



**SINGULARIDAD TECNOLÓGICA Y TRANSHUMANISMO:
CUANDO LA CIENCIA SE ENCUENTRE CON LA FICCIÓN**

**Por Federico G. Witt
Córdoba, España**

El pasado 20 de noviembre tuve la oportunidad de impartir una conferencia en el Círculo de Bellas Artes de Madrid. La charla formaba parte de la programación de un acto dentro de la VIII Semana de la Ciencia: la presentación de la Unidad de Cultura Científica de la Universidad Autónoma de Madrid, a cargo de **José Antonio López Guerrero**, Director de Cultura Científica de la UAM, y **Margarita Arroyo Macarro**, Coordinadora de Cultura Científica de la misma universidad.

Escogí hablar de un concepto que suscita controversia y siempre da lugar a entretenidos coloquios, la ***Singularidad tecnológica*** (en adelante la Singularidad). Este término hace referencia a un momento crucial de la historia, que se presupone tendrá lugar en un futuro a medio plazo, en el cual se producirá una aceleración súbita del desarrollo tecnológico. En teoría será consecuencia del carácter exponencial del propio progreso, que conducirá a la creación eventual de una inteligencia artificial superior a la humana. Esto a su vez producirá un cambio tecnológico y social tan drástico que cualquier ser humano actual, anterior a la Singularidad, es ahora incapaz de comprender, predecir o incluso imaginar cómo serán la ciencia, la tecnología y la vida postsingulares.



El nombre viene dado por analogía con las singularidades descritas en física teórica, en las que las reglas universales conocidas dejan de ser válidas y la convergencia hacia valores infinitos imposibilita definir una función.

Se piensa que el primero que mencionó el término “singularidad” fue **John von Neumann**, americano de origen austrohúngaro conocido sobre todo como padre de la Teoría de Juegos y autor de la Arquitectura de von Neumann para ordenadores, además de contar en su haber con muchos otros logros, algunos de los cuales sin embargo pasarán a la historia como capítulos poco decorosos de la misma. El caso es que fue, como íbamos diciendo, esta mente brillante y polifacética la que ya en la década de 1950 —von Neumann falleció en 1957— usara el término que nos ocupa para aludir a un momento histórico a partir del cual, como consecuencia del progreso tecnológico, “los asuntos de índole humana no podrían continuar



desarrollándose tal y como los conocemos". Esta es la transcripción literal, por Stanislaw Ulam, de la conversación que sostuvo con von Neumann:

*"One conversation centered on the ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential **singularity** in the history of the race beyond which human affairs, as we know them, could not continue."*

Von Neumann se limitó al hecho del desarrollo tecnológico y a sus consecuencias a partir de determinado momento (esa singularidad esencial histórica), sin mencionar nada acerca de alguna supuesta inteligencia artificial. Quien sí lo hizo, durante la década de 1960, fue un conocido estadístico, **Irving John (Jack) Good**. Este discípulo de Turing predijo una "explosión de la inteligencia" cuando, casi de pasada, sentaba las bases del concepto que tratamos aquí. Good se ciñó a las propias máquinas, argumentando que si se creara una de ellas que estuviera dotada de una inteligencia superior a la humana, una de sus actividades consistiría en producir a su vez máquinas más inteligentes que ella, las cuales harían lo correspondiente de forma sucesiva. Todo esto originaría una "explosión de inteligencia" que dejaría muy atrás al ser humano. Y añadió que, como consecuencia, el invento de una máquina ultrainteligente sería el último que necesitase realizar la raza humana.

"Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an 'intelligence explosion', and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man need ever make."

Pero es a partir de los 80 cuando **Vernor Vinge** popularizó la hipótesis de la que estamos hablando hoy, y a lo largo de los años la ha ido modificando hasta desarrollar el concepto tal y como se entiende en la actualidad. Él mismo lo resumió así en un ensayo de 1993: *"En unos treinta años dispondremos de los medios tecnológicos para crear inteligencia artificial superhumana. Poco después, la era humana acabará"*.

No cabe la menor duda de que Vinge, como buen escritor que es, sabe enganchar al lector con un comienzo impactante. Pero no se trata sólo de un comienzo estremecedor sino que esas dos frases resumen muy bien lo que el autor pretende comunicar en su controvertido ensayo.



