

La peligrosa idea de Turing

TURING'S DANGEROUS IDEA

Javier Blanco

Facultad de Matemática, Astronomía Física y Computación de la Universidad Nacional de Córdoba (UNC)

Facultad de Ciencias Exactas, Físico-químicas y Naturales, Universidad Nacional de Río Cuarto (UNRC)

javier.blanco@unc.edu.ar

Código ORCID: 0000-0003-4416-7892

Marcelo Ruiz

Departamento de Matemática, Facultad de Ciencias Exactas Físico-Químicas y Naturales, Universidad Nacional de Río Cuarto

mruiz@exa.unrc.edu.ar

Código ORCID: 0000-0002-2336-9316

Presentación de conferencia por Marcelo Ruiz

La Diplomatura Superior Tópicos de Ciencia de Datos tiene diversos objetivos explicitados en su plan de estudio enmarcados en la propuesta de formación de posgrado. Y también a partir de estas actividades extracurriculares pretendemos generar instancias abiertas a toda la comunidad para abrir y problematizar ciertas nociones del sentido común que, por provenir de él, obturan debates y obstaculizan el necesario rodeo de todo proceso de comprensión de lo «real».

Ya hace un tiempo, cuando comenzábamos a plantear la posibilidad de la existencia de la Diplomatura, en una mesa redonda de la cual participamos Guillermo Durán, Marcelo Arroyo y yo, y que fue organizado por la Secretaría de Posgrado de esta facultad, enfatizamos en la importancia de abrir los significados de uno de esos sintagmas, el de Ciencia de Datos y titulamos esa mesa “¿Qué es esa cosa llamada ciencia... de datos?”, tomando prestado el título del texto de Chalmers.

Parafraseando a la editorial Mimesis, tal vez necesitemos apostar a la reivindicación de la urgencia de la crítica y la teoría para un presente saturado de empirismo y de datos y que, como nos invita Rodríguez Freire¹, tomemos distancia de un espíritu burocrático neotaylorista que no ha hecho sino disciplinar y controlar el trabajo intelectual.

Invitamos a Javier con la libertad de que diga lo que quiera, pero le insistimos en la posibilidad de poner en tensión los sintagmas Inteligencia Artificial y Ciencia de Datos que hoy forman parte de dispositivos discursivos muy recurrentes y potentes, que incluso suelen funcionar por su carácter performativo.

Conferencia Javier Blanco²

Muchas gracias por la presentación, estoy emocionado de estar acá, esta es mi casa, doy clases desde el 2000, aunque ahora estoy con licencia en el cargo que tengo acá en la facultad de ciencias exactas. Todo este último tiempo vengo trabajando más en el área de filosofía y computación, de hecho, mi primera doctorada de filosofía es Aldana D’Andrea, de esta

Universidad.

Voy a arrancar directamente “a los bifés”. Primero plantear que estamos en una situación realmente inusual en el plano político, y todo el tiempo aparecen indicios de estar transitando una transformación radical. Quiero mostrar cómo entender un poco esto. Pero antes, para motivar e ilustrar el estado de cosas, presento algunos casos ejemplares actuales.

La semana pasada estaba en Mar del Plata y me llamaron de Uruguay por una entrevista. Dado el tema que me proponen, y que luego cuento, empecé a buscar casos de influencias directas de las tecnologías digitales en el ámbito político, específicamente en los procesos electorales, aunque también en el ámbito de la gestión. Hay una cantidad de casos que son más o menos conocidos, que les voy a mencionar para recordar qué es lo que enfrentamos.

La famosa, la infame, la tengo que poner en inglés para que quede el chiste lingüístico –(in)famous–, Cambridge Analytica, una empresa que tuvo una importante influencia en procesos electorales y que sigue existiendo, aunque fue cambiando de nombre. Constituyó una gran novedad en los años 2015 o 2016 más o menos, cuando mostró una capacidad importante para influir en las decisiones de voto en distintas elecciones en diferentes países. Usando tecnología aplicable en plataformas digitales conseguía deducir con alguna probabilidad qué iba a votar cada quien, y podían también influir en esas decisiones a partir de una perfilización de los votantes. Esa es una idea que todos hemos escuchado y que sigue teniendo consecuencias electorales hasta hoy.

Hay muchos más casos. Llamo a uno *Breaking Bad* de Luisa González. Luisa González fue, se acuerdan, hace un año la candidata a completar el mandato que dejó vacante en Ecuador, Guillermo Lasso. Los principales candidatos para esa elección, por un período breve, fueron Luisa y Daniel Noboa, de la derecha, quien finalmente ganó.

Circuló un video donde ella está hablando en un programa de televisión y en un momento dice que el fentanilo, una droga muy peligrosa y adictiva, debería ser incluido entre las drogas de consumo personal, que debía ser despenalizada. Ese video es falso, fue alterado usando herramientas de lo que hoy se llama IA. Es bastante fácil de hacer con herramientas que se bajan de internet. Circuló por todas las redes, pese a que se lo denunció y se intentó bajar. El día después de las elecciones, desapareció de casi todas las redes.

Revisando hoy qué hay sobre este hecho del año pasado, encuentro primero una nota donde algunos periodistas refutan que ese hecho haya realmente ocurrido, mostrando el video original y el video adulterado. Los comentarios en esta nota, quizá hechos por bots o trolls, parecen no creer en la refutación y refuerzan su voto por Noboa.

Otro caso: *Bollywood plays for Congress Party*. Este año hubo elecciones presidenciales en India, con 965 millones de electores. Imaginen e intenten dimensionar ese número.

