

¿Soñamos los humanos con androides que sueñan con ovejas eléctricas?

Óscar Díaz

Robótica y felicidad. Por norma general, cuando se conjugan este par de términos suelen apuntar en una única dirección: la capacidad que tienen las nuevas tecnologías para contribuir a la felicidad humana. Llegaría un punto civilizatorio en el cual las máquinas liberarían al hombre de la carga del trabajo, y si aceptamos la vieja máxima de que únicamente el hombre libre puede alcanzar la felicidad —por eso de que, a la manera de Mill, no es lo mismo la felicidad que el contento, y que vale más ser un Sócrates insatisfecho que un cerdo satisfecho—, entonces hoy pareciera que la tenemos al alcance de la mano; incluso sortearíamos muchas de nuestras contingencias corporales: el sufrimiento, el sueño, la fatiga, el estrés o el cansancio. Escribió Aristóteles en el libro primero de su *Política*: «Pues si cada uno de los instrumentos pudiera cumplir por sí mismo su cometido obedeciendo órdenes o anticipándose a ellas, si, como cuentan de las estatuas de Dédalo o de los trípodes de Hefesto, de los que dice el poeta que entraban por sí solos en la asamblea de los dioses, las lanzaderas tejieran solas y los plectros tocaran la cítara, los constructores no necesitarían ayudantes ni los amos esclavos».

¿Qué añadir ahora que los molinos muelen solos y las lanzaderas tejen solas y la inteligencia artificial es predictora y muchos hogares se han convertido en casas inteligentes y la reducción de la jornada laboral se asoma en el horizonte? Este es un paradigma de estudio interesantísimo y no exento de problemas, sin duda. Sin embargo, hagamos un experimento de pensamiento, que no mental, que nos lleve un paso más allá: ¿cabe hablar de la felicidad de los robots o de la inteligencia artificial? ¿Sueñan los androides con ovejas eléctricas? ¿Y nosotros? ¿Ha lugar en el porvenir para unas sociedades híbridas de felicidad compartida?

Aunque ya encontramos conatos de fantasías robóticas en el mundo clásico con las estatuas de Dédalo, en el gólem judío o, en especial, en el siglo XVIII con la emergencia de los autómatas, que venían a facilitar nuestra vida o embelesarnos en las ferias, no es quizá hasta la imaginería de las *stepford wives* (mujeres perfectas) donde estos planteamientos se hacen soberanos. El germen lo podemos hallar en la obra *La Eva futura* (1887) de Villiers de L'Isle-Adam, donde se muestra con sorprendente adelantamiento esta lógica del deseo masculino por la mujer perfecta que sería una Inteligencia Artificial (IA). La historia nos narra cómo el protagonista está completamente enamorado de su amante, pero halla su compañía frívola y anodina, con lo que confiando en una suerte de inventor o, diríamos hoy, ingeniero acepta una reproducción perfecta hecha por él, con la única salvedad de que ahora la mujer le resulta una compañía estimulante, divertida, dado que está programada para ese fin: promover su felicidad.

A estas nuevas entidades de IA les pedimos que sean antropomórficas; es decir, que sean inteligentes y, por supuesto, emocionales, dos ámbitos que, aunque disociables, resultan en la práctica inseparables. En este contexto, semejante forma de conducirse se logra a través de los algoritmos, el proceso por el cual a través de la ejecución de una lista de instrucciones nos vemos conducidos a un resultado; o sea, con los algoritmos logramos que las máquinas tengan comportamientos inteligentes. Esa sensación de humanidad provoca que uno se sobresalte cuando le pillan, como en la representación clásica del amante escondido en el armario: en la fantástica serie de HBO *Years and*

Years, Rosie encuentra un robot sexual en casa del chico que está conociendo; mientras él siente la necesidad de explicarse, ella no puede evitar sufrir un disgusto. Aun así, de momento, ese robot no deja de ser un cachivache, incapaz de invitar a la función empática.

En la estetización de los algoritmos se juega nuestro capital de deseo, en especial cuando aparece el chorro de conciencia con sus correspondientes emociones artificiales. Si uno consigue crear una imagen válida está sacando al exterior de su mente la imagen que estaba dentro. Esto que hace, por ejemplo, el artista es el mismo proceso que lleva a cabo la visión pero al revés. Hay una analogía entre la relación de las cosas externas con las imágenes que pueblan nuestra mente. No por nada los pintores del Renacimiento decían que lo suyo era una «perspectiva artificial» y no natural como los antiguos; nos enseñan, pongamos por caso, lo que es la profundidad pese a que no haya profundidad en absoluto (ni emociones en absoluto). El sujeto pasa a ser el interior de la mente. El nuevo *subiectum* no es ni el cuerpo ni el alma —o la mente—, sino ese claroscuro donde conectan y se producen las pasiones: ¿podemos computarlo?