Vinge también decía en el artículo que (*sic*) “Nos encontramos al borde de un cambio comparable a la aparición del ser humano sobre la Tierra”. Y añadió que había varias maneras de llegar a la Singularidad:

- El desarrollo de una IA que “despierte”, alcance a la humana y posteriormente la supere (posteriormente, ya en 2008, ha denominado a esta vía Escenario AI, por las siglas de Inteligencia Artificial en inglés). Este escenario seguiría la línea argumental de Jack Good.
- El desarrollo de redes de ordenadores que “despierten” como entes inteligentes (vía que ahora ha dividido en dos: el Escenario de Internet, que no necesita explicación para todo aquel que haya seguido la serie *Terminator*, y el Escenario de Gaia Digital —los propios miles de millones de microprocesadores, interconectados en redes, en el futuro podrían cobrar una conciencia propia y global).
- El desarrollo de interfaces que permitan a un humano comportarse como un ser superinteligente (Escenario IA —Intelligence Amplification—, que cualquier lector de *hard sci-fi* tendrá presente; véanse, por citar un solo ejemplo, varios casos en *Axiomático*, de Greg Egan [AJEC, 2006])
- Manipulaciones biológicas que permitan mejorar el nivel de inteligencia humana (Escenario Biomédico, reflejado por ejemplo en *GATTACA*).

Vinge ha expresado que lo más probable es que ocurra una combinación de varios de estos escenarios.

Las tres primeras posibilidades —o cuatro primeros escenarios, según se mire— dependen del desarrollo del hardware; y en la época en que escribí esto, las tendencias apuntaban a que entre 2005 y 2030 se podía lograr superar a la inteligencia humana por medio de dicho hardware. Hoy en día parece que no somos tan optimistas, no por el hecho en sí de que se consiga o no superar la capacidad de cálculo o computación del cerebro humano, asunto sobre el que hay pocas dudas, sino porque la Inteligencia Artificial, la conciencia autónoma, la capacidad de tener iniciativa y no tanto de solucionar problemas como de plantearlos, para muchos científicos y tecnólogos es una entelequia irrealizable. La IA fuerte, de hecho, no vive sus mejores momentos. La pregunta es, ¿necesita esa IA que provoque la Singularidad ser realmente consciente? ¿Debe decirse a sí misma, después de conectarse, que piensa, luego existe? Dejemos la respuesta en el aire.

Los dos últimos escenarios supondrían una mejora de la raza humana. Esto parece más factible, siempre que se superen las barreras bioéticas y



morales —habitualmente religiosas— que surjan. Y no dudo que esto no representa un problema, ya que el argumento de la pendiente resbaladiza (o efecto dominó) en materia de ética deja de ser una falacia para convertirse en un hecho incuestionable, al menos hasta ahora.

Pero sigamos. Según Vinge habría unas consecuencias:

- Las IA superinteligentes crearán a su vez IA más inteligentes que ellas, y en tiempos cada vez menores, con lo que se acelerará el desarrollo tecnológico (este era exactamente el planteamiento de Jack Good cuando expresó lo que ya hemos mencionado de una “explosión de inteligencia”).
- El hombre se quedará atrás y dejará de ejercer el control. Habrá nuevas reglas para “la realidad”. Todo lo que ocurra, si lo presencia una persona normal, supondrá una sorpresa tras otra; y los humanos actuales, nosotros, no somos capaces de comprender ni de imaginar qué ni por qué ocurrirán las cosas concernientes al ámbito social y tecnológico. En palabras de Vinge, estaremos en la misma situación en la que ahora se encuentra un petirrojo.
- Se llegará a una era posthumana.

En cualquier caso, mientras una Singularidad producida por IA completamente artificiales representaría el fin de nuestra era, la que se produjera por mejoras en el ser humano permitiría al hombre seguir controlando su entorno y su propio destino como especie dominante del planeta.

