

Conferencia de Angel González-Ferrer, Director Ejecutivo del Centro de Cultura Digital del Consejo Pontificio para la Cultura: «El dilema y los retos del desarrollo tecnológico: ¿humanización de la máquina o robotización del hombre?»»

ÁNGEL GONZÁLEZ FERRER

La Inteligencia Artificial (IA) está cada vez más presente en nuestras vidas y en nuestra sociedad pudiendo ciertamente decir que cada vez sabemos menos dónde acaba el hombre y empieza la máquina/robot o viceversa. Por ello se hace más necesario conocer el impacto de estas nuevas tecnologías, pues sólo así estaremos en mejor disposición para entender las consecuencias que pueden tener y, por tanto, tomar la decisión más conveniente.

Es importante tener unos conocimientos mínimos sobre ellas para facilitar su comprensión, por lo que la ponencia se ha dividido en cuatro partes: (i) el concepto de IA, (ii) la evolución del desarrollo tecnológico, (iii) los nuevos retos de la IA; oportunidades y riesgos, y (iv) ¿qué nos deparará el futuro?. Durante la exposición utilizamos indistintamente el término robot o máquina en contraposición al ser humano.

(i) El concepto de la inteligencia artificial (IA)

La IA está presente en todos los ámbitos de nuestras vidas; desde los robots en las cadenas de producción en las fábricas, en nuestra vida laboral y social por medio de los ordenadores y móviles (Siri, Alexa, Google Maps), los servicios financieros (banca on-line, fintechs), el comercio electrónico (Amazon, Aliexpress), el ocio (Spotify, Netflix, Playstation, Xbox), la automoción, etc. Es tanta la penetración en nuestras vidas que, queramos o no, debemos estar preparados para hacer un mejor uso del reto que se nos presenta delante.

El liderazgo del desarrollo de esta tecnología se está constituyendo alrededor de dos bloques claramente diferenciados; (1) el bloque occidental liderado por los EE. UU., y (2) el bloque oriental liderado por China principalmente. Entre ambos bloques, la UE aporta un liderazgo más avanzado, no por ello menos importante, en cuanto al desarrollo normativo y regulatorio de por dónde deben ir estas nuevas tecnologías.

El término *inteligencia artificial* es un vocablo que lo venimos escuchando y utilizando muy a menudo en cualquier foro donde nos encontremos y, cuando una palabra se pone de moda, su uso exhaustivo puede llevar a confusión o malinterpretación. Por eso a mí me gusta más hablar de *nuevas tecnologías basados en el análisis y explotación de los datos*.

A grandes rasgos podemos diferenciar tres tipos bien definidos de estas tecnologías:

- La *IA general o estándar*, que son el grupo de tecnologías que se basan en el procesamiento de grandes cantidades de datos a velocidades muy rápidas. Aquí observamos como se introducen dos conceptos sobre los que gira todas estas tecnologías: la *capacidad de almacenamiento* y la *velocidad de procesamiento* (por ejemplo, las cadenas de producción en una fábrica de montaje).
- Si a lo anterior le añadimos una mejora de los resultados a partir del aprendizaje de los datos analizados, tenemos el *Machine Learning* (o también llamado Aprendizaje Automático). Algunos ejemplos son los algoritmos de detección enfermedades en pacientes, apps como Uber, Cabify, Spotify, Amazon, que mejoran los resultados en base al análisis y proyección de patrones de comportamiento.
- Si a lo anterior le añadimos la posibilidad de que el sistema o algoritmo mejore y sea capaz de tomar decisiones propias, sin estar programado específicamente para ello, entonces tenemos el *Deep Learning* (o también llamado Superinteligencia Artificial). Basado en el principio de redes neuronales que imitan el funcionamiento de nuestro cerebro (algunos ejemplos son los coches autónomos, reconocimiento facial y de voz, algoritmos que detectan ciberacoso y cibercrimen en las RRSS, los asistentes virtuales y traductores online entre otros).

