Se deban a un mal diseño, a un mal desarrollo o al uso de conjuntos de datos inadecuados para el correcto entrenamiento de la IA.
3. Se produzcan por no disponer de sistemas basados en IA del mismo nivel en todos los ámbitos geográficos y sociales, no contribuyendo a la eliminación de brechas sociales sino, muy al contrario, a su ensanchamiento.

Pero del mismo modo que intencionada e injustificadamente se hace a la IA responsable de los sesgos que puedan presentar las soluciones, también se la suele culpabilizar de diferentes perjuicios sociales, laborales o económicos que se profetiza llegarán a la sociedad para, supuestamente, desviar la atención del verdadero origen de los mismos. Por ejemplo, en los países de menores rentas, tanto la escasez de recursos como la falta de marcos de financiación estables son un obstáculo para la adquisición y mantenimiento de sistemas basados en IA fiables y eficientes; a lo que también hay que sumar la complejidad de los procesos de contratación pública y compra de tecnología con IA, que en muchos casos son inexistentes. En estos casos la IA puede llegar a fallar ya que podría producir efectos indeseados, como identificaciones equivocadas, diagnósticos erróneos, destrucción de empleo, imputaciones falsas o cosas peores. De esos casos particulares no se puede extraer la conclusión general de que la IA funciona mal siempre, porque sencillamente no es verdad. Recurrir al síndrome del algoritmo (“la culpa es del algoritmo que falla”) no es más que una prueba de que o no se quiere explicar la verdadera razón de un mal funcionamiento, que puede haber sido intencionado, o de que realmente por falta de formación de los diseñadores no se sabe explicar la causa de los errores.

La responsabilidad tiene que recaer en quien promueve un mal uso de la IA

Son problemas tan graves y trascendentes que tenemos que abordarlos de forma decidida y urgente. Para ello, lo que parece absolutamente necesario es disponer de entes reguladores y sancionadores, de agencias locales, regionales o nacionales que actúen con plena independencia de las distintas administraciones, de forma objetiva, transparente e imparcial, para que procuren la minimización de riesgos significativos sobre la seguridad y salud de las personas que puedan derivarse del uso de sistemas de IA. Dichos entes deberían de encargarse de la supervisión y seguimiento de los productos que de alguna manera se comercializan en su ámbito de atención con el sello IA y que, sean cuales sean sus entornos geográficos, actúen coordinadamente conforme al esquema del futuro European Artificial Intelligence Board, ajustándose a la legislación europea, ya disponible en forma de borrador de reglamento del Parlamento y del Consejo Europeo. Ahí se establecen normas armonizadas en materia de IA (la conocida Ley de la IA) que establecen una métrica para evaluar el impacto social de los algoritmos en el sector industrial, exigir la transparencia algorítmica y su explicabilidad y acreditar su calidad ética.

Dado que la convivencia con la IA conllevará una automatización de trabajos y tareas que transformará la economía y el actual modelo social en los próximos 15 o 20 años, el que esa transformación sea suave y cause los mínimos perjuicios posibles solo nos concierne a nosotros, que ante todo deberíamos ocuparnos y preocuparnos de capacitar a las personas de cualquier sector social para que pudieran alcanzar el máximo nivel formativo posible en sus respectivos entornos socio-laborales.

Como dijo Federico García Lorca (1931): «Bien está que todos los hombres coman, pero que todos los hombres sepan. Que gocen todos los frutos del espíritu humano porque lo contrario es convertirlos en máquinas al servicio del Estado, es convertirlos en esclavos de una terrible organización social».