Vernor Vinge, además de profesor universitario retirado de matemáticas y científico especializado en computación, es un famoso escritor de ciencia ficción que ha recibido importantes premios y ha enfocado la Singularidad hacia el *handicap* que supone para los escritores esta incapacidad de predecir el futuro y realizar ciencia ficción dura situando la acción en algún momento postsingularidad. Y peor aún será cuanto más cerca nos encontremos de ese momento. Ya en 1993 predijo una crisis para los escritores de ciencia ficción *hard*, que hoy en día parece patente, sea esta la causa o no. Pero sus novelas *True Names* (1981), *La guerra de la paz* (1984, Saga de las burbujas 1) y *Naufragio en el tiempo real* (1986, Saga de las burbujas 2), ya exploran el fenómeno de la Singularidad desde diferentes perspectivas. *True Names*, una novela que creo que no ha sido editada en España, precursora del ciberpunk, presenta un ciberespacio muy al estilo Matrix. La acción en ella transcurre en el preciso momento en que se produce la Singularidad. *La guerra de la paz* plantea que una serie de acontecimientos, con alguna catástrofe entre medias, ha retrasado la Singularidad. Y *Naufragio en el tiempo real* narra las peripecias de un



reducido grupo de gente que por determinadas circunstancias se sitúa al margen de la sociedad, que sí vive en un tiempo postsingular. Es una pena que, por motivos ajenos al deseo de Vinge y de los lectores —una reciente publicación, al parecer presuntamente irregular, de las dos novelas de Las burbujas en un solo tomo—, la reedición de estas obras durante los próximos años se antoje una labor ciertamente difícil. Pero nunca se sabe.

Estas tres estrategias para poder escribir acerca del futuro a medio o largo plazo son algunas de las que ha usado Vinge para ser coherente con su propia hipótesis, que ha mantenido viva en todas sus obras desde entonces. En otras ocasiones ha evitado la singularidad, como en la serie de *Un fuego sobre el abismo* (1992) y *Un abismo en el cielo* (1999), en las que construye unos mundos de modo que específicamente no ocurra la singularidad. De hecho, según esta trilogía inacabada habría regiones —como las cercanas al centro de las galaxias, que él denomina "las honduras sin pensamiento"— donde la inteligencia no es posible, y otras —en los confines de la galaxia— donde plantea la existencia de seres con otra forma de pensamiento, absolutamente incomprensibles para los humanos. Y por supuesto describe varios escenarios postsingularidad. Aunque esto parezca imposible y las historias puedan resultar algo duras de leer para el no aficionado, lo consigue. De hecho, en el libro *El monstruo de las galletas* (AJEC, 2007), que incluye la novela corta homónima y la también novela corta *Acelerados en el Instituto Fairmont*, también toca el tema de la Singularidad, situando ambas novelas cuando ésta se ha producido. El último libro de Vinge, *Al final del arco iris* (Nova, 2008), está ambientado en la misma ficción que *Acelerados*...En esta novela, un individuo que sufre Alzheimer es curado y rejuvenecido y se encuentra de repente en un mundo que no reconoce. El escenario sería, al parecer, el de la Internet que adquiere superinteligencia.

Por supuesto no es el único autor que ha escrito novelas con la Singularidad como telón de fondo. Tal vez quien más provecho haya sacado de esta idea de Vinge sea Charles Stross, conocido como el antipapa loco de la Singularidad.

Raymond (Ray) Kurzweil —llamado el *Cybernostradamus* por su capacidad predictiva, y otro genio con multitud de patentes a sus espaldas, como el lector en voz alta para ciegos y el OCR entre otras—, destacado tecnólogo y hombre polifacético que ostenta quince doctorados *honoris causa* en disciplinas diversas, incluyendo la música, ha publicado tres libros clave que hacen referencia al tema de la Singularidad:

- *The Age of Intelligent Machines* (1990, escrito desde 1986)
- *The Age of Spiritual Machines* (1999)



- *The Singularity Is Near: When Humans Transcend Biology* (2005)

Los dos primeros son más de tipo predictivo cronológico y en ellos el autor ha “acertado” en cosas que el resto de la humanidad ni sospechaba, como la importancia actual de las redes p2p, por citar sólo un ejemplo.