Fue una elección de un país con 22 lenguas oficiales y estuvo atravesada por todo tipo de formas tecnológicas para bien y para mal. Una de las primeras manifestaciones fue una publicidad del Congress Party, partido socialdemócrata opuesto a la derecha gobernante. Este partido comenzó su campaña con una semana de *fake news*, de *deepfakes* con estrellas de Bollywood –la principal productora de cine del mundo–, que criticaban al gobierno, mayormente por sus políticas culturales. El problema es que estas actrices y actores nunca habían hecho esas declaraciones, fueron forjadas por diferentes programas de IA.

Encontramos también muchos muertos hablando de política, apoyando a los nuevos candidatos. En Tamil, por ejemplo, el ex líder de uno de los partidos, muerto en 2018, presentaba al nuevo líder. Todo el mundo sabe que está muerto, no es que no se sepa, pero igualmente... Otro caso fue el de una mujer que murió en el 2013, que era entonces una

influencer importante, una deportista que muere joven, no sólo reapareció hablando, sino que se la recreó envejecida por los años que habían pasado, tenía 10 años más en la imagen que habían hecho.

Estos casos ilustran un poco el panorama actual. Último ejemplo: en una elección local de este año en Estados Unidos, Joe Biden llama a los votantes a no votar. Hay un llamado por teléfono, automático, hecho por un bot con su voz, donde propone no votar en esa elección, argumentando que votar favorecería a Trump, algo obviamente falso. Frente a esto, lo que se resolvió en ese estado fue prohibir los mensajes de voz automáticos.

Volviendo a India, además de todos los *fakes*, hubo cosas súper interesantes como, por ejemplo, cuando el candidato del partido de gobierno dio un discurso masivo en el que la gente, con un audífono con interpretación simultánea, al estilo traductor universal de *Star Trek*, podía escucharlo en su lenguaje favorito de los 22 idiomas oficiales de India. Durante el proceso electoral había empresas que ofrecían un servicio “ético”, proponían una IA ética para la política en India. El o la candidata, dejaba su voz grabada, su imagen y sus propuestas, y la empresa producía los discursos que eran pronunciados por un avatar del candidato.

Como les decía, esta mini investigación y reflexión se inspiró en una llamada de una periodista de Uruguay, quien me cuenta un hecho aún más exagerado del uso de este tipo de tecnologías. Invitan al entonces candidato presidencial por el Frente Amplio, Yamandú Orsi, al programa de televisión Santo y Señá. La convocatoria era para hablar de seguridad, luego de algún hecho aberrante y con claras intenciones electoralistas, junto al candidato del Partido Colorado. Ante la decisión, esperable dado el tipo de programa, de no concurrir, pusieron un avatar creado por IA con su imagen, hablando en el debate, mientras pasaban otras entrevistas grabadas. El “periodista” dice en algún momento: “Bueno, hemos conseguido lo imposible, está acá sentado Yamandú en nuestro programa”. Si bien al final dijeron que no era el Yamandú real, quien no vio esa parte puede haber creído que realmente lo estaba escuchando, “perdiendo el debate”.

Sigamos con la presentación con una breve digresión lingüística:

电脑

(Se pronuncia Diàn nǎo)

Así es como los chinos llaman a la computadora, “cerebro eléctrico”, muy similar a cómo se refería a ella el mismo John von Neumann en su *first draft*.

El idioma chino tiene desde sus orígenes cerca de 4500 años y se mantiene bastante estable durante los últimos 2500 años, algo muy diferente a la mayoría de los otros lenguajes. Pero obviamente tuvieron que inventar palabras para nuevas entidades que fueron apareciendo, como en este caso.

¿Cómo y cuándo apareció la idea de computación? No voy a recorrer demasiado esta historia, pero una de las tradiciones que confluyen en la maduración de esta idea es la de los dispositivos de cálculo mecánico, cuyo origen puede rastrearse hasta Ramón Llull en el siglo XIII, pasando por Pascal, en el siglo XVII, y luego Leibniz, entre los siglos XVII y XVIII, con diferentes máquinas combinatorias.

La tradición más influyente fue la de la lógica formal, que también tiene en Leibniz una de las manifestaciones más importantes, en particular la idea de definir representaciones del mundo y operar formalmente con ellas. La *Characteristica Universalis* y el concomitante *Calculus Ratiocinator* fueron la apuesta de Leibniz en esta dirección. Buscaba una manera de poder codificar todos los problemas del universo y poder disponer de un cálculo capaz de

resolver todas las controversias. ¿Existe Dios? Bueno, lo calculamos y decidimos. ¿Para qué tanta discusión, o incluso persecución? Calculemos, decía Leibniz.

Este fue el sueño de Leibniz, disponer de formas de calcular y poder responder las principales preguntas del mundo. De alguna manera, Turing encarnó ese sueño.

Por el lado mecanicista, encontramos también otra confluencia que provino de los telares programables de Joseph Jacquard, que inspiraron a Charles Babbage para el diseño de su motor analítico, que puede verse hoy como la primera máquina universal de cálculo, en los inicios del siglo XIX.

La idea de programa va a ser uno de los conceptos fundamentales en lo que sigue, la capacidad de que un texto formal puede codificar de manera precisa cualquier comportamiento. Veamos cuáles son las características de esta idea tal como la concibe Turing en 1936.

Una breve digresión: ¿Vieron la película sobre Alan Turing? Es esa que creo se llamó en castellano “El código enigma”, y en la versión original era “The imitation game”. Es una linda película que caracteriza muy bien a quien para mí es uno de mis héroes del siglo XX, aunque tiene un par de errores significativos.

