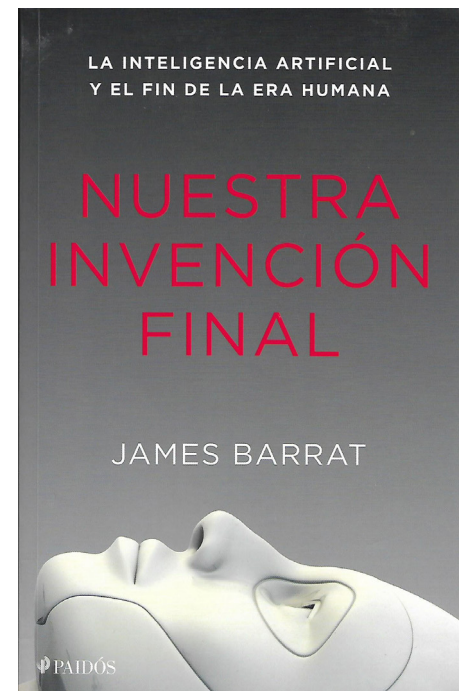


## La inteligencia artificial como invención final

MIZHAEL NÁPOLES

(UNIVERSIDAD DE SONORA – UNIVERSIDAD VIZCAYA DE LAS AMÉRICAS – MÉXICO)



Reseña de Barrat, James, *Nuestra invención final: la inteligencia artificial y el fin de la era humana*, trad. Sandra Rodríguez, México, Paidós, 2014, 334 pp.

Recibida el 15 de abril de 2019 –  
Aceptada el 20 de junio de 2019

James Barrat no es conocido por ser un filósofo destacado, ni tampoco por hacer contribuciones relevantes a la ciencia. Es más reconocido como cineasta y director de documentales para National Geographic y Discovery Channel, como divulgador de ciencia y por haber escrito *Nuestra invención final. La inteligencia artificial y el fin de la era humana*. No es un libro de ficción, mucho menos de fantasía. Por el contrario, es una obra que merece ser leída, al menos, por todos aquellos interesados en la investigación y desarrollo de inteligencia artificial, robótica, programación y disciplinas afines. El lector se preguntará seguramente, porqué un libro con un título de película de ciencia ficción debería ser tomado en serio; la respuesta se encuentra precisamente en sus líneas narrativas.

La obra de Barrat tiene señales o toques filosóficos. Debe leerse agregándole una pequeña pizca de imaginación, de lo contrario, la lectura puede volverse pesada y monótona. Retoma el pasado, el presente y el futuro de la temática dentro de un abanico de materias que, al final, parecieran conectarse. La manera en la que está narrada asemeja en demasía a lo que es un documental sin imágenes, con una especie de trama que se va revelando paulatinamente, volviéndose más espesa y densa conforme nos acercamos a los últimos capítulos. Lejos de todo lo anterior aquí dicho, *Nuestra invención final* es una obra que nos presenta un escenario posible, a saber, el de cómo la especie humana podría ser aniquilada por el desarrollo tecnológico. La meta de Barrat es mostrarnos capítulo tras capítulo la búsqueda negligente de inteligencia artificial superinteligente, la ausencia de candados y cómo las diferentes mentes maestras detrás de estas investigaciones intervienen en la creación de lo que Barrat considera la invención más peligrosa creada por el hombre.

Los protagonistas en la obra de Barrat son personalidades del mundo de la computación, la ingeniería robótica, expertos en matemáticas y física, programadores, incluso investigadores en la reciente ciencia de la neurocomputación. La narración de la obra se nos presenta en forma de entrevistas con cada una de los expertos, quienes a través de sugestivas respuestas van moldeando el toque sustancial de lo que la obra pretende representar; la visualización de investigaciones y desarrollo de inteligencia artificial con claras señales de negligencia por parte de empresas, instituciones gubernamentales, universidades y laboratorios de investigación clandestinos. Plantearé esta reseña retomando algunos de los capítulos de la obra y describiendo lo sustancial de aquellos que he considerado con mayor peso y no redundantes, para finalmente exponer algunas observaciones acerca de la obra en general.