Estos tipos de tecnologías tienen la característica de que se basan en algoritmos que tienen la capacidad de modificarse, no sólo en función de los datos que reciben como entrada, sino en las retroalimentaciones, patrones y proyecciones de los datos de salida. La estructura de estas nuevas tecnologías es bastante simple en el sentido de que sólo hace falta tres componentes: datos de entrada, elemento de procesamiento y datos de salida. Es en el elemento de procesamiento o también llamada *caja negra o "core"* —porque es donde radica toda la inteligencia y procesamiento de los datos mediante comparativa de distintas variables, acceso a datos históricos, cruce de patrones y modelado— donde se centra todo el enigma del resultado final al no poder predecir el funcionamiento interno de este elemento, pues no se tiene la misma salida para los mismos datos de entrada. A todo esto, en 1956 John McCarthy lo llamó inteligencia artificial en oposición a la inteligencia humana, por no llamarlo inteligencia no-humana o no-biológica.

Probablemente es más apropiado hablar del término inteligencia sin ningún apellido o adjetivo que le acompañe. El término inteligencia es más propio de los seres vivos al haber una parte de interrelacionalidad subjetiva, que tiene que ver con los sentimientos y emociones, que por tanto no existe en una

máquina. Así un robot no se pone triste si pierde a su dueño o desarrollador, mientras que una persona o un perro sí lo hacen ante la pérdida de un ser querido. Entonces, ¿podemos decir que todos los seres vivos son inteligentes como el ser humano?. Lo que sí podemos afirmar es que, siendo conscientes los seres vivos, la conciencia humana es racional pues tenemos la capacidad de discernir, que nos viene dada por la libertad a la hora de la toma de decisiones, mientras que la de los animales es una conciencia irracional guiada por los impulsos naturales. Por tanto, toda la parte de emociones, conciencia y subjetividad tiene que ver con el hombre y no está presente en una máquina o robot. Es en este punto donde se libra actualmente la batalla de las principales reflexiones filosóficas para saber qué se necesita para que la inteligencia de una máquina llegue a compararse a la de un ser humano algún día.

Una máquina o robot representa fielmente lo que ha programado el desarrollador que diseña el algoritmo y, por tanto, estará muy influenciado por los sesgos de esa persona si no marcamos unas pautas muy claras. La única “subjetividad” que puede tener un robot es la que le infiere su programador. Aunque un robot sea capaz de superar al hombre en precisión, rapidez, destreza, etc., no es capaz de “recordar” nada de lo que ha hecho, es decir, actúa en función de la foto fija que tiene en cada momento, pero luego no es capaz de planificar a continuación una acción sencilla. Son muy buenos en una tarea específica pero no son capaces de desarrollar varias tareas, ni de manera simultánea, como el ser humano. Por tanto, hablar del *cociente o coeficiente intelectual (IQ)* del hombre y las máquinas no tiene sentido. Una máquina o robot no tiene coeficiente intelectual y compararlo con el del hombre sería un error.

Consideramos al hombre que es inteligente en la medida que logra unos objetivos predefinidos, es decir, tenemos éxito en su consecución. Por analogía, decimos que una máquina es inteligente porque también desempeña unos objetivos o tareas con éxito, pero a diferencia de los humanos, no son sus objetivos sino somos nosotros los que debemos proporcionarles esos objetivos para que lo ejecuten a una velocidad y precisión mucho mayor que la nuestra. Lo que consigue la máquina es una *emulación de las capacidades del ser humano* —una a la vez—, sin llegar a ser realmente cómo él o ella. Lo que realmente se está comparando es la *aptitud*, la capacidad de hacer las cosas, y no la conciencia de las máquinas. Otra característica de esta emulación es que las máquinas *adolecen de experiencia subjetiva*. Un ejemplo muy descriptivo es cuando nos damos un golpe o si nos cortamos con un cuchillo, todos sentimos el mismo tipo y grado de dolor en función de la capacidad de aguante de la persona, pero una máquina no conoce esa sensación. Un golpe en un robot como mucho produce una deformación en su estructura, pero no hace un gesto de retirarse porque no

percibe la sensación. Lo máximo sería intuir lo que se puede sentir, pero sin llegar a medirlo.

(ii) Evolución del desarrollo tecnológico

Actualmente el desarrollo tecnológico está yendo tan rápido que los seres humanos no somos capaces de gestionar adecuadamente los nuevos retos que representa la tecnología con la misma velocidad a la que se está desarrollando. Pensemos que en los últimos 300 años hemos asistido a 4 fases diferenciadas de evolución tecnológica:

- *1ª revolución industrial o de la MECANIZACIÓN* (desde mediados del s. XVIII hasta comienzos del s. XIX). Esta etapa se caracterizó por la aparición de inventos como la máquina de vapor y la energía hidráulica.
- *2ª revolución industrial o de la ELECTRICIDAD* (desde mediados del s. XIX hasta la I Guerra Mundial). En esta etapa aparecieron inventos como las líneas de producción en fábrica, la producción masiva, la aparición del acero y nuevas formas de energía (luz y gas).
- *3ª revolución industrial o de la INFORMÁTICA, también llamada 1ª revolución digital* (desarrollada entre los años 1960 a 1990). Esta etapa destacó por la aparición de los dispositivos electrónicos representado por el paso de lo analógico a lo digital, del HW al SW y la explosión de Internet y nuevas formas de conectividad.
- *4ª revolución industrial o de la DIGITALIZACIÓN, también llamada 2ª revolución digital* (desde el año 2010 hasta nuestros días). Esta es la etapa en que aparecen nuevas tecnologías que influenciarán a la sociedad con la explosión de los datos y la elevada integración entre el hombre y las máquinas.

Se comienza ya a hablar de una *5ª revolución industrial o de la SINGULARIDAD TECNOLÓGICA*, donde se vislumbra un futuro aún más incierto sobre cuál será el verdadero rol de la tecnología en la humanidad. Por un lado, se señala que es en esta etapa donde será posible trasladar la conciencia humana a una máquina y donde el desarrollo exponencial del progreso humano será tan grande, que no tendremos la capacidad de asimilarlo y, por tanto, nos debemos fiar de las máquinas (ley de retornos acelerados).

Cada una de estas evoluciones anteriores fue una auténtica revolución en la sociedad de su época, pero si bien las primeras revoluciones tuvieron unos períodos más largos y varias generaciones para asimilar los cambios introducidos, en las últimas dos revoluciones los tiempos han sido mucho más reducidos entrando en un continuo ciclo de cambio, y cada vez más complejos, difícilmente de asimilar por la mente humana. Sin embargo, debemos aprovechar toda la potencialidad que ofrece la tecnología para ponerla al

servicio del Bien Común, donde la tecnología esté verdaderamente al servicio de la humanidad, y donde la máquina complemente –y no sustituya– al hombre en aquellas tareas intrínsecas como ser (interrelación, subjetividad, emociones). El gran reto es cómo interactuamos con esas tecnologías avanzadas para hacer un buen uso.

(iii) Los nuevos retos de la IA; oportunidades y riesgos

Dada la elevada integración entre el hombre y la máquina con este tipo de nuevas tecnologías, existe un riesgo elevado de que el ser humano pierda la centralidad de nuestra sociedad en beneficio de aquellas, pasando a convertirnos en un espectador más sin participar en la toma de decisiones de por dónde debe ir el desarrollo de estas nuevas tecnologías. Es importante que la última decisión –el voto de calidad– no quede en mano de una máquina, sino que siga residiendo en el ser humano. La cuestión es cómo retener (siempre) el control sobre elementos más potentes que nosotros.

Puesto que esperamos que las máquinas cumplan los objetivos que les marcamos los humanos, estos deben estar muy bien *definidos y experimentados* para que no se genere ninguna duda en los robots sobre cuáles son las preferencias que gobiernan esos objetivos y puedan ejecutarlos correctamente. De esta manera nos aseguraremos de que la actuación de un robot sea lo menos invasiva a los intereses del hombre y sigan fielmente el desarrollo inicial para el que se han programado. Es importante que el ser humano siga manteniendo su *autonomía*. Si el desarrollo de la IA es a costa de perder la autonomía del ser humano, entonces no será una IA beneficiosa. Nadie en su sano juicio quiere una IA descontrolada de la misma manera que nadie quiere tampoco una guerra nuclear. Si sabemos hacerlo bien, nos podremos beneficiar de toda la potencialidad positiva que tienen estas nuevas tecnologías.

Está en nuestras manos y para ello se necesita mayor *claridad y transparencia* para saber cómo funciona y cuáles son los límites que no se deben traspasar, sobre todo porque no hay un “botón de retorno” que deshaga el daño que hayamos podido causar. En base a esta secuencia, se nos aparece un abanico grande tanto de oportunidades como de riesgos.

Se considerará que existe un nicho de oportunidad si:

- La usamos como *un instrumento y no es el propio fin de las cosas*. La tecnología debe ser una herramienta y no un arma de esclavitud.
- El *acceso a los medios se hace más universal*. El acceso a estas tecnologías no debe ser exclusivo sólo de unos pocos.
- Logramos que sea un *facilitador (enabler) de la vida* que nos permita aumentar la calidad y esperanza de vida. Permitiendo disfrutar de las cosas

más propias de la naturaleza del hombre, fomentando las emociones y la interrelación con otras personas.

