

Aceleración tecnológica e inteligencia artificial. ¿Hasta dónde podríamos cambiar?*

Technological acceleration and artificial intelligence.
How much could we change?

Aceleração tecnológica e inteligência artificial. Até onde podemos mudar?

Felipe Arocena**

Universidad de la República, Montevideo, Uruguay

Sebastián Sansone***

Universidad de la República, Montevideo, Uruguay

Cómo citar: Arocena, F., y Sansone, S. (2022). Aceleración tecnológica e inteligencia artificial. ¿Hasta dónde podríamos cambiar? *Revista Colombiana de Sociología*, 45(2), pp 301-326.

DOI: <https://doi.org/10.15446/rsc.v45n2/89851>

Este trabajo se encuentra bajo la licencia Creative Commons Attribution 4.0.

Artículo de reflexión

Recibido: 10 de agosto del 2020 Aprobado: 26 de abril del 2021

* Este trabajo se desarrolló en el marco del Proyecto de Investigación: “La cultura en el mundo de la aceleración tecnológica”, Departamento de Sociología, Facultad de Ciencias Sociales, Universidad de la República, Uruguay; <https://cienciassociales.edu.uy/departamentos/departamento-de-sociologia/>

** Doctor en Ciencias Humanas, Profesor Titular, Universidad de la República, Uruguay
Correo electrónico: felipe.arocena@cienciassociales.edu.uy-ORCID: <https://orcid.org/0000-0002-1666-8449>

*** Aspirante a magister en Sociología, Profesor Asistente, Universidad de la República, Uruguay
Correo electrónico: sebastian.sansone@fder.edu.uy-ORCID: <https://orcid.org/0000-0003-0174-0265>

Resumen

¿Cuán rápido nos estamos moviendo en el siglo xxi? ¿qué tan rápido nos movimos en el pasado? ¿cómo podemos captar con alguna objetividad la aceleración tecnológica que, por supuesto, termina impactando cada dimensión de la vida social? ¿hasta dónde podemos proyectar con fundamentos razonables el desarrollo de la inteligencia artificial impulsada por la aceleración tecnológica? En este trabajo ensayaremos algunas respuestas a estas preguntas apoyándonos en la reflexión de tres reconocidos especialistas como Nick Bostrom, Max Tegmark y Ray Kurzweil, quienes provienen de campos disciplinares diversos como la filosofía, la física y el futurismo. Es cierto, todos parece que sentimos esa aceleración de la vida en paralelo con la velocidad de la tecnología, pero ¿cuán rápido vamos? ¿cómo podemos compararnos con el pasado? ¿cuándo empieza ese vértigo? y, más interesante aún, ¿hasta dónde puede llegar esa aceleración? Para responder estas preguntas seguimos el método tradicional, dialogamos con expertos que han hecho de ellas su objeto de estudio. Es por eso por lo que en este trabajo construimos un análisis interdisciplinario con un físico, un filósofo y un inventor, tres reconocidos pensadores que expanden los horizontes convencionales y nos ayudan a intuir ese futuro singular en ciernes, tan impredecible en sus formas y figuras concretas, como predecible en su proyección por las tendencias de la aceleración tecnológica. Piense el lector que el dispositivo móvil ahora en su bolsillo para comunicarse tiene un millón de veces más memoria y es cien mil veces más rápido que la computadora que nos puso en la luna. Y esta es la tendencia de la exponencialidad de la evolución de las tecnologías de la información, que se duplican en capacidad cada un año y ocho meses desde hace un siglo. ¿Qué tecnología llevarán sus hijos en los bolsillos, o más bien implantada en sus cuerpos, dentro de cincuenta años?

Palabras clave: cambio social, ética, futuro, superinteligencia, tecnología.

Descriptor: ciencia y sociedad, ética de la ciencia, inteligencia artificial, sociedad futura.

Abstract

How fast are we moving in the 21st century? How fast did we move in the past? How can we capture with some objectivity the technological acceleration that, of course, ends up impacting every dimension of social life? How far can we reasonably project the development of artificial intelligence driven by technological acceleration? In this work we will test some answers to these questions based on the reflection of three renowned specialists such as Nick Bostrom, Max Tegmark, and Ray Kurzweil, who come from diverse disciplinary fields as philosophy, physics, and futurism. It is true, we all seem to feel that acceleration of life in parallel with the speed of technology, but how fast are we going? how can we compare with the past? when does that vertigo start? And, more interestingly, how far can that acceleration go? To answer these questions we follow the traditional method, we talk with experts who have made them their object of study. That is why in this work we build an interdisciplinary analysis with a physicist, a philosopher, and an inventor, three renowned thinkers who expand the conventional horizons and help us to intuit that singular future in the making, as unpredictable in its forms and concrete figures, as predictable in its projection by the trends of technological acceleration. Let the reader think that the mobile device now in your pocket to communicate has a million times more memory and is a hundred thousand times faster than the computer that put us on the moon. And this is the trend of the exponentiality of the evolution of information technologies, that has doubled in capacity every year and eight months for a century. What technology will your children carry in their pockets, or rather implanted in their bodies, in fifty years?

Keywords: ethics, future, social change, superintelligence, technology.

Descriptors: artificial intelligence, ethics of science, future society, science and society.

Resumo

Quão rápido estamos nos movendo no século 21? Quão rápido nos movemos no passado? Como capturar com alguma objetividade a aceleração tecnológica que, claro, acaba impactando todas as dimensões da vida social? Até onde podemos projetar razoavelmente o desenvolvimento da inteligência artificial impulsionado pela aceleração tecnológica? Neste trabalho iremos testar algumas respostas a estas questões com base na reflexão de três especialistas de renome como Nick Bostrom, Max Tegmark e Ray Kurzweil, oriundos de diversas áreas disciplinares como a filosofia, a física e o futurismo. É verdade, todos nós parecemos sentir essa aceleração da vida em paralelo com a velocidade da tecnologia, mas quão rápido estamos indo? como podemos comparar com o passado? quando começa essa vertigem? E, mais interessante, até onde essa aceleração pode ir? Para responder a essas perguntas seguimos o método tradicional, conversamos com especialistas que fizeram deles seu objeto de estudo. É por isso que neste trabalho construímos uma análise interdisciplinar com um físico, um filósofo e um inventor, três pensadores renomados que ampliam os horizontes convencionais e nos ajudam a intuir esse futuro singular em construção, tão imprevisível em suas formas e figuras concretas, como previsível em sua projeção pelas tendências de aceleração tecnológica. Deixe o leitor pensar que o dispositivo móvel agora em seu bolso para se comunicar tem um milhão de vezes mais memória e é cem mil vezes mais rápido que o computador que nos colocou na lua. E esta é a tendência da exponencialidade da evolução das tecnologias de informação, que duplicaram de capacidade a cada um ano e oito meses durante um século. Que tecnologia seus filhos levarão no bolso, ou melhor, implantada em seus corpos, daqui a cinquenta anos?

Palavras-chave: ética, futuro, mudança social, superinteligência, tecnologia.

Descritores: ciência e sociedade, ética da ciência, inteligência artificial, sociedade.

Introducción

Comencemos por definir brevemente qué es la aceleración tecnológica, de acuerdo con el filósofo y sociólogo alemán Harmut Rosa:

Es el aumento deliberado de velocidad de los procesos orientados a metas específicas del transporte, la comunicación y la producción que, en su conjunto, pueden denominarse *aceleración tecnológica* [...] Esta última puede ser definida como un incremento de producción total por unidad de tiempo; es decir, en el número de kilómetros cubiertos en una hora, o en bytes de datos transferidos por minuto, o en el número de automóviles producidos por día. (Rosa, 2016, pp. 21-34)

Naturalmente la *aceleración tecnológica* tiene consecuencias de todo tipo en la vida social, pero existen otras dimensiones relevantes desde las que podemos aproximarnos al concepto de aceleración que el propio Rosa analiza en su obra *Alienación y aceleración* (2016), por ejemplo, la aceleración social o la aceleración del ritmo de vida. En otro trabajo (Arocena y Sansone, 2020) hemos discutido su teoría de la aceleración, en este en particular nos dedicaremos a entender específicamente en qué etapa estamos de la aceleración tecnológica y su probable evolución en la inteligencia artificial (IA) en el corto y largo plazo.

Si queremos comprender la aceleración en la innovación tecnológica, lo primero que debemos abordar es la Ley de Moore. Gordon Moore descubrió en 1965 que la potencia de la computadora se duplicaba cada año porque la cantidad de elementos que podían ensamblarse en un circuito integrado se multiplicaba por dos cada año. Luego corrigió su estimación una década más tarde y se dio cuenta de que el poder de los ordenadores se duplicaba cada dos años; poco después de esto, se acepta ampliamente que la duplicación es cada dieciocho meses (año y medio) y se mantiene así desde entonces.