En 2001 publicó un ensayo muy discutido, *The Law of Accelerating Returns*, que podríamos traducir (confirmado por un asistente a la conferencia) como *Ley de los rendimientos acelerados*. En este ensayo expresó que la Ley de Moore, formulada en 1965 por el cocreador de Intel, Gordon E. Moore, y que actualmente sigue cumpliéndose —aproximadamente cada dos años se duplica el número de transistores por pulgada en un circuito integrado; también se enuncia a veces como que cada dos años se duplica la velocidad de procesamiento de datos en un ordenador—, se puede extender más allá de los transistores, hasta el procesamiento de información o capacidad de computación del sistema “mundo”. Cada vez que la ley se va a incumplir, cuando un muro se presenta infranqueable en el horizonte del progreso tecnológico, aparece otra tecnología u otro material que permite superar esa barrera. Y, según Kurzweil, esta tendencia nos llevará indefectiblemente a la Singularidad hacia 2045. Ese año la inteligencia no biológica, siempre según el autor, va a ser mil millones de veces más poderosa que toda la inteligencia humana actual. Signifique esto lo que signifique.

Kurzweil asegura que “Necesitamos aproximadamente 10 cuatrillones (10^{16}) de cálculos por segundo (cps) para proporcionar un equivalente funcional a todas las regiones del cerebro. Algunos cálculos son más bajos que estos en un factor de 100. Los Supercomputadores ya efectúan 100 billones (10^{14}) de cps, y llegarán a 10^{16} cps al finalizar esta década. Varios supercomputadores con un cuatrillón de cps ya están proyectados, con dos esfuerzos japoneses que buscan los 10 cuatrillones de cps al finalizar la década. Para el 2020, van a estar disponibles 10 cuatrillones de cps por aproximadamente mil dólares.”

La pregunta que me planteo al llegar aquí, y estoy de acuerdo con varios especialistas que acudieron a la conferencia, es si como elemento de juicio se debe considerar la capacidad o velocidad de computación o si la inteligencia es otra cosa. Parece que el asunto es bastante complicado.

El caso es que Kurzweil afirma que el patrón de desarrollo tecnológico culminará de un modo el que las distinciones entre máquinas y humanos se verán borradas a finales del siglo XXI, debido a lo que agrupa como GNR — Genética molecular, Nanotecnología y Robótica (incluyendo la biónica y los implantes quirúrgicos)—, lo que garantizará la continuidad humana sobre el



planeta. Pero se tratará de post-humanos, humanos versión 2.0, como bromea el propio Kurzweil.

Al contrario que otros futurólogos y escritores, que manejan futuros no muy gratos a los ojos del ser humano, Kurzweil es un tipo de lo más optimista. Según él, la existencia humana no correrá peligro cuando las IA y superhumanos superinteligentes dominen la Tierra, ya que estarán dotados de principios éticos muy avanzados, e incluso de una moralidad sofisticada. Serán ecologistas, conservacionistas... y pacifistas por convicción. Al ser tan sumamente inteligentes, se darán cuenta de que es una estupidez provocar guerras, esquilmar el planeta, destruir la biodiversidad y perjudicar a sus congéneres. Personalmente, permítidme que dude de que exista esta relación estrecha entre inteligencia y utopía. Estoy convencido de que los abogados tendrán más trabajo cuanto más avanzada (e inteligente) sea una sociedad. Pero esta es tan solo una apreciación personal.

Kurzweil predice que llegará el transhumanismo. Y se declara transhumanista convencido. Incluso antes de producirse esa fusión hombre-máquina, el propio ser humano antes de 2050 habrá logrado vencer a la enfermedad y el envejecimiento —primero ralentizándolo y luego deteniéndolo—, para a continuación llegar a idear procesos de rejuvenecimiento. Y todo ello culminará en la inmortalidad. Eso a los transhumanistas, junto con las mejoras biónicas, les parece fenomenal, frente a los críticos —sobre todo los luditas pero también algunos ideólogos arrepentidos, como Francis Fukuyama— que esgrimen argumentos tales como que el transhumanismo no es bueno ni sano, que va *contra natura* y que no se debe desafiar a las leyes de la naturaleza ni seguir nuestros impulsos e intereses sin respetar el principio de cautela. Pero resulta que Kurzweil tiene motivos personales que tal vez supusieran el pistoletazo para su lucha a favor del transhumanismo —de hecho milita, o incluso preside, algunas organizaciones transhumanistas—: a los 35 años le diagnosticaron una diabetes tipo II, una especie de intolerancia a la glucosa, y desde entonces está obsesionado con los cuidados personales. Su objetivo es alargar su vida y finalmente hacerse inmortal. En primer lugar pretende (literalmente) “reprogramar su metabolismo”. Para ello ha comenzado por crear su propio régimen dietético —litros diarios de agua alcalina y té verde, abundante vino tinto y 250 píldoras de suplementos al día, que actualmente ha rebajado a 150, y se inyecta más suplementos los fines de semana—. Practica deporte y lleva un estricto régimen de chequeos y análisis de sangre, y asegura no comer nada que no sea comida orgánica y sin azúcares. Su objetivo es ir cambiando el metabolismo lo suficiente como para llegar a vivir el momento en el que la terapia génica y la biónica sean un hecho, lo que le curará por completo y a su vez le permitirá llegar al momento en el que la nanotecnología —y de nuevo la biónica— sea capaz