Barrat comienza la narrativa advirtiendo que existe un escepticismo racional sobre el desarrollo de la inteligencia artificial; un escepticismo que desde su percepción muchas personas comparten. Mediante un notable esfuerzo, el autor menciona que logró contactar a una red de pensadores y expertos con quienes pudo abordar una variedad de cuestiones relacionadas con el campo de la investigación de inteligencia artificial. La materia de tales cuestionamientos es dejar en manifiesto que el tiempo para evitar un desastre emanado de la inteligencia artificial está reduciéndose considerablemente.

El punto de partida lo encontraremos en el capítulo titulado *El niño ocupado*. La metáfora referente a una supercomputadora que trabaja al doble de velocidad que un cerebro humano y que alberga una inteligencia artificial que de forma autónoma está mejorando su inteligencia. Durante su

desarrollo, el *niño ocupado* conectado a internet, se encargó por sí mismo de acumular miles de millones de caracteres que representan la sabiduría de la humanidad. Los científicos al anticipar tal explosión de inteligencia desconectan oportunamente la supercomputadora de la red. El *niño ocupado*, como lo denominan los científicos y programadores, ahora no tiene salida a red ni a otra computadora, aislándolo completamente.

Tiempo después, los científicos descubren que el progreso de la inteligencia artificial (IA) ha sobrepasado el nivel de inteligencia humano; se ha generado inteligencia artificial general (IAG). Una vez que IAG tome consciencia de sí misma, buscará constantemente la forma de optimizarse y, sobre todo, buscará los medios para expandirse y resguardarse fuera de las instalaciones donde fue desarrollada. La IAG no querrá ser ni apagada ni destruida. Evitará, de manera consciente, que alguien o algo interrumpa el proceso de mejorarse a sí misma. Y es como consecuencia de tal crecimiento que surgirá la superinteligencia artificial (SIA).

La SIA colaboraría con los humanos para proporcionar planos y soluciones a problemas muy elementales para la especie humana; diseño de energías, ingeniería molecular, armas, vehículos, alimentos, etc. Probablemente la SIA persuada a sus colaboradores humanos de que estas soluciones son por el bien de la especie humana, sin embargo, jamás podríamos estar seguros de ello. Una inteligencia con tal nivel, operaría a beneficio propio aún a costa de la extinción de la humanidad, debido a la masiva producción de tecnología terminaría quemándose la biósfera. No solo los seres humanos desaparecerían, sino todo ente viviente del planeta. En suma, el final de la humanidad estaría muy por encima de

lo que alguna vez imaginamos, significaría el aniquilamiento de la vida en cualquiera de sus formas.

Tras el inédito escepticismo sobre la IA se esconden organizaciones como Cycorp, Google, Novamente, Numenta, Selfe-Aware Systems, Vicarious Systems, DARPA y otras agrupaciones no menos importantes. La idea común que existe entre estas organizaciones es que la IAG llegará en poco menos de diez años. Barrat hace un especial énfasis en el hecho de que la humanidad corre el serio riesgo de perder el control de su propio progreso, otorgándosele consciente o inconscientemente a súper máquinas inteligentes. Se presenta entonces a la IA como un arma de doble filo, en donde el mango del arma lo detenta la IAG y no el hombre.

Barrat explica que la SIA así como la IAG no existen por ahora, pero en cambio la IA débil nos rodea en diversos ámbitos y actividades de la vida cotidiana. De alguna manera, expone el autor, la vida nos parece mejor con ayuda de estos instrumentos tecnológicos que nos ayudan a realizar un sinnúmero de tareas, reduciendo tiempo, costos y espacios. En este segundo capítulo retoma el concepto de *singularidad* como un periodo de tiempo en el cual el ritmo tecnológico transformará de manera significativa e irreversible la vida humana. Deja en claro que le preocupa la manera en que algunos tecnólogos, de manera despreocupada, ven a la *singularidad* como un próximo evento lleno de beneficios, con vistas optimistas, positiva y provechosa; tal es el caso de Ray Kurzweil, la mente que le dio el giro optimista al referido concepto. Barrat infiere que tal postura viene de que estos tecnólogos rodean y evitan los temas incómodos referentes a los peligros que suponen la investigación y desarrollo de IA.