La velocidad de la miniaturización es una historia paralela. A principios de los años setenta, un semiconductor tenía el tamaño de diez micrómetros, mucho más delgado que un cabello humano. En 1985 era diez veces más delgado, y luego a principios de siglo se redujo a cien nanómetros... y se sigue contando. Esta aceleración puede describirse en cualquier gráfico que muestre cómo se mueven los datos exponencialmente: “tan lejos, tan hacia arriba, tan hacia la derecha” (Bridle, 2018, p. 81). Bridle argumenta que la Ley de Moore, en lugar de una ley, es una proyección y que llegamos a un punto en el que podríamos experimentar demasiado ruido y como resultado del “código desechable e ineficiente y características redundantes, la velocidad del software se reduce a la mitad cada dieciocho meses” (Bridle, 2018, p. 81), proceso conocido como Ley de Gate; lo que podría entenderse como una especie de némesis de la ley de Moore.

En la actualidad, sin embargo, muchos científicos prestigiosos creen que la proyección de la tasa de poder tecnológico de Moore es correcta (o al menos no refutable), así como lo demuestra el fallecido Stephen Hawking.

Imagínese, escribió en su libro póstumo *Breves respuestas a las grandes preguntas* (2019), que las computadoras continúan ganando energía a este ritmo solo por un siglo más (un parpadeo en el tiempo histórico). Es muy probable que, si esto sucede, una explosión de inteligencia artificial surgirá dentro de este marco de tiempo. Esto es así porque se producirá una mayor aceleración ya que una inteligencia mayor, superior, o que sobrepase la humana, presumiblemente tendrá la capacidad de mejorar y mejorar recursivamente. El físico inglés, tal vez el científico más reconocido desde el brillo de Einstein, afirma, además, que:

si los ordenadores continúan siguiendo la Ley de Moore, duplicando su velocidad y su capacidad de memoria cada dieciocho meses, el resultado será que los ordenadores probablemente adelantarán a los humanos en inteligencia en algún momento en los próximos cien años. (Hawking, 2018, p. 228)

Para apoyar esta idea tiene en cuenta las últimas innovaciones en computación cuántica, con las cuales se producirá una nueva aceleración exponencial “cambiando todo”; por supuesto, incluyendo la transformación de la vida tal como la conocemos y la biología. A propósito de esto último hay que resaltar que ya la larga, lenta y milenaria cadena de la vida de la selección natural darwiniana se está viendo alterada a medida que avanza la selección artificial gracias a la decodificación del genoma humano completo y a la técnica de edición de genes de Crispr¹, que está totalmente disponible.

Solo una ayuda más para tratar de comprender qué tan rápido nos estamos moviendo en la aceleración tecnológica. Ray Kurzweil (*La singularidad está cerca*, 2012, original de 2005) afirma que durante la mayor parte de la historia de la humanidad el imaginario dominante sobre los procesos evolutivos fue lineal, mientras que con las innovaciones recientes en *infotech*² se ha adquirido un patrón exponencial cada vez más rápido. La primera tecnología de información registrada por la historiografía que desarrollaron los humanos fue el lenguaje, que tardó varios cientos de miles de años en aparecer, luego la escritura se desarrolló hace diez mil años, la imprenta demoró cuatrocientos años, el teléfono solo necesitó cincuenta años para expandirse masivamente, el teléfono celular lo hizo en siete años, y teléfonos inteligentes en tres. Si bien no podemos predecir el futuro, la firme convicción de Kurzweil es que sí podemos predecir que la *infotech* evolucionará exponencialmente de una manera predecible, como lo ha hecho en el pasado reciente. Un teléfono inteligente es mil millones de veces más poderoso por dólar que una computadora de hace sesenta años, una como la que nos puso a los humanos en la luna.

-
1. Crispr: proviene de las siglas en inglés *Clustered Regularly Interspaced Short Palindromic Repeats*. La tecnología Crispr es una herramienta molecular utilizada para editar o corregir el genoma de cualquier célula.
 2. Información y tecnología. En el habla anglosajón es común la utilización. En el mundo hispanohablante no está tan arraigado.

Esta aceleración exponencial en *infotech* y en el significativo crecimiento del poder de las computadoras por dólar (computadoras millones de veces más rápidas y otro tanto más baratas respecto a las que nos permitieron salir de nuestro planeta), continuarán provocando cambios extraordinarios en todos los dominios de nuestras vidas, incluida la salud, la esperanza de vida y la aparición de inteligencia artificial, tal vez primero de nivel humano hasta alcanzar la *superinteligencia*. Esta singularidad aparecerá cuando la curva exponencial del progreso en *infotech* alcance su impulso vertical con tal velocidad que el mundo tal como lo conocemos dejará de existir. La “rodilla” de la curva exponencial del progreso tecnológico, cuando se vuelve casi vertical, podría ocurrir a mediados de este siglo (Kurzweil, 2012). Si bien es discutible que esta tendencia pueda continuar mucho tiempo más, la posición de Kurzweil es que hay más evidencia que la confirmaría antes que refutarla. Los recursos en el planeta Tierra en esta perspectiva no serían un límite porque la fase de singularidad incluye la colonización del espacio exterior (Moravec, 1998; Joy, 2000; Schwab, 2016; Kelly, 2017; Sadin, 2018; McEwan, 2019; Arocena, 2017)

Hemos definido resumidamente qué entendemos por aceleración tecnológica, toca ahora anticipar a qué nos referimos con inteligencia artificial. Esto es más complicado porque existen diversas maneras de entender qué es la inteligencia. Hay personas muy inteligentes para determinadas cosas y nulas para otras, hay inteligencia emocional e inteligencia racional, hay inteligencia para ser creativo estéticamente o para resolver teoremas matemáticos, hay personas muy inteligentes para fabricar cosas con sus manos y hay otras que no pueden clavar un clavo, pero pueden memorizar el *Martín Fierro*, de Hernández. Aunque sea muy complejo definir *inteligencia* y el ejercicio de la misma, al menos tratemos de ser honestos con el idioma en el cual escribimos. El Diccionario de la Real Academia Española [RAE] nos ofrece, en la entrada *inteligencia*, varias acepciones: a) “la capacidad de entender o comprender”, b) “la capacidad de resolver problemas”, c) “conocimiento, comprensión, acto de entender”, d) “habilidad, destreza y experiencia”, e incluye además e) “sustancia puramente espiritual”³. Así mismo, la RAE nos proporciona una definición mínima, aunque aceptable, de la *inteligencia artificial*: “disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana como el aprendizaje o el razonamiento lógico” (RAE, s.f.). Max Tegmark coincide con esta definición esencial y propone que podemos entender la inteligencia como “la capacidad de alcanzar objetivos complejos” (2018, p. 70), y entre los objetivos complejos podemos incluir varias de las dimensiones mencionadas. A su vez Tegmark complementa la definición previa de inteligencia artificial sosteniendo que: “la inteligencia artificial actual suele ser estrecha: cada sistema solo es capaz de lograr objetivos muy específicos, mientras que la inteligencia humana es extraordinariamente amplia” (p. 105).

3. Véase <https://dle.rae.es/inteligencia?m=form>

En los próximos apartados profundizaremos sobre estas definiciones de aceleración tecnológica e inteligencia artificial. Discutiremos cómo estos conceptos son pensados por tres reconocidos especialistas provenientes de disciplinas diversas: Bostrom, un filósofo; Tegmark, un físico; y Kurzweil, un inventor futurista. ¿Cómo proyectan ellos la aceleración tecnológica y cómo evalúan las maneras en que la inteligencia artificial podría llegar a afectarnos en el futuro?

Nick Bostrom: hacia la *superinteligencia artificial*

Bostrom advierte sobre su obra *Superinteligencia*: “no es un libro sobre la aceleración tecnológica” (2018, p. 1), ni tampoco sobre la posibilidad de la emergencia de una singularidad en el sentido de “los escritos de Ray Kurzweil y otros” (p. 2). Sin embargo, sí es un análisis “de la idea relacionada con la singularidad [...] de una explosión de inteligencia, especialmente la perspectiva de una superinteligencia artificial” (p. 2). La convicción de Bostrom es que hay razones fundadas para tomarse en serio esta última posibilidad y a demostrar esa afirmación dedicará todo el libro. No es el único que llega a esta conclusión (Baum, et ál., 2011; Tegmark, 2018, Hawking, 2018).

Una IA de nivel humano podría definirse como aquella: “que pueda desempeñar la mayoría de las profesiones humanas al menos igual de bien que un ser humano típico” (Bostrom, 2018, p. 19). Los principales expertos en este tema estimaron que, si no hay algún acontecimiento extraordinario que interrumpa la innovación científica, llegaríamos a este nivel de IA de nivel humano en relativamente poco tiempo. Para el 90 % de los expertos consultados se podría alcanzar antes del año 2075 y para la mitad se alcanzaría antes de 2040. Con todas las dudas que puedan emitirse en contra de estas predicciones y tomando en cuenta que tantas veces otros científicos se han equivocado antes, esta innovación parece estar muy, muy cercana. En particular uno de los investigadores más renombrados en escribir la historia de la IA, Nils Nilsson (2009), respondió de manera muy congruente con las estimaciones anteriores, ya que para él hay 10 % de probabilidades de que se llegue a este nivel antes de 2030, 50 % de que se llegue antes de 2050 y 90 % de que se alcanzará antes de 2100 (Bostrom, 2018, p. 19).