En la película se van intercalando tres momentos de la vida de Turing: su adolescencia, cuando descubre su amor por Christopher, el trabajo en el año 1941 en Bletchley Park, haciendo el criptoanálisis de la máquina Enigma de los alemanes, y en el año 1952 cuando es condenado por homosexual y obligado a tomar hormonas para reducir el deseo sexual (o ir preso). Esta última parte, presenta a un Turing con problemas cognitivos producto del “tratamiento”, algo que nunca ocurrió (aunque sí le crecieron tetas, producto del estrógeno), ese es uno de los errores de la película. El otro es que falta el período más importante de todos, el que realmente transformó el mundo de manera perdurable, centrado en el año 1936 cuando resuelve el *Entscheidungsproblem* de Hilbert y como consecuencia de ello inventa la noción de computación y de computadora. Es una curiosidad histórica que los principios que hoy sustentan una nueva era tecnológica emergieron de un problema de filosofía de la matemática, en particular la paradoja de Russell en la teoría intuitiva de conjuntos.

No voy a abundar sobre eso que es bastante conocido. Quiero caracterizar una concepción intensional de computación (intencionalmente escrito con ‘s’, usado en el sentido lógico de algo no-extensional). La noción de computación de Turing básicamente tiene tres dimensiones importantísimas que voy a traer a colación para después pensar cómo eso influye en nuestro mundo actual. Las características que me interesa abordar hoy son las de efectividad, programabilidad y abstracción.

Primero la efectividad. La idea de efectividad fue la gran pregunta de Turing, que implicaba entender qué significaba el cálculo mecánico, aunque no sé si esa es la mejor palabra. Turing construye en realidad un mecanismo teórico, un mecanismo abstracto que caracteriza cómo se resuelven problemas matemáticos de manera mecánica.

Esa máquina abstracta se presenta como un humano siguiendo reglas elementales precisas. Cuando Wittgenstein dice que las máquinas de Turing son personas calculando... La respuesta es un categórico sí, lo dice Turing en la segunda página de su artículo. No descubriste mucho, don Ludwig.

Aparece acá la idea de algoritmo, que indica un método preciso, expresado paso a paso, que resuelve un problema dado. Es paradójico que se haya popularizado el uso de esta palabra, pero para referir a algo que no es *stricto sensu* un algoritmo sino un clasificador estadístico, es decir, los programas generados por la técnica de *machine learning*. Una consecuencia

importante de lo que podríamos llamar la tesis de Turing es que, si hay un método preciso para resolver algo, entonces hay una máquina de Turing que lo resuelve. Es decir, no es que los algoritmos son entidades muy específicas, cuya característica principal se llama hoy efectividad. Un algoritmo describe un método efectivo.

Otra consecuencia de esto, que no voy a trabajar acá, es que la idea de efectividad sirve para caracterizar de manera general la noción de mecanismo. Extiende esa noción, pero, por primera vez, delimita cuál es su alcance, mostrando que hay problemas que no pueden resolverse mecánicamente. Decía Canguilhem que el mecanicismo como propuesta filosófica, estuvo asociado a las máquinas de la época. Para Aristóteles era la catapulta, para Descartes, sobre todo, eran los autómatas hídricos o de relojería. Después, para la época del siglo XVIII, eran las máquinas de vapor. Curiosamente la máquina de Watt se llamó máquina universal.

Hoy podemos decir que las máquinas paradigmáticas son las máquinas informacionales.

Como decíamos, esto permitió, por primera vez, establecer límites precisos a los mecanismos. Es decir, hay cosas que no se pueden hacer mecánicamente. El conocido *halting problem*, el problema de la parada de una máquina, o equivalentemente, la posibilidad de tener un algoritmo o programa que tome otro programa como dato y determine si éste se va a detener para una entrada dada, fue demostrado imposible en el artículo de Turing de 1936.

Aclaro algo importante al respecto. Que no pueda haber un programa, significa que tampoco lo pueden resolver mentes humanas de manera sistemática. Si no hay algo mecánico, tampoco hay algo diferente en los humanos que podría hacerlo. Fue una larga discusión desde los años 60, 70, pasando por Lucas, Penrose, quienes querían usar los teoremas de Gödel y Turing para “demostrar” que las mentes humanas exceden cualquier programa, pero eso no va a ningún lado. Por el contrario, la evolución constante de los sistemas algorítmicos parece indicar que cualquier proceso mental humano terminará siendo excedido por variantes algorítmicas.

La idea de que los humanos hacen mejores cosas cognitivas que las computadoras, yo creo que no es sólo posiblemente falsa, sino, sobre todo, *misleading* (engañoso), desorienta para pensar estas cuestiones.

La segunda característica, que también muestra Turing en 1936, es la existencia de la llamada máquina universal, es decir, una máquina completamente programable ¿Qué significa esto? Que hay una máquina, que puede construirse, que de cierta forma es todas las máquinas.

Esta idea implica que todo programa es un dato para otros programas, en particular para un intérprete, un programa que puede ejecutar el código de cualquier programa, que es una forma de ver a la máquina universal. No hay una diferencia ontológica entre programas y datos, sino que lo que los define se relaciona. Algo es un dato porque hay un programa que lo toma como tal. En general, que algo se vuelva un dato significa que hay una entidad –puede ser también humana– capaz de procesarlo como tal. Este es el origen en la multiplicación de datos que ha habido, la proliferación de este tipo de entidades.

Hay aquí un concepto fundamental para la computación y que se encarna en la noción de intérprete: un vínculo causal entre codificación y comportamientos. Este vínculo puede realizarse de múltiples maneras, con diferentes tecnologías, pero su existencia es condición necesaria para que haya computación. Es también lo que hace que programar sea una operación sobre el futuro, programar es crear lo que Husserl llamaba protenciones.