Por el contrario, se considerará que existe un riesgo si:

- Tenemos un excesivo *individualismo y dependencia de la tecnología*. El llamado paradigma tecnocrático o enfoque unidimensional y reduccionista de que la tecnología es la única que da sentido a la vida y al funcionamiento de la sociedad (Papa Francisco, *Laudato Si'*)
- Promovemos una *magnificación de los sesgos y vulnerabilidades*, incrementando la brecha con los más desfavorecidos.
- Permitimos la *manipulación de las masas* para un beneficio concreto comercial o político (Brexit y Cambridge Analytics, influencia en elecciones, etc.)

Tenemos antes nosotros unos retos importantes como el de desarrollar una *IA éticamente responsable* que no sólo tenga en cuenta la opinión los tecnólogos, sino que recoja también los intereses de la sociedad plural y global (stakeholders). La IA ha de ser *confiable* y que nos proporcione la tranquilidad de que es para un buen fin en el que todos nos beneficiemos, a la vez que nos debe facilitar y complementar un *trabajo digno* y no una sustitución del trabajo, siendo importante la formación y reciclaje de la fuerza laboral en aquellos puestos de trabajos que aportan menos valor.

Estas tecnologías nos deben permitir adaptarnos a los cambios de la sociedad actual sin perder la identidad como personas. La tecnología no parece tener límite, pero los recursos de este mundo sí son finitos. Si no somos capaces de compaginar el uno con el otro tenemos un serio riesgo de liquidar nuestra civilización tal y como la conocemos actualmente.

(iv) ¿Cuál será el destino de nuestro futuro?

Hasta ahora hemos hecho un repaso de conceptos, implicaciones, bondades y perversidades, oportunidades, riesgos y retos de estas nuevas tecnologías para llegar a la pregunta que nos deberíamos hacer todos: ¿qué queremos como seres humanos o como sociedad?, ¿qué nos deparará el futuro?.

A lo largo de la exposición, hemos visto que hay mucha dualidad en el análisis donde la elección va a depender de nosotros y del ejercicio de nuestra libertad, que es lo que realmente nos diferencia como seres humanos y excepcionales (así hemos sido creados por Dios). Lo verdaderamente importante es que seamos conscientes de las consecuencias que tiene cada elección. Desde mi visión, considero una serie de puntos importantes para tener en cuenta:

- Se necesita hacer un esfuerzo importante desde la *educación* para entender las implicaciones que tienen estas tecnologías desde bien temprano. Esta educación debe extenderse por todas las fases del ciclo educativo hasta llegar a las universidades y empresas, pues es ahí de donde saldrán y ejercerán profesionalmente los programadores que influencien a las máquinas.
- Debemos evitar crear una IA cuyos resultados sean *indescifrables* para la sociedad. Nos corresponde ser rigurosos en la comprensión de lo que estamos creando y, cuanto más sofisticado el resultado final, más rigurosos debemos ser en la fase de definición.
- Tenemos que ser *críticos* en nuestra capacidad de análisis y decisión porque estas nuevas tecnologías no son neutras, siendo capaz de manipular a las personas por el gran volumen de datos que maneja sobre nosotros.
- Somos responsables de promover el *bien común y bienestar* de la humanidad en nuestra sociedad. Esta elección va a estar sometida a la capacidad de decisión de la gente joven y están empezando a tener conciencia de querer vivir en un mundo mejor.
- Tenemos que seguir potenciando las *relaciones interpersonales* propias del ser humano. Tenemos la capacidad de discernir entre lo que es bueno -y funciona para todos- y lo que es malo, por lo que debemos poner la tecnología a nuestro servicio para potenciar las emociones y subjetividad.

Si como sociedad somos capaces de hacer bien las cosas de acorde a nuestra naturaleza humana, de escuchar a todas las partes siendo crítico en el análisis y de poner al ser humano en el centro de las decisiones, entonces tendremos unas *tecnologías al servicio de la humanidad*. Si por el contrario, pensamos sólo en el beneficio de unos pocos y no somos inclusivos, le damos prioridad al fin y somos esclavos de las tecnologías dejándonos llevar por ellas, entonces tendremos una *sociedad robotizada*. Estamos a tiempo, pero la decisión es sólo nuestra.