Si se llegara al nivel anterior de IA general, como se le llama al nivel de un sistema capaz de realizar la mayoría de las profesiones con el mismo nivel de eficiencia que los humanos, esta no sería la última parada en la innovación tecnológica, la próxima sería la creación de un nivel de IA todavía superior, es decir una superinteligencia artificial. El propio Bostrom está bastante convencido de que el tren de la innovación se detendrá poco tiempo en la estación de nivel humano de IA y velozmente evolucionará a un nivel infinitamente superior en un salto de singularidad. Una cita ya clásica del matemático I.G. Good, quien fuera parte del equipo de Alan Turing, ilustra la razón de este salto:

Definamos una máquina ultrainteligente como aquella que puede superar con creces todas las actividades intelectuales de cualquier

hombre por muy listo que sea. Puesto que el diseño de máquinas es una de esas actividades intelectuales, una máquina ultrainteligente podría diseñar máquinas incluso mejores; entonces habría, sin duda, una ‘explosión de inteligencia’, y la inteligencia humana quedaría muy atrás. Por ello, la primera máquina ultrainteligente es el último invento que el hombre necesita crear, contando con que la máquina sea lo suficientemente dócil como para decirnos cómo mantenerla bajo control. (Bostrom, 2018, p. 4)

Good todavía tenía su imaginario restringido a máquinas ultrainteligentes, ya que la fusión entre tecnologías de información y la biología aún no se había producido. ¿Cuáles podrían ser los caminos para que se produzcan salto de inteligencia artificial semejante? Bostrom propone cinco trayectos plausibles.

Un primer recorrido posible⁴ para lograr un sistema con inteligencia general de nivel humano es a través de la evolución del poder computacional en asociación con el descubrimiento cada vez más sofisticado y acelerado de cómo funciona un cerebro humano. Una vez que se decodifiquen cuáles son las características esenciales del funcionamiento cerebral, estas podrían replicarse en una máquina capaz de desempeñarse como tal. En este caso el paso inicial sería crear un sistema seminal de inteligencia artificial de nivel general, cuya habilidad principal sería su capacidad de mejorarse a sí mismo aprendiendo no solo en sus contenidos (software) y mejorándolos, sino además en su estructura (hardware) y reformulándola. Estas máquinas tendrían “sentido común y una capacidad eficiente para aprender, razonar y planificar para afrontar complejos retos de procesamiento de información a lo largo de un amplio rango de ámbitos naturales y abstractos” (p. 3). Pero no debemos caer en el error de querer antropomorfizar estos sistemas. Lo anterior, de lograrse, utilizaría estructuras seguramente muy alejadas de lo que entendemos como una mente humana y no tendría sentimientos de amor, odio o celos como los de los sapiens. Aunque tampoco habría que descartar que uno de sus múltiples objetivos fuera replicarlos. El avance del poder de cálculo computacional y el avance en el descubrimiento de cómo funciona el cerebro, ambos potenciándose entre sí, hacen factible, con todas sus indeterminaciones e incertidumbres, esta trayectoria hacia una máquina que sea una semilla de inteligencia artificial de nivel general capaz de mejorarse recursivamente a sí misma indefinidamente. Este es el primer camino posible analizado por Bostrom hacia la superinteligencia, es decir, hacia la aparición de una inteligencia artificial superior a la humana en todas las dimensiones de la inteligencia ya mencionadas en la introducción. Pero hay otros itinerarios posibles, un poco más alejados del anterior, sobre el futuro de la inteligencia artificial.

4. Específicamente nos basamos aquí en el segundo capítulo del libro de Bostrom (2018): “Caminos hacia la superinteligencia”.

Un segundo sendero factible para alcanzar una inteligencia artificial general de nivel humano es bastante diferente al anterior y se basa en lograr escanear el cerebro humano con suficiente precisión para que una computadora pueda generar las estructuras neuronales. Mediante técnicas de laminación del cerebro, el escaneado microscópico y la reconstrucción en tres dimensiones en un ordenador, se lograría la emulación completa de un cerebro humano que quedaría subido en la nube. Este camino es una copia directa de la naturaleza ya que es el cerebro real que se reproduciría en un ordenador, con todas sus partes, conexiones y funcionamiento neuronal. “Hay buenas razones —nos dice Bostrom— para pensar que las tecnologías que harán posible lo anterior son alcanzables, aunque no en un futuro próximo” (2018, p. 31). Se trata de un futuro no muy cercano que, en realidad, parece mañana mismo ya que la estimación es que “las capacidades previas requeridas podrían estar disponibles a mediados de siglo, aunque con un gran intervalo de incertidumbre” (Bostrom, 2018, p. 34). Treinta años apenas.

Un tercer recorrido hacia una inteligencia artificial superior a la humana actual está balizado por la ingeniería genética. El proceso en este carril se daría a través de la mejora y amplificación de la inteligencia humana biológica. Bostrom escribe casi al mismo momento en que la tecnología de edición genética Crispr se estaba inventando, aunque aún no había sido dada a conocer públicamente y por lo tanto no la cita en su libro. Aun sin conocer este salto enorme en la capacidad de editar genéticamente embriones humanos que se traducirán en mutaciones permanentes mediante selección artificial, ya el autor anticipaba lo que podría ocurrir. Si bien descarta un plan eugenésico de la población en el que se obligue a determinados individuos con ciertas propiedades intelectuales superiores a que se crucen con otros similares porque esto chocaría con nuestro sistema de valores morales y políticos (lo que fue un deseo explícito del nazismo, pero antes también del darwinismo social positivista del siglo XIX que se planteaba el mejoramiento de la especie), hay otras maneras de lograr este objetivo. La edición genética es el más probable. Los avances en la decodificación del genoma humano y los avances de la ingeniería genética hacen viable en un futuro cercano que:

un embrión podría ser diseñado con la combinación preferida exacta de rasgos genéticos de cada padre. Los genes que no están presentes en ninguno de los padres también podrían ser empalmados, incluyendo los alelos que están presentes de manera poco frecuente en la población, pero que pueden tener efectos positivos significativos en la cognición. (Bostrom, 2018, pp. 40-41)

En poco tiempo esta selección artificial arribaría a un incalculable mejoramiento de las capacidades cognitivas del propio cerebro biológico. El lapso de tiempo para que esto pueda ocurrir también es de unas pocas décadas:

si tenemos en cuenta escenarios que se prolonguen hasta la segunda mitad de este siglo y más allá, hay que tener en cuenta la aparición probable de una generación de poblaciones genéticamente mejoradas —votantes, inventores, científicos— con la magnitud de la mejora incrementándose rápidamente en décadas posteriores. (Bostrom, 2018, p. 44)

Esa población genéticamente mejorada probablemente ya no sea de nuestra especie sapiens y “no hay ninguna razón para suponer que el homo sapiens haya alcanzado el vértice de la eficacia cognitiva para un sistema biológico” (Bostrom, 2018, p. 44)⁵.

Un cuarto camino por el que podría alcanzarse la inteligencia artificial de nivel humano o la superinteligencia superior a la humana es el de la interfaz entre ordenadores y cerebros, entre lo biológico y lo tecnológico, el camino de los ciborgs. Si bien los avances en este sentido han sido también notables y no se vislumbran límites en la integración de la tecnología en el cuerpo humano, Bostrom no percibe que por este lado se pueda llegar a una inteligencia artificial superior. Las razones que da son varias y entre las más importantes está que esta interfaz es extraordinariamente poderosa para ayudar a personas con dificultades físicas como la sordera, la ceguera, o con problemas motrices, o con enfermedades como el párkinson, cáncer o alzhéimer; pero en personas sin alguno de estos problemas someterse a implantes e interfaces muy agresivas de tecnologías avanzadas puede resultar demasiado costoso para el propio individuo. “A pesar de estas reservas, la ruta ciborg hacia la mejora cognitiva no es del todo sombría” (Bostrom, 2018, p. 47), sostiene el autor, pero aún no le ve la misma potencialidad que a las anteriores.