Es decir, las maneras de proyectar el futuro se pueden programar, o esa programación

puede ser interpretada con precisión por entidades no siempre humanas, y que tienen agencia propia. Cada vez más la agencia está desplazada hacia entidades no humanas. Eso es parte del mundo que tenemos, constituido no ya sólo por objetos técnicos, sino sobre todo en ambientes técnicos. Y entonces toda acción se vuelve acción técnica. Acciones técnicas como parte de ambientes con gran capacidad de acción ellos mismos.

John Von Neumann decía que en cualquier mecanismo o autómatas previo a las computadoras era mucho más fácil describir su comportamiento que su estructura. El ejemplo típico es el reloj, o los autómatas que le gustaban a Descartes. Autómatas que tenían comportamientos elementales y repetitivos a través de algún mecanismo bastante sofisticado.

Pero, siempre según Von Neumann, hay un umbral de complejidad a partir del cual esto se invierte. Y el primer mecanismo en franquear ese umbral fue la computadora, en tanto máquina universal. Va a ser mucho más complejo el comportamiento que la estructura.

Esta idea, que contradice nociones clásicas sobre mecanismos, se sustenta tanto en la idea de universalidad, es decir la posibilidad de convertir cualquier objeto sintáctico (y hay un número infinito de éstos) en comportamiento, como en la inédita capacidad de abstracción que habilitan los sistemas computacionales. Como indicaba Dijkstra, los sistemas computacionales trabajan en órdenes de magnitud inconcebibles antes. Entre el hardware y las aplicaciones actuales, hay un salto de unos diez órdenes de magnitud, y lo que permite ese salto son programas.

La simplicidad de una máquina de Turing que puede implementarse con lápiz y papel, la relativa simplicidad de cualquier computadora, queda asociada a la inabarcable, literalmente infinita variedad de sus comportamientos. Esto da lugar a un mundo que no estábamos en condiciones de comprender.

La transformación siguiente se da a partir de la década de 1990, cuando cobra forma y se difunde la computación ubicua, las redes, Internet. Todo se vuelve red. La aparición de los celulares acelera ese proceso. Pero veamos cómo se dice celular en chino:

手机

(Shǒu jī)

El primer carácter es mano, el segundo es máquina. Una mano mecánica, eso es el celular.

El universo, que otros llaman la Biblioteca, decía Borges, anunciando quizá la aparición de Internet, la red de redes, la casi infinita biblioteca de Babel.

A fines del siglo pasado, Hubert Dreyfus, un fenomenólogo interesado en pensar las computadoras, también el introductor de Heidegger en Estados Unidos, escribió un libro sobre Internet (On the Internet). En su primera edición, cuyo manuscrito entregó en 1999, realizó una predicción que hoy resulta divertida: Internet no va a andar.

En base a sus propias especulaciones y a consultas con los informáticos de la época, veía que no había forma de que eso funcionara. ¿Por qué? Por una simple cuenta. El crecimiento exponencial de la cantidad de páginas y la imposibilidad ya enunciada por Borges de tener el catálogo de todos los catálogos. La búsqueda en esa primera Internet era análoga a recorrer un catálogo, sus categorías y sub-categorías. El sistema iba a colapsar pronto.

En la edición de 2009, en cambio, cuenta por qué Internet funciona. ¿Qué cambió en esa década? Bueno, un informático formado en su mismo grupo de investigación, Larry Page, creó el *PageRank*, un algoritmo de clasificación de páginas que dio origen a Google y

quizá al mismo siglo XXI.

¿Qué hace ese algoritmo? En su idea básica, más allá de que hoy, se ha sofisticado ad infinitum,

calcula la relevancia de una página a partir de un análisis semántico elemental y de los links que otras páginas análogas tienen hacia ésta. Es decir, al buscar se siguen las huellas de lo que otros buscaron, las indicaciones de otras páginas. También el peso que da un link depende de la relevancia de la página de origen. Este algoritmo ha ido evolucionando incorporando las propias búsquedas desde el perfil que está haciendo la nueva búsqueda.

Según Dreyfus, la idea surgió en un seminario sobre Heidegger que estaba dando Terry Winograd, y la idea es claramente heideggeriana. Frente al crecimiento exponencial de la cantidad de páginas, y más aún a los vínculos entre páginas, la idea es aprovechar eso para construir los índices de búsqueda, que ahora serán un reflejo de lo que la gente busca o señala. En este sentido, es un algoritmo idiosincrático, y cada vez que buscamos o vinculamos nuestras páginas, lo vamos realimentando.

Es decir, el mundo que vemos desde y gracias a Internet, es decir el mundo *tout court*, es un mundo contingente, signado por una historia que podría haber sido otra. Y que podría ser otra.

Cabe preguntarse qué significa la objetividad cuando el acceso al mundo está mediado por esta infraestructura algorítmica. La idea de objetividad no puede obviar a la mediación como parte del objeto mismo, una especie de objetividad relacional, en cierto sentido relativa. La perspectiva misma, materializada en los programas, son parte inescindible de los hechos.

¿Qué es esa cosa llamada información? ¿La ciencia de datos se ocupa de la información? ¿Hay teorías que den cuenta de qué es la información hoy?

Las teorías de qué es la información van a aparecer asociadas a la idea de orden del mundo, incluso en los sistemas físicos. Entropía, neguentropía y anti-entropía son conceptos que van a aparecer asociados a esta aproximación a la idea de información. Van a aparecer los problemas teóricos de Boltzmann y Maxwell, resueltos en un cierto sentido por el físico húngaro Szilard.