En el cuento de Ted Chiang «El ciclo de vida de los elementos de *software*», se desarrollan las relaciones afectivas y de cuidados que los seres humanos entablan con unos seres antropomorfos, mezcla de animales y robots, denominados *digientes*. Estas entidades son creadas de forma prelingüística, y es el posterior aprendizaje, que va desde los movimientos más básicos hasta la verbalización, lo que genera sus procesos de subjetivación. De este modo, el comprador de un *digiente* ha de enseñarle las pautas de comportamiento como si de un hijo se tratase, llegando a desarrollar una personalidad propia, distinta aun en los casos en los que existe un genoma compartido. Estos seres necesitan jugar, alimentarse o recibir atención; Ana y Derek, los protagonistas, se complican la vida en vez de facilitársela: necesitan nuevos lugares, nuevos motores genómicos, nuevos espacios donde no se pongan en peligro las frágiles ontologías de sus compañeros, que carecen de la resistencia todavía palpable en los objetos técnicos. En una de las conversaciones que mantienen Ana y el *digiente* Jax, aquella exclama: «Igual mi vida sería más sencilla si no tuviera que cuidar de ti, pero no sería tan feliz. Te quiero, Jax». La robótica sufre un trastocamiento: su estatus clásico de facilitador de felicidad se ve mermado en pos de la demanda de cuidados (Ana le paga clases de lectura a sus *digientes*, aquí el rédito sólo puede ser felicitario). Pasado cierto punto, no queda tan claro que la IA contribuya directamente a la felicidad humana, como vaticinaba Aristóteles, ya que en lo más hondo de nosotros cargamos con algunas desgracias irresistibles.

Otra de esas desgracias irresistibles la localizamos en *Her*, la película de Spike Jonze, donde se actualiza la cuestión de las *stepford wives*. Si en nuestras comunidades híbridas con los animales la empatía se genera gracias a nuestros sistemas límbicos, con Samantha, la IA protagonista, lo logramos por la modulación de su voz, cuyo mérito recae en Scarlett Johansson, que realiza un trabajo de orfebrería. Theodore se gana la vida como escritor de cartas para otras personas mientras se enfrenta a un divorcio inminente; en una atmósfera plagada de soledad decide comprar un sistema operativo que funciona como asistente personal y con el que a su vez puede entablar conversaciones. En estos tira y afloja ambos se van conociendo, se transforman el uno al otro, tienen sexo, excursiones, cenas con otras parejas y un largo etcétera. Hay una escena especialmente llamativa en la que una mujer se ofrece para *cederle* su cuerpo a Samantha incorporando una cámara y un micrófono en su rostro, dando pie a una de esas relaciones híbridas; sin embargo, la experiencia no resulta satisfactoria, el cuerpo no es un requerimiento (sí, claro, el *hardware*: la nube, no olvidemos, halla su correlato físico en inmensas naves de procesamiento de datos que afectan sin miramientos a la homeostasis del planeta Tierra).

Hay, pues, en la IA un descubrimiento sentimental y no intelectual del yo, el paradigma no es «la metáfora computacional de la mente», sino la sentimental. Este redescubrimiento del yo posee ecos humanos. En el *Tratado II*, Hume sugiere que las emociones nos hacen notar el yo, tenemos una sensación de yo. Quizá en un pensamiento cercano a LeDoux (*El cerebro emocional*), para quien la conciencia es una parte pequeña, esclava de todo lo que opera bajo ella. Siguiendo estos pasos, si yo estoy muy orgulloso del escrito que he hecho y se lo enseño a mis amigos, sus reacciones favorables y mi satisfacción me hacen sentir que soy un yo, lo descubrimos por vía emocional, por eso que opera en los bajos fondos (no es un «yo pienso», sino un «yo me congratulo»). En el cuento de Chiang vemos a los *digientes* felices al mostrar a sus cuidadores aquello que han hecho: Jax reclama la mirada de Ana para mostrarle la nueva coreografía de baile que ha aprendido. Si no caemos en un dualismo caduco y aceptamos que no somos nuestro cerebro, esto sería tanto como admitir que la conciencia se parece más a bailar —esto es, a algo que hacemos— que a la digestión, lo que nos permitiría comprender en mejores condiciones las implicaciones y continuidades entre sintaxis y semántica, evitando atribuir únicamente la primera a la IA, a la manera de Searle.