El quinto y último camino factible hacia la superinteligencia artificial es el de las redes colectivas. En este caso lo emergente sería el crecimiento hacia una inteligencia colectiva de nivel superior. Naturalmente aquí el punto de partida es la red de redes, o sea la internet, a la cual ya miles de millones de personas están conectadas diariamente. No es descabellado imaginar que las infinitas mejoras de la red tanto en algoritmos y en información, como en las máquinas cada vez más potentes que están conectadas, puedan terminar produciendo, junto a protocolos de interacción cada vez mejores, algo así como una inteligencia unificada emergente. El buscador de Google es hoy probablemente el mayor sistema de inteligencia artificial. Ya hace varias décadas el científico Vernor Vinge (1993) habló de la posibilidad de que internet “despertara” y en caso de que algo así ocurriese, en ello consistiría la “singularidad tecnológica”, término aparentemente creado por

5. Dentro del camino biológico de selección artificial de la especie una variante es la de la clonación de individuos con características cognitivas sobresalientes, lo cual parece más improbable por ahora. Aclaremos que la clonación de sapiens y la edición genética cuyos efectos sean mutaciones que pasarán a generaciones futuras, están prohibidas legalmente en muchos países así como entre la propia comunidad científica.

él. Ese despertar sería la emergencia de algún tipo de inteligencia unificada capaz de sintetizar y aprovechar de alguna manera el poder cognitivo de los miles de millones de usuarios y el poder computacional de un número de máquinas similar conectadas individualmente o en subredes.

Sea a través del aumento del poder de sistemas inorgánicos (máquinas), u orgánicos (biológicos), sea a través de una interfaz entre tecnologías informáticas y organismos biológicos, si no hay alguna disrupción que interrumpa la actividad científica y la innovación, Bostrom ve plausible que se llegará a la creación de inteligencia artificial, primero de nivel humano, luego de un nivel superior al humano. Hay probabilidades de que esto pueda ocurrir en el lapso del siglo XXI, lo podrían ver nuestros nietos y, tal vez, también muchos de nuestros hijos. Esta es la conclusión más importante del libro de este experto, filósofo respetado por todos los gurús en innovación tecnológica.

Las características de este sistema de inteligencia artificial no pueden predecirse. No sabemos cuánta empatía podrá tener con los *sapiens*, así como tampoco sabemos si lograremos que se aliñe con algunos de los principios o valores del humanismo, por ejemplo, y tal como se puede apreciar en los códigos morales de las religiones universales o con los valores más caros para occidente con la libertad individual, la igualdad y la solidaridad. No lo sabemos porque no podemos antropomorfizar un sistema de ese tipo cuya distancia en la percepción del mundo podrá ser tan larga como la de un ser humano y un piojo de oveja, y cuyos objetivos nos son inimaginables. Esta inteligencia artificial ¿desarrollará algo parecido a una conciencia de sí?, ¿cómo valorará a nuestra especie? ¿se expandirá fuera del planeta?, ¿qué debemos y podemos hacer ante esta proyección?

Es fácil negar la posibilidad de que algo de lo anterior pueda ocurrir y argumentar que todos estos caminos son apenas conjeturas irresponsables de ciencia ficción sin fundamentos suficientes que, en el fondo, tienen como objetivo propagar el miedo civilizatorio para facilitar la dominación. A ello diremos que esta es una posición demasiado necia y hasta parece pecar de obtusa. Por más incierta que pueda resultar esta posibilidad de desarrollo de superinteligencia, negarla puede ser ni más ni menos que taparse los ojos ante la incertidumbre de los futuros posibles.

Volvamos al tema de la aceleración. Supongamos, propone Bostrom, que tuviésemos la posibilidad de mover la palanca de un acelerador universal hipotético. Podemos inclinarla hacia delante y entonces todo “el ritmo en el que las características macro estructurales de la condición humana se desarrollan” se acelera; la inclinamos hacia atrás y desaceleramos; no la movemos y entonces se mantiene el ritmo actual. ¿Usted qué haría?, ¿aceleraría aún más, frenaría, o seguiría con la velocidad como está ahora? Dejando de lado su voluntad individual, ¿qué piensa que sería mejor para la humanidad en general y para la vida del *sapiens* en la Tierra? (Hanson, 2009).

Max Tegmark: ¿una explosión de inteligencia?

Si Nick Bostrom es el principal filósofo del “Instituto del futuro de la humanidad”, de la Universidad de Oxford en Inglaterra, el físico Max

Tegmark es el director del “Instituto para el futuro de la vida”, y profesor de física y cosmología del MIT, en Estados Unidos. Ese instituto sobre el futuro que preside Tegmark es “una organización de investigación y divulgación dirigida por voluntarios en el área de Boston que trabaja para mitigar los riesgos existenciales que enfrenta la humanidad, particularmente el riesgo existencial de la inteligencia artificial avanzada” (Future of Life Institute, 2015). Tegmark ha publicado recientemente su libro *Vida 3.0* y el cuarto capítulo se titula precisamente “¿Explosión de inteligencia?”. Dedicaremos los siguientes párrafos a analizar el pensamiento de Tegmark sobre la aceleración tecnológica, los límites que podría trascender la inteligencia artificial y a establecer coincidencias, discrepancias y matices con Bostrom.

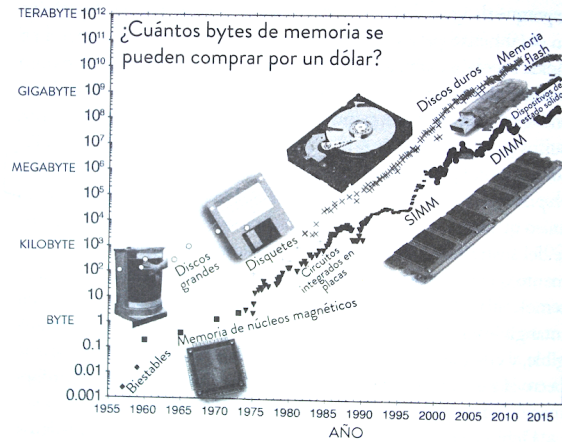
Partamos de una afirmación clara y contundente que Tegmark acepta e incluye como una de las conclusiones de su segundo capítulo:

Una vez que la tecnología es el doble de potente, a menudo puede utilizarse para diseñar y construir tecnología que es a su vez el doble de potente, lo que desencadena una duplicación reiterada de la capacidad, en la línea de la Ley de Moore. El coste de la tecnología de la información se ha reducido a la mitad aproximadamente cada dos años durante un siglo, haciendo posible la era de la información. (2018, p. 105)

Este fenómeno de crecimiento exponencial se verifica para todos los campos de la computación y también para nuevos campos tecnológicos como la secuenciación del genoma, la memoria o la decodificación del cerebro.

Tegmark, además, hace referencia a Kurzweil y lo que este definió como “la ley de rendimientos acelerados” (2018, p. 90) por cuya evolución exponencial se llegan a cifras astronómicas. Y para intentar dar una imagen del alcance de esta exponencialidad, Tegmark invoca la teoría del *Big Bang*, aquella que afirma que nuestro universo se expandió exponencialmente desde un punto, “por el cual cada paso de duplicación provocaba el siguiente” (2018, p. 91). Hemos hablado ya de la ley de Moore y lo haremos nuevamente más adelante. Solo quisiéramos que el lector intente hacerse una composición de lugar intuitiva de lo que representaron cien años de duplicación de la capacidad tecnológica cada dieciocho meses y, más aún, que intente imaginar lo que representará que esto siga así otro siglo. Es verdad, el resultado no es fácil de intuir, pero lo que sí podemos saber es que las innovaciones en esta trayectoria serán tan potentes que las del presente, en comparación, parecerán más obsoletas y primitivas que las cachilas Ford T. Tegmark lo llama una posible explosión de inteligencia que nos llevaría a la superinteligencia artificial. Pero toma recaudos, nadie puede afirmar con total certeza que efectivamente esto ocurrirá y la incertidumbre es la que predominará. A continuación, se presenta una gráfica (figura 1) que muestra la aceleración de la cantidad de computaciones por segundo a lo largo de los últimos cien años. Invitamos al lector que continúe la proyección de esta exponencialidad para los próximos cincuenta años.

Figura 1. Cálculos por segundo por cada mil dólares



Fuente: Tegmark (2018, p. 90)⁶.