de hacernos inmortales. De hecho asegura que sus análisis de fluidos indican que hasta ahora ha retrasado su envejecimiento más de 10 años. Ya ha reservado, de todos modos, su propio criogenizador en ALCOR, supongo que por si no llega a cruzar alguno de esos dos puentes que le harán inmortal.

En cualquier caso, Kurzweil predice que a partir de la Singularidad la evolución biológica humana dejará de tener importancia y que toda la evolución será tecnológica. Esto implica algo que a mi entender es importante: al menos en lo que respecta al ser humano, y posiblemente el alcance sea mayor, de producirse este cambio que vaticina Kurzweil estaríamos hablando de un salto desde un escenario de evolución darwiniana —cambios genéticos al azar e impulso de la evolución mediante selección natural, proceso lento y de rumbo errático que, al menos desde que el *Homo sapiens* está sobre la faz de la Tierra, no ha producido variaciones en nuestra especie— a otro lamarckiano —rápido, guiado, determinista, y con herencia (e incluso actualizaciones periódicas de hardware y software) de caracteres adquiridos—. Es más, este sería el momento de hablar sobre esa putativa brecha social que se produciría dependiendo de la diferencia de poder adquisitivo que habría entre unos y otros representantes de nuestro género. Pero es una materia de debate que dejaremos para otra ocasión.

Algunos sitios donde se puede leer sobre este tema:

<http://mindstalk.net/vinge/vinge-sing.html>

<http://www.singinst.org/>

<http://www.jornada.unam.mx/2008/02/21/index.php?section=ciencias&article=a02n1cie>

<http://axxon.com.ar/not/173/c-1730024.htm>

http://webpri.wmaker.net/tendencias/forum-El-transhumanismo_m18086.html

<http://guerrerosdesingularidad.wordpress.com/category/singularidad/>

<http://www.acceleratingfuture.com/people-blog/?p=1611>

<http://www.spectrum.ieee.org/jun08/6306>

<http://proyectoliquido.net/h2blog/121/especial-h2blog-la-singularidad-tecnologica/>

<http://proyectoliquido.net/h2blog/236/singularidad-tecnologica-cuando-los-humanos-trasciendan-la-biologia-entrevista-a-ray-kurzweil/>

<http://es.wikipedia.org/>



Ficha de la conferencia:

Fecha: Jueves 20 de noviembre de 2008, de 19:00 a 21:00

Localidad: Madrid

Lugar: Círculo de Bellas Artes, C/Alcalá 42

Ponentes: Dr. Federico G. Witt, Doctor en Biología Molecular, UAM; Prof. José Antonio López Guerrero, Director de Cultura Científica, UAM; Dña. Margarita Arroyo Macarro, Coordinadora Cultura Científica, UAM.

Autora: Dr. Federico G. Witt, Córdoba, España.

Artículo inédito. Teorema Z. www.libroandromeda.com

El autor: Dr. Federico G. Witt, ha cedido a Libro Andrómeda el derecho de publicación de esta obra en nuestra web, con la siguiente condición, de acuerdo con las opciones de protección de los derechos de propiedad intelectual existentes para la difusión en internet:

Reconocimiento – Sin obra derivada – No comercial: El material creado por un artista puede ser distribuido, copiado y exhibido por terceros si se muestra en los créditos. No se puede obtener ningún beneficio comercial. No se pueden realizar obras derivadas.