

Derecho penal en tiempos de IA

-

Parece imposible sustraerse a la relevancia que están adquiriendo instrumentos como los LLM (Large Language Model). Es enorme su eficacia para generar lenguaje natural humano muy similar al que empleamos en la actividad profesional. A juzgar por su uso y la reacción en el profesorado, ya está afectando a la actividad docente. Bien por su simulación de la autoría, o bien por su uso a modo de «aprendiz de brujo», se han encendido las alarmas. No hay ámbito docente que quede al margen. Y no menos importante es lo sucedido en el ámbito de la investigación: la dificultad de acceso a las fuentes y la plasticidad de su empleo para generar contenidos nos está haciendo dudar de nuestra tarea. Algo semejante está teniendo lugar en la actuación profesional de los abogados y otros agentes del foro. Si no fuera porque llegan ya noticias de sanciones deontológicas por un uso irresponsable, parecería que la máquina iba a desplazar al ser humano.

Ante lo que ya vemos –y lo que está por venir– es comprensible reaccionar con fascinación. Fascinación ante lo que parece tener de inteligencia. Es una reacción muy humana. Recuerda a lo que se contiene en el Quijote en «la aventura de la cabeza encantada», «una de las más raras aventuras, o, por mejor decir, novedades que imaginarse pueden» (II, 62). Se cuenta en el relato cómo su anfitrión presenta a don Quijote el artilugio de una cabeza «hecha y fabricada por uno de los mayores encantadores y hechiceros que ha tenido el mundo», «que tiene propiedad y virtud de responder a cuantas cosas al oído le preguntaren», «podrá vuestra merced prevenirse de lo que querrá preguntar, que por experiencia sé que dice verdad en cuanto responde», pues «con tal traza y tal orden estaba fabricada». Todo el relato muestra la admiración –«admirado quedó don Quijote»– ante tal artificio. Hoy diríamos mejor, fascinación: «¡oh cabeza!, que lo sabes todo».

Frente a la fascinación por lo sorprendente de los instrumentos de IA, es comprensible también la reacción de rechazo, precisamente por tratarse de algo artificial. No hace falta verse golpeado por sus «alucinaciones», aunque comiencen a ser ya datos comunes. Con un mínimo de contacto experto con estos instrumentos, enseguida se perciben sus limitaciones. Sobre todo, la limitación por la sombra de su falta de autenticidad, la carencia de esa chispa de gracia que acompaña a los productos del arte. La primera impresión, la verosimilitud, cede pronto su lugar a cierta

decepción. Lo artificial emula lo humano, pero dista mucho de lo verdaderamente humano y que aquí falta: esa clave de inteligencia al objetar argumentos o rebatir falacias; o también los errores, que lejos de ser defectos, son prueba de la autenticidad de lo escrito. Por supuesto que con el tiempo la experiencia hará mejorar el uso eficaz de los instrumentos. Aunque su rápida adaptabilidad a los modos del usuario son llamativos, dejan un cierto rastro de artificiosidad no humana que nos previene. El relato de Cervantes lo descubre enseguida: «la fábrica era de esta suerte: la tabla de la mesa era de palo, pintada y barnizada como jaspe, y el pie sobre que se sostenía era de lo mismo, con cuatro garras de águila que dél salían, para mayor firmeza del peso. La cabeza, que parecía medalla y figura de emperador romano, y de color de bronce, estaba toda hueca, y ni más ni menos la tabla de la mesa, en que se encajaba tan justamente, que ninguna señal de juntura se parecía. El pie de la tabla era ansimesmo hueco, que respondía a la garganta y pechos de la cabeza, y todo esto venía a responder a otro aposento que debajo de la estancia de la cabeza estaba. Por todo este hueco de pie, mesa, garganta y pechos de la medalla y figura referida se encaminaba un cañón de hoja de lata, muy justo, que de nadie podía ser visto. En el aposento de abajo correspondiente al de arriba se ponía el que había de responder, pegada la boca con el mismo cañón, de modo que, a modo de cerbatana, iba la voz de arriba abajo y de abajo arriba, en palabras articuladas y claras; y de esta manera no era posible conocer el embuste». «En la opinión de don Quijote y de Sancho Panza, la cabeza quedó por encantada y por respondona, más a satisfacción de don Quijote que de Sancho».