¿A qué niveles de la inteligencia artificial hemos llegado hoy?, ¿dónde estamos parados?, ¿cuáles son algunos de los alcances más notables? Un ejemplo concreto del proceso de aceleración de la IA en el que se detiene, como también muchos otros expertos en el tema lo destacan, es el esfuerzo de *Deep Mind* (la sección de Google encargada de desarrollar IA), primero con *AlphaGo* y luego con *AlphaZero*, siendo ambos, programas para el juego del Go. El Go tiene origen chino y se juega en un tablero cuadrado de 19 líneas horizontales y 19 líneas verticales, dando lugar a 361 intersecciones donde cada jugador va poniendo piedras negras y blancas con el objetivo de dominar la mayor cantidad de terreno. El número de posiciones posibles en este juego es tan elevado que los calculistas han estimado que son más que la cantidad de átomos existente en el universo. Por esta razón es un juego que combina una gran dosis de intuición junto con una necesaria estrategia racional. En el año 2015 se creó la primera versión de *AlphaGo* y ya para el año 2016 el campeón mundial de este juego Lee Sedol reconoció su impotencia ante el programa que lo venció todas las veces que lo enfrentó. La estructura de este programa de inteligencia artificial es una base de datos de miles de partidas previas entre humanos y entre *AlphaGo*, y entre este y una réplica de sí mismo; a su vez, le sumaron una red de información que pudiese predecir la jugada más probable en respuesta a otra realizada. Juntaron ambas informaciones y el resultado fue un programa imbatible para cualquier humano. No suficiente con esto, dos años después en 2017, *DeepMind* creó un nuevo programa llamado *AlphaZero*. Pero en este caso la estructura fue completamente diferente ya que le introdujeron solamente

6. Si graficáramos la aceleración de la cantidad de bytes de memoria por dólar durante los últimos sesenta años la exponencialidad sería prácticamente la misma (Tegmark, 2018, p. 80).

las reglas del juego del Go sin ninguna base de datos (a diferencia del programa anterior), y lo programaron para que jugando contra sí mismo fuese aprendiendo y mejorándose cada vez más. El resultado fue sorprendente, en dos horas *AlphaZero* podía vencer al mejor humano y enseguida derrotó al previo *AlphaGo*. Este mismo programa *AlphaZero* aprendió también a jugar al ajedrez y fue capaz en unas pocas horas de perfeccionarse hasta que logró ser insuperable para cualquier programa previo de ajedrez, así como para cualquier humano. El gran descubrimiento de estos ejemplos es que una vez que una IA logra aprender a aprender en su área específica, no hay capacidad humana que pueda sobrepasarla en esa área, y que además se mejora a sí misma recurrentemente. Como ya advertimos, se trata de IA estrecha, es decir, muy superior a los humanos en ese campo, pero incapaz de resolver otro tipo de desafíos en otros escenarios. Al analizar estos casos concretos Tegmark escribe: “Tanto Nick Bostrom como Ray Kurzweil han insistido en lo difícil que es prever estos avances en IA” (2018, p. 112) ya que la mayoría de los expertos antes de que estos programas se materializaran descartaban que pudiesen inventarse en un futuro cercano. Pero la realidad demostró que fue en apenas dos o tres años. En parte, esta es la confirmación de la hipótesis de Irving Good de 1965, pero en dos campos específicos como son el Go y el ajedrez. Una vez que un sistema es capaz de mejorarse recurrentemente a sí mismo, se convierte en el último invento humano. ¿Será *AlphaZero* el último invento en estos dos juegos, será tan dominante en el futuro, tendrá la capacidad de mejorarse a sí mismo indefinidamente?

Un segundo ejemplo tiene que ver con el lenguaje. Precisamente en este campo pocos meses atrás nosotros mismos nos sorprendimos. Hasta hace aproximadamente cinco años quienes escribimos en español debíamos contratar un traductor para que nuestros artículos científicos pudieran ser presentados a publicaciones en inglés. Esto tenía un costo considerable de varios cientos de dólares, demoraba semanas, y el traductor debía tener tiempo disponible. Muchas veces el idioma era una barrera que limitaba el acceder a una audiencia mayor en inglés, así como a las mejores revistas que también se publican en este idioma. Poco tiempo atrás, sin embargo, probamos el traductor de Google, elaborado por *GoogleBrain*. La sorpresa fue extrema porque el resultado fue casi perfecto, con mínimos errores. En unos segundos la IA de *GoogleBrain* tradujo el artículo del español al inglés. Otra profesión que quedará obsoleta, los traductores humanos. Los últimos avances de traducción de este sistema se han basado en la utilización de redes neuronales profundas y el salto cualitativo fue espectacular. Hay otros traductores gratuitos online, por ejemplo *DeepL*, una compañía alemana que, en su sitio web, sostiene que su objetivo es “entrenar inteligencia artificial para que entienda y traduzca textos” y que “las redes neuronales expanden las posibilidades humanas, superan las barreras del lenguaje y acercan culturas”⁷. En el año 2017 *DeepL* lanzó en internet el traductor

7. Véase <https://www.deepl.com/home>

gratuito para cualquier idioma con una excelente calidad. Si tenemos en cuenta el test de Turing, no estamos lejos de que una máquina pueda expresarse por escrito con un nivel suficiente como para engañar a su interlocutor de que se está comunicando con otro humano. Dominar el lenguaje será un paso esencial en el avance de la IA hacia un nivel humano. Tegmark (2018) resalta el salto en el campo de la traducción, no solo de textos escritos, sino de discursos hablados, pero sostiene que aún los ordenadores no son capaces de entender lo que traducen. Si bien un ordenador puede utilizar casi perfectamente las palabras, no capta sus referencias en el mundo, no sabe realmente lo que significa una mujer, o un pez, o un escritorio. Ya llegará ese momento, vaticina el autor. Los sistemas artificiales de traducción funcionan casi a la perfección con textos científicos o de información más objetiva porque pueden ser traducidos literalmente. Otros escritos más subjetivos, como poesías, novelas o cuentos, es decir obras con componentes estéticos determinantes, aún no se traducen bien con estos sistemas artificiales. La razón es que un buen traductor debe interpretar el texto y muchas veces expresarlo en el otro idioma mediante asociaciones muy alejadas de la literalidad. Algo de esto ocurrió recientemente cuando los sistemas de Twitter y Facebook intentaron erradicar de sus redes el lenguaje que alentaba el odio contra comunidades o personas. En estos casos los sistemas de IA fueron incapaces de tomar en cuenta los contextos culturales y entender cómo, dependiendo del contexto en que se emite, una expresión similar puede tener significados radicalmente diferentes⁸. Es muy probable que también se lleguen a solucionar estas dificultades en un lapso breve y la propia estructura de estos programas que pueden aprender a aprender lo anticipan.

Podríamos continuar con infinidad de ejemplos de los últimos avances en IA en todas las áreas de la vida social que nos dejan sin aliento: en finanzas, en armamentos, en arte, en medicina, en información, en la investigación del cosmos, en nanotecnologías. También Tegmark comienza por interrogarse ¿cuándo se daría el primer gran paso, es decir cuándo aparecerá una inteligencia artificial general de nivel humano, aquella que podría realizar cualquier tarea cognitiva igual que el humano mejor dotado para desarrollarla? En este sentido propone que hay al menos cuatro grandes grupos de posiciones en torno a la pregunta. A un primer grupo lo llama el de los tecnoescépticos; son aquellos que no admiten que esto pueda suceder ni en veinte, cincuenta o cien años, y para quienes los sapiens estaremos en control todavía mucho tiempo más. (Tegmark personaliza esta posición en dos científicos renombrados como Andrew Ng y Rodney Brooks). A un segundo grupo lo denomina el de los utópicos digitales, aquellos que creen que el avance de la innovación terminará generando un tipo de vida artificial mucho más inteligente que la vida biológica y ese salto mejorará nuestra existencia, será positivo y no deberíamos

8. Véase, en particular, sobre este tema “Changing Relationship between Artificial Intelligence and Humans” (Unesco, 2019^a; 2019b).

asustarnos ni enlentecer la innovación. (En este grupo incluye personas renombradas como el fundador de *Google* Larry Page, a Hans Moravec y a Ray Kurzweil). El tercer grupo es el de los que integran el movimiento en pro de una IA benéfica (el propio Tegmark se incluye en este grupo junto a Bostrom, Elon Musk, Vernor Venge, Bill Gates, Stephen Hawking, Stuart Russel). La proposición de estos últimos es que hay una posibilidad real de que se produzca el salto hacia una IA general de nivel humano en este siglo y que a su vez tendría probabilidades de seguir avanzando hacia una IA superior y que una de las tareas actuales es intentar profundizar en aprender más cómo controlar que ese nuevo sistema opere en el sentido de defender ciertos valores del humanismo. Este objetivo sería justamente el del Instituto del futuro de la humanidad. ¿Cuándo se podría materializar esta posibilidad?, la respuesta, en este caso de Tegmark, es que no hay consenso y sí incertidumbre, pero una estimación que proporciona es que, para los expertos reunidos en enero de 2015 en una conferencia en Puerto Rico, la opinión predominante fue que para la mitad del siglo que viene sería esperable. Un cuarto y último grupo es el de los *neoluditas* que, así como el movimiento en los comienzos de la primera revolución industrial destruían los telares y las cosechadoras, tiene una percepción muy crítica sobre la aceleración de la tecnología y promueven su moratoria indefinida ya que estiman que los resultados serán negativos. A los representantes de esta posición de detener el avance tecnológico los identifica dentro del movimiento antiglobalización y en parte del movimiento ecológico y menciona específicamente como uno de los más notorios al ecologista estadounidense Bill McKibben (2003)⁹. El lector podrá ubicar fácilmente a estos cuatro grupos en relación a lo que harían con la hipotética palanca de aceleración tecnológica descrita en la sección anterior, quiénes la moverían para acelerar, quiénes frenarían y quienes seguirían más o menos como ahora, pero tomando mucho mayor precaución¹⁰.