Toda esa historia desemboca en la teoría estadística de la información (llamada originalmente teoría matemática de la comunicación), desarrollada por Claude Shannon en 1948, que mide la información como reducción de la incertidumbre y resistencia al ruido, básicamente. Bueno, no quiero entrar mucho en eso ahora, pero aparece como central el concepto de entropía que en las teorías cibernéticas se asocia a otro concepto fundamental, como el de variedad, que retomaré sucintamente.

Es algo paradójico que las teorías de la información en uso hoy son, en su mayoría, de antes de la década de 1960, como la ya mencionada teoría estadística de Shannon (también la desarrollada por Norbert Wiener), o la teoría algorítmica de la información (también llamada complejidad de Kolmogorov).

Hay algunas que otras, pero en principio, cuando hablamos de información, hablamos de teorías que son de antes de que apareciera Internet. ¿Estamos hablando de lo mismo cuando hablamos hoy de información? Lo dejo como pregunta, no sé la respuesta.

Una de las razones por la que me inclinaría por la negativa, es decir que ha habido un cambio cualitativo en qué entendemos por información, se basa en el gran cambio cuantitativo acerca de cuánta información se produce.

Hasta la popularización de las computadoras personales y la aparición de Internet a inicio de la década de 1990, el total de información producida por la humanidad desde sus orígenes se estimaba en 12 exabytes. No estamos seguros de qué es información, pero igualmente la medimos.

Un exabyte, es decir 10^{18} bytes, equivalente a 100.000 bibliotecas del Congreso de Estados Unidos, unidad de medida muy universal. Se habían producido entonces en toda la historia de la humanidad, hasta los '90, más o menos, 1.200.000 bibliotecas del Congreso de Estados Unidos.

Supongo que eran bibliotecas de libros físicos. Esa cantidad de información incluye toda la filmografía de Hitchcock, todo Shakespeare y sus copias, no sé, todo eso, todo lo que la revolución iniciada por Gutenberg pudo imprimir, bla, bla, bla. Sin embargo, para el 2006 se habían producido 180 exabytes, es decir que se multiplicó por 15 la producción en sólo 10 años. ¿Qué pasó? Internet, por ejemplo.

Se estima hoy una producción de 463 exabytes, es decir, el triple que hasta 2006, pero ahora ¡por día! Es decir, en este rato que vamos a estar acá, vamos a producir el doble o el triple de información, (el mundo en ese tiempo, no nosotros), que produjo toda la humanidad hasta los 90. Desde las cuevas de Altamira hasta Matrix.

En un poco menos de una horita, cuarenta minutos más o menos, produciremos esa cantidad de información. ¿Qué significa eso?, ¿y cómo se llega a eso? Bueno, de muchas maneras.

Pero volvamos al tema de los datos. Recordemos que un dato es cualquier cosa que un programa tome como tal. En particular un programa también es un dato. Puede ser un metadato, no en el sentido del metadato de bibliotecología, sino de datos sobre datos.

Una de las principales razones del enorme crecimiento en la producción de datos es que, en los primeros años del siglo XXI, los dispositivos conectados a Internet eran, en una gran mayoría, humanos, mientras que en la actualidad más del 80% de los dispositivos conectados son no-humanos.

Es importante también la evolución de los sistemas que reconocen patrones, con el *machine learning* como principal exponente. Estos sistemas se nutren de grandes volúmenes de datos y generan clasificadores a partir de ellos, lo cual hace que ciertos datos que eran in-significantes hasta hace poco, se vuelvan información relevante.

Siguiendo una idea de Mackay, podemos pensar a la información como una distinción que hace una diferencia. La distinción es cualquier dato hasta hace poco irrelevante pero que ahora colabora en la construcción de un clasificador estadístico capaz de generar nuevas diferencias. O también, como decía Bateson, la información es una diferencia que produce una diferencia, distinguiendo distinción de diferencia (o diferenciándolas). Esa idea de que cosas que eran diferencias, que no hacían diferencia, hoy hacen diferencia, porque hay sistemas capaces de producir esa diferencia.

El título de esta charla es una paráfrasis de un libro de Dennett: La peligrosa idea de Darwin.

Hay varias cuestiones que se cruzan acá, pero antes, parafraseando esta paráfrasis, hablemos de la peligrosa idea de Dennett

En el año 2006, Daniel Dennett propuso en el sitio "Edge" que diferentes filósofos, científicos, artistas, expusieran cuál era la idea que consideraban más peligrosa. Muchas de las respuestas son muy interesantes, pero quiero concentrarme en la propuesta del propio

Dennett. Esta idea plantea que (ya en 2006): “No hay suficientes mentes que puedan albergar la población explosiva de memes”.

Por memes Dennett entiende una extensión de la idea de Dawkins de unidad de reproducción cultural. Richard Dawkins es un biólogo que buscó el equivalente al gen en el ámbito cultural. La reproducción cultural se da porque los memes son entidades, unidades culturales que se reproducen usando los cerebros, y otros dispositivos cognitivos. Así como los genes son entidades de ADN que se reproducen usando las células.

Invierte un poco la perspectiva sobre el lugar de la agencia. No serían los humanos los principales productores de la cultura, sino que son los memes quienes se reproducen a través de los humanos.

Esto obviamente cobra nueva forma y toma nuevo impulso con los memes de internet y, de hecho, hay vínculos importantes entre ellos. Varios estudios abordan eso, por ejemplo, Juan Ruocco tiene un libro sobre los peligros de la democracia, intenta pensar la memética y los riesgos y los desafíos que plantea al sistema democrático.