Posiblemente, la fascinación y el rechazo deberían dejar paso a una reacción más prudente y sensata sobre las enormes potencialidades de estos instrumentos para la actividad intelectual de los humanos, como también sobre sus limitaciones. Y entre estas, la principal no me parece su escasa fiabilidad –que sin duda lo es–, sino sobre todo la creación de una suerte de apariencia de perfección y verosimilitud como seña de identidad de los productos que proporciona. Pero es verosimilitud en lugar de veracidad. Resulta muy fácil acabar otorgando certeza a lo que proviene de estos instrumentos; es posible que nos esté influyendo el sesgo de fiarnos más de lo mecánico que de la persona. Somos así.

Hay que reconocer que ya está modificando nuestros modos de proceder. Ciertamente hay tareas en el aprendizaje y la docencia que están cambiando, como también comienzan a ser otros los modos de acceder y trabajar la información disponible y los resultados alcanzables. No faltan estudios sobre los efectos del uso de estos instrumentos en despachos de abogados y Facultades de Derecho. Hay tareas que en otro tiempo aportaban valor a una investigación, y que ahora pasan a ser secundarias, en la medida en que estén al alcance de cualquiera. Pero también hay tareas que tuvieron valor y lo siguen teniendo ahora, incluso mayor. En particular, adquieren gran relevancia la originalidad de las propuestas en función de la personal visión de sus autores, la agudeza en las críticas, o la singularidad de los enfoques. Pero es claro que, junto a eso, está modificando nuestro modo de proceder. Quien haya afrontado la evaluación de propuestas para la publicación de lo que siempre hemos llamado «originales» puede confirmarlo. Como también podrá decirlo quien haya tenido que valorar trabajos y ensayos de los estudiantes en los últimos meses. La suma perfección de los escritos resulta sospechosa, porque resta fuerza al valor tan humano de los errores y defectos del autor, y su superación con esfuerzo. Eso es –o era, hasta el momento– el arduo proceso de aprendizaje. Tendremos que seguir ideando nuevas formas de evaluar lo que hasta ahora eran el examen o la prueba tradicionales. Ante tal avalancha de cambios y estrategias, se impone la prudencia y sensatez, para no dejarnos llevar ni por la fascinación de unos, ni por el rechazo manifiesto de otros.

Que el tema es relevante en la educación universitaria y la Administración pública parece claro. Y no es extraño que haya sido objeto de atención en instancias diversas, que alertan de su abuso y uso indebido, y proponen pautas para su utilización prudente: desde la *Guía para el uso de IA generativa en educación e investigación* de la UNESCO (2024), hasta la encíclica *Magnifica humanitas* (15 de mayo de 2026). El panorama se muestra preocupante, y resulta imposible sustraerse a él. Lo que no parece aceptable es no hacer nada y pensar que se trata de un avance técnico más, o que nuestra tarea en la academia no se verá afectada por las modas de la utilidad o la técnica.

Nunca como ahora me parece que adquiere relevancia sobresaliente la prudencia. Por supuesto, la referida al modo de afrontar estos instrumentos. Pero sobre todo me refiero ahora a la actividad prudencial de quien tiene que juzgar. Y esto lo hace el juez, como también el evaluador. La prudencia apela a la sensatez sobre los medios para alcanzar el fin. Ignorar esta tecnología sería insensato. Resultaría imprudente ignorar lo que nos pueden proporcionar estos avances, siempre que se empleen con sentido sobre sus limitaciones y con la conciencia de lo que es específicamente humano, en donde no cabe ser reemplazados por la tecnología. Ser reemplazados o sustituidos: ahí es donde deberíamos poner límites.

Vuelvo a Cervantes. El relato acaba desvelando cómo detrás de la máquina no se hallaba un genio malvado ni un homúnculo, sino un sobrino del anfitrión: por cierto, «estudiante agudo y discreto», «el cual, estando avisado de su señor tío de los que habían de entrar con él en aquel día en el aposento de la cabeza, le fue fácil responder con presteza y puntualidad a la primera pregunta; a las demás respondió por conjeturas, y, como discreto, discretamente». Pero hizo pasar por verdaderas las respuestas. Y era un estudiante.

Pablo Sánchez-Ostiz Gutiérrez