RayKurzweil: la singularidad

Es muy difícil, incluso para las personas inmersas en los epicentros del desarrollo tecnológico, mostrar cuándo se dará el salto sin precedentes en el mundo de la tecnología, por ejemplo, que la aceleración de la IA llegue y supere el nivel de inteligencia humana. Raymond Kurzweil en el año 2005 titulaba a una de sus obras *La singularidad está cerca*, provocando un poco de espanto en el mundo tecnológico, y ni que hablar en el social. Era un concepto complejo, difícil de digerir y aún hoy lo sigue siendo. Sus preocupaciones respecto a la IA son centrales dado el hecho de que ve factible que llegará un momento en que el humano superará sus limitaciones

9. McKibben, Bill (2003). *Enough: Staying Human in an Engineered Age*, Times Books, New York.
10. Muy cerca de la posición de McKibben y linderas con los neo-luditas estarían casi todas las variantes de las teorías del Buen vivir concebidas en América Latina y con fuertes raíces en los movimientos indígenas (Hidalgo-Capitán y Cubillo-Guevara, 2017).

biológicas y se irá paulatinamente fusionando con las máquinas. A este punto de trascendencia biológica lo denomina *singularidad*¹¹ (Chalmers, 2010).

Kurzweil observa que la vida en la tierra ha evolucionado de forma bastante perceptible en cuatro eras de las seis que él prevé, cada una utilizando los resultados de la anterior para crear la siguiente; la singularidad llegará en la quinta era. La primera era es la de la física y química con el comienzo del universo, la segunda es la era biológica y del ADN, compuestos basados en carbono que se volvieron más complejos; la tercera era es la de los cerebros y los sistemas nerviosos, momento cuando los animales comienzan a reconocer patrones; la cuarta era es la de la tecnología, cuando se combina el pensamiento racional y abstracto con el pulgar oponible y se comienza a crear tecnología como herramientas básicas hasta máquinas mecánicas automatizadas (Kurzweil, 2012). Llegado a determinado punto de desarrollo tecnológico, arribaremos a lo que denomina la quinta era:

será el resultado de la fusión entre el enorme conocimiento alojado en nuestros cerebros y la enormemente superior capacidad, velocidad y agilidad para compartir información de nuestra tecnología. Así, la quinta era permitirá que nuestra civilización humano-máquina trascienda las limitaciones de las tan solo cien billones de conexiones extremadamente lentas del cerebro humano. (Kurzweil, 2012, pp. 22-23)

Esta era acoge a los conceptos de superinteligencia de Bostrom, y de explosión de inteligencia de Tegmark, puesto que es aquí donde se daría ese momento en que la IA iguale o supere nuestra inteligencia en todas sus dimensiones. Hay más. Si colocamos la lupa en la parte de la gráfica exponencial de la cuarta era de la evolución humana observaremos que la tecnología tiene un cambio exponencial con forma de “S”. El advenimiento del siguiente paradigma tecnológico ocurrirá cuando el anterior muestre una meseta, lo que le da cierta forma de letra “S” corrida “_ | -” tan característica que hemos mencionado anteriormente. Esto, por supuesto, deja entrever un panorama que, si bien puede sonar de ciencia ficción, hoy está siendo seriamente tratado y, de hecho, hay un cierto consenso respecto a que nos permitirá ir más allá de las limitaciones biológicas, lo que provocará fabulosos cambios en las sociedades desde todo punto de vista pero en particular sobre la mortalidad y la composición de las sociedades, como él mismo indica:

el control sobre nuestros destinos, nuestra mortalidad estará en nuestras propias manos, podremos vivir tanto como queramos (que es un poco diferente a decir que viviremos para siempre), comprenderemos enteramente el pensamiento humano y

11. Recordemos que hay quienes atribuyen el concepto de singularidad a Vernor Vinge (1993), bastante anterior a la obra de Kurzweil en cuestión.

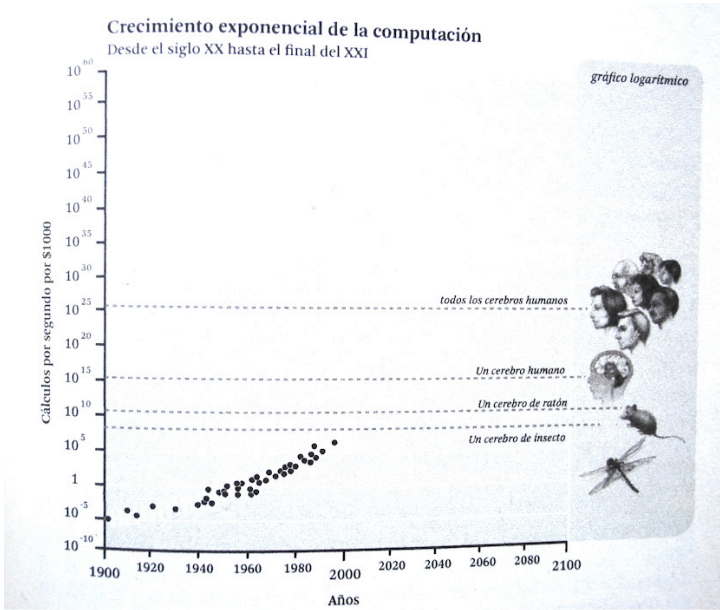
expandiremos y aumentaremos enormemente su alcance. Como consecuencia, al final de este siglo la parte no biológica de nuestra inteligencia será billones de billones de veces más poderosa que la débil inteligencia humana producto de la biología. (Kurzweil, 2012, p. 9)

En otras palabras, luego de alcanzada la singularidad, estaremos más cerca de volvernos organismos (si aun cupiese la palabra) verdaderamente universales, es decir, de poder explorar el universo mediante nuestros propios recursos intelectuales, lo cual dará inicio a la sexta etapa de evolución.

¿Cómo llega Kurzweil a semejante proyección?, ¿es apenas un charlatán, tiene fundamentos científicos serios e interesantes para su teoría? Sus argumentos no difieren mucho del conjunto de razonamientos de los dos autores tratados antes. La idea detrás de la singularidad es, en el fondo, la misma: que el ritmo del cambio tecnológico se está dando de manera exponencial. También él parte de lo que Gordon Moore propuso por la década de los sesenta, y desarrolla una teoría en la cual, según sus cálculos, la humanidad irá cada vez acercándose a la fusión con sus propias creaciones tecnológicas, y que este proceso, es relativamente inminente. Repitamos una vez más lo mostrado en la introducción de este artículo, la Ley de Moore que Kurzweil cita: “cada veinticuatro meses era posible doblar el número de transistores que se podían encajar en un circuito integrado” (2012, p. 64). Lo que mostraba era, pues, lo siguiente: mientras más pequeño se transforma el recipiente físico, más rápido circula el electrón produciendo mayor velocidad de circulación provocando así que se pueda volver a innovar más rápido, porque habrá un espacio físico más reducido en el que el electrón tiene menos que recorrer. Corrigiendo y ampliando la ley, Kurzweil afirma:

Tal y como hemos visto [...], la Ley de Moore se refiere específicamente al número de transistores en un circuito integrado de tamaño fijo, y a veces esta ley ha sido todavía más específica al ser expresada en términos de tamaño del dispositivo de los transistores. Pero la medida más apropiada para reflejar el rendimiento por unidad de coste es la velocidad computacional por unidad de coste. (2012, p. 78)

De aquí Kurzweil entiende necesaria la readaptación de la ley porque esta se estanca en determinados espacios físicos. Moore no previó la posibilidad de que en el futuro existiría tecnología que excediese la materialidad de los circuitos integrados, la cual será la que permita la cantidad de computación y memoria suficiente para conseguir inteligencia de nivel humano: por ejemplo, una computación molecular en 3-D, computación con moléculas, o la computación cuántica a la que se refería Hawking.

Figura 2. Cálculos por segundo por cada mil dólares

Fuente: Kurzweil (2012, p. 77).

Ahora pensemos con la lógica exponencial, como ilustra la gráfica (figura 2) de Kurzweil: ¿qué tiene que ver que podamos comunicarnos mejor con que se desarrollen las técnicas informáticas? La premisa que se sostiene es que, a mayor capacidad de transmitir información, más rápido se dará el siguiente cambio computacional debido a la fluidez con la que la información llega, a la velocidad de los intercambios y la subsiguiente mayor interconectividad global. La proyección de Kurzweil que vemos en la gráfica, indica que las computadoras llegarán a tener un nivel de conocimiento similar al humano entre 2020 y 2040. Poco tiempo después, en el tercer quinto del siglo XXI, acontecerá la singularidad, o la superinteligencia o la explosión de inteligencia.