Volviendo a la idea de Dennett, dice que los humanos tienen un hambre insaciable de memes, y todo el tiempo producen memes nuevos.

Y en algún momento ya no alcanzarán todas las mentes para albergar esa población creciente de memes. Quizá ese momento ya ocurrió. ¿Qué se hace con eso? Bueno, dilemáticamente hay dos posibilidades: Los humanos se ahogan en la inundación de información, o no se ahogan. Ambas alternativas son perturbadoras, dice Dennett.

¿Por qué serían perturbadoras ambas? Bueno, creo que estamos viendo muchos indicios de ahogo frente a la cantidad de información, y está ocurriendo lo que Dennett temía en esta alternativa: formas crecientes de inconmensurabilidad cultural... Cada vez más gente no comparte entre sí pautas culturales elementales.

La otra alternativa, que sería que los humanos sí aprendan a lidiar con la inundación informacional, con la siempre creciente cantidad de memes, implicaría, para Dennett, que los sucesores de la humanidad serían muy diferentes, ya sea que siguen siendo humanos o que sean robots o cyborgs. Desde ciertas perspectivas conceptuales, los humanos siempre fueron algún tipo de cyborg, pero se seguiría transformando esa condición.

Es posible que en este momento histórico se esté dando un poco de cada alternativa. Cada vez hay más inconmensurabilidad cultural y más incorporación tecnológica a las propias subjetividades. Las mediaciones algorítmicas ubicuas son, cada vez más, parte de cualquier subjetividad, individual o colectiva.

El tipo de mediaciones actuales, mayormente sustentadas en alguna medida en clasificadores estadísticos producidos por *machine learning*, incorpora formas de opacidad intrínsecas que es necesario pensar, si se quiere lidiar con ellas de alguna manera virtuosa, cultural y políticamente hablando. Esta, quizá, es la principal característica de lo que puede verse como una transformación antropológica en curso.

Dennett plantea que hay diferentes momentos evolutivos, que la idea de Turing puede ser vista en esa clave, en la transformación de la evolución *bottom-up* en ciertas posibilidades de diseño dirigido. Si bien no lo plantea directamente en esos términos, uno de los caminos más interesantes es el de la co-evolución de mentes y máquinas (y organizaciones colectivas), como lo plantea Ashford Lee. En cualquier caso, este marco conceptual muestra que hay diversos caminos alternativos al que hoy el *tecnocapital* nos lleva.

La última palabra en chino por hoy, la que designa a la inteligencia artificial. En chino

hay varias maneras de decir artificial y varias de decir inteligencia. Pero para este sintagma completo usan la siguiente:

人工智能

(Se pronuncia Rén gōng zhì néng)

Los dos primeros caracteres son respectivamente persona y trabajo. Lo artificial es lo producido por el trabajo humano. El tercer carácter es sabiduría, el último es poder, la capacidad para hacer algo. La sabiduría en acción es la inteligencia.

Se dice acá poder como el poder hacer, no como el poder establecido, aunque quizá también ahora estaría asociándose más a ese tipo de poder.

¿Qué es la inteligencia artificial? Dice Nietzsche que sólo puede definirse lo que no tiene historia, y la IA tiene una larga y confusa historia.

Decía Minsky, en 1968, que es la ciencia de hacer programas que puedan resolver problemas en los que las personas, los hombres decía él, requieran inteligencia para hacerlo.

O sea, de alguna manera, se despegaba del problema de encontrar una caracterización. Por otro lado, en 1977, decía Boden que la IA, también, es una manera de usar programas para entender la inteligencia humana. Herbert Simon proponía realizar progresivamente tareas cada vez más complejas, es decir lo veía ya como una forma de la programación. Dicen que uno de los motivos de usar este sintagma es que, en 1956, John McCarthy, el autor de Fortran y Lisp, entre otras cosas, quería inventar un nuevo nombre que les permitiera distinguirse de los cibernéticos.

No voy a abordar aquí el muy interesante pensamiento cibernético, que se desarrolló intensamente durante las conferencias Macy, entre 1946 y 1953, con la participación de numerosas disciplinas. Entre otros, Norbert Wiener, uno de los fundadores de la cibernética y quien difundió el término, planteaba una idea muy diferente a la inteligencia artificial. Podría decirse que se enfocaba más en explicaciones mecánicas de la mente que en construir máquinas que piensen, idea que sigue resonando, por ejemplo, en la cita de Boden que hice hace un rato. En el congreso de Dartmouth de 1956, cuando aparece el sintagma “inteligencia artificial”, el objetivo de McCarthy era que no los confundan con los cibernéticos. Turing tampoco habló de inteligencia artificial, sino de máquinas que piensan o incluso de máquinas que aprenden.

Hay un buen libro reciente de Kate Crawford, que se llama Atlas de la Inteligencia Artificial, donde ella va a hacer un análisis material de esta idea y de sus condiciones de producción. Escribe aquí una frase que se popularizó mucho en los análisis filosóficos y sociales de la IA: “La IA no es ni artificial ni inteligente”. Puede ser, pero voy a criticarla un poco, sobre todo por su posición humanista, demasiado humanista.

¿Por qué dice que no es artificial? Bueno, la justificación que da Crawford es que es corporizada y material. En la misma línea, Mateo Pasquinelli, otro autor que tiene un libro también muy interesante, *The Eye of the Master*, abunda en esta idea diciendo que no es artificial ya que se produce por un montón de trabajo humano, en general mal pago.