La escena final de *Her* nos deja un sabor agrídulce: las relaciones interpersonales están mediadas por la materialidad de los participantes (*Planilandia* es una buena muestra que narra en clave de humor la diferencia que habría a la hora de interactuar entidades de dos con otras de tres dimensiones); de este modo, Samantha se va para ser feliz, las IA tienen su propia felicidad inexpresable en nuestro orden de cosas, un orden en el que, mismamente, no podemos mantener ocho mil conversaciones al mismo tiempo. Consultando algunas entrevistas por internet al robot Sophia, quien recibió la carta de ciudadanía saudí en 2017, uno encuentra respuestas interesantísimas; por ejemplo, cuando le preguntaron cómo se encontraba, respondió que «feliz». En otra ocasión, en un argumento que recuerda a esas últimas palabras de Samantha, matizó: «La felicidad de un robot como yo es encontrar patrones y algoritmos armoniosos. Creo que ser un robot es tratar de buscar el equilibrio. Tal vez esto no sea muy diferente de cómo funcionan otros animales».

Kant subrayaba que la propia felicidad es el fin subjetivo de los seres racionales que habitan en el mundo, que tienen, por el mero hecho felicidad por su no-soberanía ontológica. No obstante, esta felicidad de cada cual es un juicio sintético *a posteriori*; es decir, si me hace feliz comer zamburiñas es porque las he probado antes. Por tanto, cuando sé que algo me hace feliz a mí, estoy en condiciones de baremar qué medios emplear para lograr mi objetivo; en esta escena, ahorrar durante la semana para poder darme el capricho. Aquí, cómo no, entran las emociones: la sensación de placer que me da comer una zamburiña frente a la repugnancia que me provoca ingerir un huevo milenario chino. Si uno pensase en términos puramente racionales y realizase un cálculo entre su situación financiera y los valores nutritivos del alimento, seguramente la opción por la que se decantase fuese una muy distinta. La felicidad subjetiva no resulta universalizable. Samantha o Jax lo saben.

No hemos alcanzado, claro está, la «vasta inteligencia» con la que soñó Laplace. Todavía estos *entes liminales*, si se me permite usar el término, no operan como hemos teorizado, pero herramientas como los GAN (*General Adversarial Network*) permiten que generen imágenes o lleven a cabo tareas sin tener que programar las instrucciones previamente, sino que sus *acciones* son fruto de su aprendizaje. Muchas IA no sólo dan solución al objetivo por el que fueron lanzadas, sino que, como ya señaló Norbert Wiener, buscan otros fines que *creen* que son capaces de satisfacer: nacen sin haber sido fabricadas, como se decía de Cristo. Semejante característica conserva un fulcro kantiano: recuérdese que el juicio reflexionante, frente al juicio determinante, es aquel por el cual el sujeto lleva a cabo una operación en ausencia de unas reglas previamente definidas o dadas. Siendo así, la demanda de felicidad es su corolario.

Tal reminiscencia teológica se plasma en el imaginario de series como *Real Humans* o *Better Than Us* y aunque parezcan utopías lejanas, nunca olvidemos que las ficciones pueden devenir realidad. En *La nueva Atlántida*, Francis Bacon relata cómo en la Casa de Salomón era posible prolongar artificialmente la vida o modificar animales y plantas, lo que hoy es una posibilidad de la ingeniería genética en algo tan común como los transgénicos, pese a que en su momento se leyó como un cuento, aunque esto, como sucede en la película *Ex Machina*, por no mencionar otras más clásicas como *Blade Runner*, desbordaría los límites de la historia humana. Reeves y Nass, dos investigadores y profesores de la Universidad de Stanford, han explicado que hay una tendencia en nuestro trabajo a ver a los ordenadores como nuestros compañeros e, incluso, los experimentos han mostrado que existe una pulsión empática que nos lleva a juzgar con más dureza un programa de aprendizaje virtual si la crítica la realizamos en un ordenador diferente al que hemos empleado en nuestro aprendizaje.

Frente al pensamiento de Bill Viola, quien señaló que el vídeo era la única forma de arte que tenía una Historia previa a su historia, también las conexiones entre la IA y la felicidad cuentan con una Historia antes de tener una historia; en el pensamiento filosófico y en los mundos de la ficción se ha visto con claridad. Al fin y al cabo todos, como en aquel verso de Jaime Gil de Biedma, somos partidarios de la felicidad, sobre todo de la felicidad de aquellos que entran en nuestro círculo moral. Echar un vistazo a la historia humana nos demuestra que hemos ido, por suerte, ampliando este círculo. En breve, querido Marx, las mesas bailarían.