Se podrá objetar, y con mucho criterio, que el crecimiento exponencial de la tecnología, la computación y la IA adolecerán de enormes retrocesos (o al menos serios estancamientos) según el devenir de la humanidad: por las guerras, pandemias, crisis económicas, etc. Incluso, como se mencionó, algunos como Bridle hablan de una némesis de la Ley de Moore. Se podrá también indicar que dicha gráfica solo contempla las sociedades occidentales, es decir aquellas que están dentro de la órbita de Estados Unidos y con fuertes antecedentes culturales con raíces europeas. Esto no parece impedir que el crecimiento siga siendo exponencial.

El estallido de la primera guerra mundial en 1914, la gripe española de 1918, la crisis económica en 1929, el comienzo de la segunda guerra mundial en 1939, la polarización global a partir de 1948, las guerras de

Corea (1950) o Vietnam (1955), los procesos decoloniales, la crisis del petróleo de 1973 y un largo etcétera, no pudieron contener el desarrollo de las fuerzas impulsoras de la tecnología. Tampoco parece sostenerse la teoría de que es un fenómeno solamente de la cultura occidental. El caso de China luce interesante para mostrar como al impulsar mayor fluidez informativa se pueden generar enormes avances en los campos informáticos. El caso de la ex URSS también podría demostrarlo, al pasar de tener una población predominantemente agrícola a enviar cohetes al espacio en cincuenta años e iniciar una competencia con Estados Unidos sobre el dominio espacial, mediante la importación, adaptación y posteriormente desarrollo de determinadas tecnologías de información y de comunicación vigentes por aquellos años.

El actual contexto de pandemia global por el COVID-19, distinto al caso de la gripe española de hace cien años, podría darnos alguna pista. Un apagón tecnológico, es decir, una caída global como consecuencia de alguna crisis sanitaria podría arrasarlo con todo lo vinculado al desarrollo tecnológico, podría detener esos avances dado que para que haya un siguiente salto debe existir una plataforma de base para que se dé; debe haber algo previo para poder generar un conocimiento nuevo en tecnología. De hecho, este punto de vista de que se genere una enorme interconectividad y que la IA llegue a dominar el mundo de manera supraestatal es una de las críticas que aborda Eric Sadin en *La silicolonización del mundo: la irresistible expansión del liberalismo digital* (2018). Igualmente, podemos señalar que esta crítica adolece de un problema y es que, y dado que en su fase actual de evolución la intercomunicación es tan sólida, los servidores globales de almacenamiento están tan dispersos y la tecnología está tan legitimada por los humanos, sin ninguna distinción social, étnico-racial, económica, etcétera, que resultaría muy difícil que se produzca un apagón tecnológico, por ejemplo, con alguna clase de ciber ataque masivo, algo también lejos de suceder por las enormes medidas de seguridad que se han puesto en práctica. Parece que la frase que hemos citado de Quino acerca de la detención del mundo, al menos en el pensamiento de Kurzweil, es algo que no pasará en el plano de la tecnología y su aceleración constante.

La singularidad, según Kurzweil, cambiará el mundo en el que vivimos de forma rotunda:

Es un tiempo venidero en el que el ritmo del cambio tecnológico será tan rápido y su repercusión tan profunda que la vida humana se verá transformada de forma irreversible. Aunque ni utópica ni distópica, esta era transformará los conceptos de los que dependemos a la hora de dar significado a nuestras vidas, ya sea en lo que se refiere a modelos de negocios o al ciclo de la vida (incluyendo la muerte) [...] La idea fundamental que subyace de la inminente singularidad es que el ritmo de cambio de la tecnología creada por el hombre se está dando y que sus capacidades se están ampliando a ritmo exponencial. (2012, pp. 7-8)

En la postsingularidad, no habrá distinción entre humano y máquina o entre realidad física y realidad virtual¹². Si se pregunta sobre lo que seguirá siendo inequívocamente humano en un mundo así, la respuesta es simplemente esta cualidad: la nuestra es la especie que inherentemente busca expandir su alcance físico y mental más allá de sus limitaciones actuales y que logrará trascender la evolución biológica de las especies con la evolución tecnológica y la inteligencia artificial.

Conclusiones

El objetivo central de este trabajo ha sido entender el grado de aceleración de la innovación tecnológica y cómo esta se proyecta en el futuro de la inteligencia artificial. Para lograrlo nos hemos basado en los aportes de tres teóricos de la tecnología: Nick Bostrom, Max Tegmark y Ray Kurzweil. Bostrom, mediante su término *superinteligencia* intenta mostrar que, aunque primero de forma tímida y luego más frecuentemente, la inteligencia artificial se acerca a un punto en el cual superará la inteligencia de nivel humano.

Los tres autores utilizan como base para sus conjeturas a la Ley de Moore y el crecimiento exponencial del poder de la tecnología. También los tres nos advierten sobre la necesidad de controles en los avances de la inteligencia artificial. Incluso para quien es considerado el más optimista de los tres, como sería el caso de Kurzweil, se acepta la necesidad de “estar activamente preocupados sobre su lado oscuro (y) El reto más importante es resolver qué particularidades de la renuncia son a la vez factibles y deseables” (Kurzweil, 2012, p. 472). Los límites éticos, el hasta dónde, a qué renunciar, son cuestiones que no pueden quedar de lado si apuntamos a que los desarrollos tecnológicos terminen beneficiando a la humanidad. Los cambios en genética, nanotecnología y robótica ponen sobre la mesa una cuestión muy evidente y es cómo hacer para que no sean utilizadas para fines destructivos. La tecnología cada vez logra ser más accesible para más personas, solo es necesario tener el conocimiento para saber qué hacer con ella. Los hackers de hoy en día podrían poner en jaque a los gobiernos, apropiándose de la inteligencia artificial, de nano robots o robots y utilizarlos para fines opuestos al beneficio de los humanos. Los peligros son tan evidentes que Steven Hawking y otros 149 científicos firmaron una carta abierta en Puerto Rico en la que se puede leer: “Recomendamos una mayor investigación dirigida a garantizar que los sistemas de IA cada vez más capaces sean robustos y beneficiosos: nuestros sistemas de IA deben hacer lo que queremos que hagan” (Future of Life Institute, 2015).

La regulación de la tecnología en genética, nanotecnología y robótica y su interfaz con la IA es, según varios expertos, uno de los principales desafíos de la próxima década; así lo han considerado también diversos organismos multilaterales como la Unesco (2018; 2019^a y 2019^b) y la Unión Europea y en consecuencia ya se han organizado varios congresos internacionales para discutir el tema. Nuestro punto de vista, con base en lo analizado en

12. Véanse, en este sentido, Arocena (2017) y Bostrom (2001).

este trabajo, es que esto debería hacerse no desde la perspectiva extrema de pretender detener la innovación tecnológica; tampoco compartimos la posición opuesta de que la aceleración tecnológica es de por sí positiva y por eso no hay que hacer nada. Nuestra convicción podría ser definida como la de un apoyo escéptico a la regulación.

Este apoyo escéptico se basa en que la aceleración tecnológica y la innovación en IA es de tal magnitud que en determinado momento se crearán sistemas, organismos, o máquinas que tendrán agencia propia y serán más inteligentes que sus creadores. Entonces será extremadamente difícil que hagan lo que queramos que hagan. La preocupación por la aceleración de la IA ya se plasmó en protocolos éticos, pero, como afirma Hawking “algunas personas no podrán resistir la tentación” de experimentar en este y en otros de los campos mencionados como la ingeniería genética y la nanotecnología, y entonces se perforarán esos límites éticos, políticos, filosóficos y científicos y habrá “una carrera de seres auto diseñados, que se irán mejorando a un ritmo cada vez mayor” (Hawking, 2018 p. 115). Este tipo de tentación de no renunciar a hacer lo que se había acordado que no se haría, ya ocurrió el año pasado con los gemelos chinos a los que se les modificó su ADN artificialmente para inmunizarlos contra el VIH, a pesar del compromiso ético de la comunidad científica de no innovar en humanos con la edición genética. Es que como hemos intentado mostrar con la discusión de este trabajo, es extremadamente difícil ejercer el control sobre la aceleración tecnológica en general y en particular en la inteligencia artificial. A pesar de ello, hay que continuar y profundizar los esfuerzos de regulación (Gómez-Sellés, 2019).

Las proyecciones de la aceleración tecnológica en la inteligencia artificial pueden parecerles a algunos lectores aún materia de ciencia ficción más que de ciencias sociales. Es verdad que las tendencias analizadas en este trabajo son eso, apenas proyecciones informadas por especialistas. Podrán ocurrir o no, y si acontece algo parecido, ello podrá suceder en los plazos estimados, demorar más, o llegar anticipadamente. Ya sabemos que lo que parecía ciencia ficción hace una o dos décadas hoy es realidad y que en lapsos cada vez más cortos emergieron innovaciones que antes no habíamos todavía imaginado. Las ciencias sociales no pueden quedarse atrás y es bueno recordar que parte de sus tareas más significativas es anticipar futuros.