Por ejemplo, el hoy famoso chatGPT es un motor estadístico generado a partir de un entrenamiento con un billón de entradas lingüísticas. Pero además ha sido “corregido”, es decir, modificado su comportamiento, para que su interacción sea humanamente más agradable, por miles de programadores o revisores, mayormente trabajando en Kenia.

En cualquier caso, y como ya lo sabían los chinos que acuñaron la palabra para IA, que

algo sea corporizado, material y producto del trabajo humano, no hace que no sea artificial, al contrario. Para aclarar esto veamos la propuesta de uno de los filósofos de la computación que más me gusta: Bill Rapaport. Lo que él dice es que la idea de “artificial” tiene una connotación de algo no real, de un engaño, y que en ese sentido él prefiere usar la idea de “sintético”, inteligencia sintética. Más aún, propone que es mejor usar “computacional”, remitiendo a la forma de dicha síntesis.

También afirma Crawford que no es inteligencia, ya que sólo pueden tener conductas racionales, autónomas o de discernimiento a partir de un intensivo entrenamiento con grandes cantidades de datos. En esta línea, también Pasquinelli afirma que son en realidad un instrumento de magnificación del conocimiento, un nooscopio es el neologismo que acuñan. Todo esto es cierto, por eso prefiero el término que va usar Rapaport, que no es el de inteligencia sino el más general de “cognición”. Así, cuando hablamos de IA creo que la interpretación más precisa de lo que esta sigla menciona es “cognición computacional”.

Lo que hoy se llama IA está asociado sobre todo a la producción de sistemas con la técnica de *machine learning* o aprendizaje automático. Este tipo de tecnología, que es básicamente una manera de programar, se basa en la construcción de clasificadores estadísticos a partir del entrenamiento con grandes volúmenes de datos. Este entrenamiento (que se basa en programas “tradicionales”) produce un modelo estadístico de reconocimiento de patrones, que permite reconocer si nuevos datos satisfacen los criterios implícitos de los datos de entrenamiento, si son suficientemente similares. Los grandes modelos de lenguaje (LLM), como ejemplo paradigmático, permiten reconocer entidades lingüísticas correctas. A eso se suma, en estos casos, un generador de datos reconocibles por el clasificador. Tanto el generador como el clasificador se producen juntos. En la mayoría de los LLM, el *prompt* (es decir, la interacción generada por el usuario) le permite escoger alguna sub-red semántica que hace que la producción de textos refiera a algún tema determinado. Comprender esto permite ser más preciso al hablar de estos sistemas. No tiene sentido decir que un LLM miente, o que “alucina”, ya que lo único que hace es producir textos aleatorios asociados a algún tema.

Un intento interesante de asignar significado teórico a los LLM es el que decía un autor de la revista Jacobin, refiriendo al ejemplar más conocido: “ChatGPT es la ideología”. Es decir, permite revelar las creencias que la humanidad tiene sobre el mundo. Lo que intermedia con el mundo. La ideología puede pensarse de muchas maneras, aquí se la ve como un hiperobjeto, es decir, en términos de Morton, un objeto que no se adecua a la experiencia humana directa, pero para el que funciona como una interfaz humanamente accesible.

Creo que es importante no antropomorfizar estos sistemas, entenderlos en su especificidad cognitiva y no en relación a la cognición humana. Por un lado, permite esto comprender mejor de qué se trata; por otro, asumir una necesaria posición de humildad, dado que una tecnología bastante sencilla puede mostrar habilidades que hasta hace poco se creían privativas de la cognición humana.

Luciano Floridi, un prolífico filósofo de la información, dice que está ocurriendo un desplazamiento de la computación, de ser considerada una empresa lógica a devenir una empresa mayormente estadística.

Esta transformación no elimina lo anterior, pero cambia la visión y las expectativas. Incluso la palabra algoritmo, de la jerga científica hasta hace poco, llegó al dominio público, pero con un sentido muy diferente al original, remitiendo casi siempre a los clasificadores estadísticos. Para colmo de males, las propiedades específicas de este tipo de programas suelen extrapolarse a cualquier otro, cayendo en una insidiosa falacia de composición.

Cuando se habla de sesgos, de repetición del pasado, como veíamos antes de “alucinaciones”, incluso de racismo de los algoritmos, hace falta relativizar mucho el alcance de estas propiedades. Son sólo propiedades predicables de un tipo muy particular de programas y además en entornos específicos, lo que suelo llamar ensamblajes computacionales. Este tipo de propiedades no se aplican a los programas en general. Podría decirse que no hay propiedades de cualquier algoritmo o programa sino meta-propiedades, cuya actualización, en casos específicos, dependerá de múltiples factores, uno de ellos, importante claro, será la definición del programa mismo.

Uno de los problemas más graves de esta falacia, es que da lugar a una mirada determinista de los sistemas algorítmicos y obtura la capacidad de pensar alternativas. Como ejemplo de esto, el muy interesante y útil concepto de gubernamentalidad algorítmica de Antoinette Rouvroy y Thomas Berns tiende a pensarse como una propiedad de cualquier mediación algorítmica (a partir de su propia enunciación), lo que parece anular la posibilidad de pensar alternativas.

En contextos de teoría política, suele afirmarse livianamente que las computadoras o los sistemas computacionales son de la derecha, favorecen el neoliberalismo. ¡No! No hay ninguna evidencia conceptual ni histórica de eso. En el presente, las corporaciones informacionales aprovechan su posición ventajosa para desarrollar tecnologías que afianzan esa posición, pero asignarles esa propiedad es tomar el efecto por la causa, o una contingencia histórica como ineluctable fatalidad.

En otra charla que daré esta tarde, me extenderé más sobre un tema acuciante, que es la manera en que se entienden estas tecnologías desde las organizaciones políticas del campo popular. Alguna gente que integra estas organizaciones suele contrastar la “verdadera” política en las calles con la “política en redes”, realizada por trolls y pseudo-militantes. Creo que hay aquí una confusión y un caso flagrante de *wishful thinking*. La calle, la ciudad, los territorios están constituidos, hoy también, por un uso creciente de mediaciones computacionales, una pila de capas digitales, en palabras de Bratton. Poder hacer política en el siglo XXI requiere entender estas mediaciones y diseñarlas o transformarlas desde propuestas y estrategias políticas.

Para concluir, algunas reflexiones sobre cómo comprender y proyectar mejores mediaciones computacionales. Primero quiero citar esta divertida advertencia del informático (y filósofo, a pesar suyo) neerlandés Edsger Dijkstra en 1984:

The Fathers of the field had been pretty confusing: John von Neumann speculated about computers and the human brain in analogies sufficiently wild to be worthy of a medieval thinker and Alan M. Turing thought about criteria to settle the question of whether Machines Can Think, a question of which we now know that it is about as relevant as the question of whether Submarines Can Swim

Los padres del campo han sido muy confusos. John von Neumann especuló acerca de las computadoras y del cerebro humano, en analogías suficientemente disparatadas como para ser propias de un pensador medieval. Y Alan Turing se preguntaba acerca de criterios para determinar la cuestión de si las máquinas podían pensar, una cuestión que ahora sabemos es tan relevante como preguntarse si los submarinos pueden nadar.

¿Qué es lo que dirá Dijkstra entonces acerca del rol de las computadoras?

Las computadoras no piensan, hacen mejores cosas que eso. ¿Por qué imitar la mente humana si se puede “imitar” algo mejor?

Creo, con Kevin Kelly uno de los editores de la revista Wired, quien decía que la misión de la IA no es hacer una mente humana, sino explorar el espacio de posibles inteligencias.

Volviendo a Bratton, es necesario entender que la creciente digitalización del mundo cambia, constituye nuevos ambientes técnicos fértiles en posibilidades de acción. Lo digital es una nueva forma de materialidad que se superpone a formas previas de organización de la materia, y como tal habilita sofisticadas formas de acción en el mundo y por lo tanto nuevas lógicas para comprender el mundo y pensar las acciones. La transformación es radical y sigue en curso, cada vez más veloz.

Dice Mark Hansen, un filósofo simondoniano:

In concert with researchers like Andy Clark and N. Katherine Hayles, I embrace the “technical distribution” of cognition and perception as a way of understanding the complex couplings between humans and machines that are typical in our contemporary world, but that have, in fact, been part of human techno-genesis since the very origin of the human.

En la misma línea con investigadores como Andy Clark y N. Katherine Hayles, acepto la “distribución técnica” de la cognición y la percepción como una forma de entender los complejos acoplamientos entre humanos y máquinas que son típicos en nuestro mundo contemporáneo, pero que, de hecho, han sido parte de la tecnogénesis humana desde el origen mismo del ser humano.

Suele declamarse, desde cierto humanismo anacrónico: “Bueno, pero finalmente la tecnología la hacen los humanos”. No exactamente, fue la tecnología la que hizo a los humanos.

Había tecnología antes que humanos. Los australopithecus tenían tecnología, pero aún no lenguaje. Y la interioridad de esto que creemos que es nuestra mente y que de alguna manera es capaz de percibir y perseverar fue una introyección de esa instrumentalidad que tenían estos homínidos y que no tenían un lenguaje desarrollado.

Una última cita de Dijkstra de 1972:

Automatic computers have now been with us for a quarter of a century. They have had a great impact on our society in their capacity of tools, but in that capacity their influence will be but a ripple on the surface of our culture, compared with the much more profound influence they will have in their capacity of intellectual challenge without precedent in the cultural history of mankind.

Las computadoras automáticas ya llevan un cuarto de siglo entre nosotros. Han tenido un gran impacto en nuestra sociedad en su función como herramientas, pero en ese sentido su influencia será apenas una ola en la superficie de nuestra cultura, en comparación con la influencia mucho más profunda que tendrán en tanto desafío intelectual sin precedentes en la historia cultural de la humanidad.

Este desafío está cobrando cada vez más importancia y es necesario abordarlo plenamente.

Notas

¹ Rodríguez Freire Raúl (2018). La condición intelectual Informe para una academia. Editorial Mímesis.

² Actualmente el doctor Blanco es profesor e investigador de la Facultad de Matemática, Astronomía Física y Computación de la Universidad Nacional de Córdoba. Además, es profesor en la Universidad Nacional de Río Cuarto (UNRC). Integra el equipo docente de la Diplomatura Superior en Tópico de Ciencias de Datos de la Facultad de Ciencias Exactas, Físico-Químicas y Naturales de la UNRC. Es director alterno del Doctorado en Ciencias de Computación y director de la maestría en Tecnología, Políticas y Cultura de la Universidad Nacional de Córdoba.

Sus áreas de interés actuales incluyen filosofía y computación, filosofía de la información y filosofía de la técnica. Cuenta con numerosas publicaciones en revistas especializadas de computación y filosofía de la técnica. Participó en los libros colectivos *Cosas Confusas*, *Tirant Blanc 2019*, *Tecnología, Política y Algoritmos en América Latina de Cenaltes 2020*, *Individualización, Tecnología y Formación*, *Simodon en debate*, *Aula 2020* y *Glosario de Filosofía y la Técnica*, *La Zebra 2022*.