Entre los principales problemas de la aceleración tecnológica y la evolución de la inteligencia artificial hay que destacar que son muy pocas empresas las que dominan el campo. Hay muy pocas personas, con demasiado dinero, concentradas en pocos lugares geográficos, que determinan el rumbo del futuro en estas dimensiones. Esto despierta inquietudes de todo tipo, desde el aumento en la brecha de desigualdad económica y social entre personas, entre países, y entre regiones, hasta dilemas éticos acerca de cuáles valores deberían ser los que orienten las innovaciones. ¿Quiénes deciden y cómo acerca de los límites éticos de sus creaciones? Hay esfuerzos significativos por mejorar en estos dilemas a nivel de los países y a nivel de instituciones supranacionales. Por ejemplo, en materia de ingeniería

genética ya se han adoptado muchas precauciones entre las que sobresalen la obligatoriedad de destruir los embriones utilizados para investigación luego de dos semanas de gestados; no hacer modificaciones genéticas en humanos que se transmitan a las futuras generaciones; no utilizarla para diseñar humanos con capacidades sobresalientes, sino exclusivamente para tratamientos somáticos. En inteligencia artificial la preocupación no ha sido menor, pero parece que los consensos son aún mucho más débiles y difusos. Entre ellos podemos destacar que nunca una decisión sobre la vida y la muerte de un humano debería quedar en manos de un sistema de IA; impedir que los sistemas artificiales puedan auto replicarse en el medio exterior, por el riesgo que ello conlleva; aplicar el principio de precaución, que inhibe el desarrollo de cualquier innovación hasta tanto no saber con cierta certeza que las consecuencias no serán peores que sus beneficios. Obviamente habrá discrepancias en este último aspecto que tal vez no lograrán ser saldadas en el mediano plazo.

Finalmente, la conclusión más importante de este recorrido que hemos hecho sobre la proyección de la aceleración tecnológica en la IA es que las probabilidades de que algún sistema de inteligencia artificial (sea digital, biológico, o híbrido; sea un organismo biológico, una máquina o su interfaz) evolucione y adquiera más capacidades que los propios humanos en todas las dimensiones que asociamos con el concepto de inteligencia, en un tiempo cercano, no son cero. Si esto es correcto, para las ciencias sociales en general y en particular para la sociología los desafíos son múltiples. Entre los principales cometidos de esta disciplina, como sencilla y claramente lo expresó Wright Mills, uno de sus exponentes más reconocidos, es que nuestra imaginación sociológica, informada por nuestras trayectorias intelectuales, contribuya a pensar otros mundos posibles. Ya imaginar los futuros posibles con el penetrante avance de las tecnologías implica todo un campo disciplinar. ¿Que peso y viabilidad tendrán otros proyectos sociales alternativos a la imaginación tecnológica, por ejemplo, aquellas propuestas vinculadas a la expansión de la mente con sustancias psicotrópicas, o asociadas con una vida más en armonía con la naturaleza y la comunidad? ¿Podrán ser complementarios o son antagónicos? ¿Cómo se transformarán las sociedades con el impacto de las tecnologías en casi todas las esferas, por ejemplo, cuando el mundo del trabajo nos depare nuevas profesiones aún no creadas junto a la extinción acelerada de las actuales? ¿Se agudizan inevitablemente las desigualdades socioeconómicas ante el futuro tecnológico ya que las innovaciones más disruptivas parecen ser resultado de grandes inversiones de capital en universidades o compañías con más presupuesto que muchos países latinoamericanos? ¿Cómo revemos teóricamente uno de nuestros principales axiomas, verdadero cimiento de la sociología, de que somos los seres humanos los que construimos nuestra realidad social? ¿Seguiremos en el comando aún si construimos sistemas que nos superen en inteligencia y nos trasciendan biológicamente? ¿Qué nuevos mitos inventaremos, qué tipo de solidaridad nos mantendrá integrados en sociedad? ¿Cómo integremos la realidad virtual cuando

los hologramas estén lado a lado con nuestros cuerpos? ¿Cómo serán los nuevos tipos de dominación? Hay tantas otras preguntas para incorporar, apenas mencionamos ahora algunas de las que balizarán la conversación sociológica de las próximas décadas.

Referencias

- Arocena, F. (2017). *¿Vives en una simulación computacional? La diaria*. Recuperado de <https://ladiaria.com.uy/articulo/2017/5/vives-en-una-simulacion-computacional/>
- Arocena, F. y Sansone, S. (2020). Hámsters en la rueda. Aceleración y cuarta revolución industrial, *Civitas, Revista de Ciências Sociais*, 20(2), 221-233. DOI: <https://doi.org/10.15448/1084-7289.2020.2.33886>
- Baum, S., Goertzel, B. and Goertzel, T. (2011). How Long Until Human-Level AI. Results from an Expert Assessment. *Technological Forecasting and Social Change*, 78(1), 185-195. DOI: <https://doi.org/10.1016/j.techfore.2010.09.006>
- Bostrom, N. (2018). *Superinteligencia*. España: TEELL Editorial.
- Bostrom, N. (2001). ¿Are you living in a computer simulation? *Philosophical Quarterly* 53(11), 243-55. DOI: [HTTPS://DOI.ORG/10.1111/1467-9213.00309](https://doi.org/10.1111/1467-9213.00309)
- Bridle, J. (2018). *New Dark Age: Technology and the End of the Future*. Londres: Verso.
- Chalmers, D. (2010). The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*, 17(9-10), 7-65. DOI: <https://doi.org/10.1002/9781118922590.ch16>
- Future of Life Institute (2015). Prioridades de investigación para una inteligencia artificial robusta y beneficiosa. *Future of Life Institute*. Recuperado de <https://futureoflife.org/2015/10/27/ai-open-letter/>
- Future of Life Institute (2019). *Max Tegmark with Yuval Noah Harari*. Recuperado de <https://futureoflife.org/2019/12/31/on-consciousness-morality-effective-altruism-myth-with-yuval-noah-harari-max-tegmark/>
- Gómez-Sellés, J. (2019). Nacimiento de los dos primeros bebés modificados genéticamente. Análisis del tratamiento de la noticia en España desde el punto de vista de la comunicación de la ciencia. *ArtefaCTos*, 8(2). DOI: <https://doi.org/10.14201/art2019825371>
- Hanson, R. (2009). *Tiptoe or dash to the futures* [blog]. Recuperado de <http://www.overcomingbias.com/2009/12/ptiptoe-or-dash-to-future.html>
- Hawking, S. (2019). *Breves respuestas a las grandes preguntas*. Buenos Aires: Crítica.
- Hidalgo-Capitán, A. L. y Cubillo-Guevara, A. P. (2017). Deconstrucción y genealogía del “buen vivir” latinoamericano. El (trino) “buen vivir” y susdiversos manantiales intelectuales”. *Revue internationale de politique de Développement*, 9. DOI: <https://doi.org/10.4000/poldev.2517>
- Joy, B. (2000). Why the future doesn't need us. *Wired Magazine*. Recuperado de <https://www.wired.com/2000/04/joy-2/>
- Kelly, K. (2017). *Lo inevitable: entender las 12 fuerzas tecnológicas que configurarán nuestro futuro*. Madrid: Teell Editorial.
- Kurzweil, R. (2012). *La singularidad está cerca*. Berlín: Lola Books.

- McEwan, I. (2019). *Machines Like me*. New York: Decker Edge.
- McKibben, B. (2003). *Enough: Staying Human in an Engineered Age*. New York: Times Books.
- Moravec, Hans P. (1998). When Will Computer Hardware Match the Human Brain? *Journal of evolution And Technology*, 1. Recuperado de <https://jetpress.org/volume1/moravec.htm#:~:text=As%20production%20techniques%20for%20those,in%20home%20computers%20before%202030>.
- Nilsson, N. (2009). *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. New York: Cambridge University Press.
- Schwab, K. (2016). *La cuarta revolución industrial*. Barcelona: Debate.
- Sadin, E. (2018). *La silicolonización del mundo: la irresistible expansión del liberalismo digital*. Buenos Aires: Caja Negra editores.
- Tegmark, M. (2018). *Vida 3.0. ¿Qué significa ser humano en la era de la inteligencia artificial?* Barcelona: Taurus.
- Unesco (2018). *Roundtable on the Ethical Implications of Big Data and Artificial Intelligence*. Recuperado de https://web.facebook.com/watch/live/?ref=watch_permalink&v=510453616140032
- Unesco (2019a). *Second Roundtable on the Ethics of Genome Editing, "Impact of Genome Editing on our Health and Environment"*. Recuperado de http://webcast.unesco.org/live/vod/2019/shs/20190212_shs_room-02/en/
- Unesco (2019b). *Second Roundtable on the Ethics of Artificial Intelligence, "Changing Relationship between Artificial Intelligence and Humans"*. Recuperado de http://webcast.unesco.org/live/vod/2019/shs/20190312_shs_room-02/en/
- Vinge, V. (1993). The Coming Technological Singularity: How To Survive in the Post-Human Era. *Vision-21: Interdisciplinary science and Engineering in the Era of Cyberspace*, 11-22.