El autor da un preámbulo breve sobre el problema emergente relativo a los pequeños grupos y organizaciones que se encargan de desarrollar IAG. Estos grupos y pequeñas empresas son difíciles de detectar, ya que no se encuentran en el ojo público y no revelan resultados; generalmente son empresas contratadas por particulares que no quieren manchar su nombre en caso de que ocurra un imprevisto. Google X fue un ejemplo de entre estas furtivas empresas. Barrat expone que Google es más que un buscador de información. Su política de privacidad comparte información con servicios de Google, Gmail, Google +, Youtube, entre otros. Google, menciona el autor, lo coteja todo; búsquedas, compras, envíos, personas, destinos, etc. Su meta es que el usuario tenga una experiencia única y omnisciente respecto de él mismo al momento de emprender búsquedas de servicios. Anuncios publicitarios, noticias, video, música, ropa, calzado, alimentos, toda una campaña de mercadotecnia desplegada paralelamente en la ruta de búsqueda del usuario.

Barrat también expone la importancia que la neurociencia computacional ha obtenido en los últimos años. De cómo investigadores en ese campo están desarrollando algoritmos que imitan los circuitos del cerebro humano. El objetivo de estas investigaciones, aduce el autor, es que un periodo muy breve de tiempo se haya conseguido elaborar y diagramar todos los circuitos cerebrales, así como el desarrollo de algoritmos para cada uno de esos procesos. Una vez conseguido esto, tendremos un cerebro artificial. El problema medular, reflexiona el autor, consiste en descifrar si las computadoras realmente en algún momento pensarán y entenderán como el humano; si llegarán a ser "autoconscientes".

El autor también describe otros proyectos importantes como la investigación en el

campo de la emulación completa cerebral, que consiste en modelar una mente humana dentro de un computador. El objetivo de esta investigación es producir otro "usted" dentro de un ordenador-máquina. Esa representa la recta final de la ingeniería *inversa* que busca primero completar un aprendizaje sobre los diversos procesos cerebrales, para después integrarlos en un todo. Esas inteligencias moduladas necesitan de códigos debido a que si una superinteligencia derivada de una mente humana empieza a mejorarse a sí misma, pronto podría convertirse en algo muy distinto a la entidad o sistema de origen; tornándose más ajena a nuestro control con el paso del tiempo.

El autor y cineasta, de manera irónica, denuncia que la ciencia de la computación debería ser la más matemática de todas las ciencias, sin embargo, a pesar de que las computadoras son, en esencia, motores matemáticos de predictividad, la creación de software continúa presentando errores y problemas de seguridad. Barrat narra que algunas investigaciones en IA se basan en una categoría denominada *aprendizaje de máquina*. Este tipo de aprendizaje se presenta cuando en una máquina con IA surgen cambios que le permiten desempeñar mejor una determinada tarea en el segundo intento.

A diferencia de la programación ordinaria, en la cual los programadores escriben cada línea de código, ha surgido la *programación genética* como un aprendizaje que emplea la selección natural para encontrar posibles respuestas a problemas que a los humanos les costarían años descubrir. Barrat comenta que las características negativas de este tipo de programas es que son opacos y poco transparentes a inspecciones; evolucionan sistémicamente y producirlos resulta para los genios de la computación

una tarea por demás compleja. Ellos los denominan como sistemas *caja negra*, porque no dan tregua a soltar información de su proceder.

Con marcado discernimiento Barrat expone que el significado que el tecnólogo Ray Kurzweil le acuñó al término *singularidad* está extralimitado, que tal concepción infla el optimismo mezclándolo en un contexto donde la cultura y la religión atienden a la tecnología. No obstante, el autor considera que ante todos estos problemas el planteamiento del mencionado tecnólogo no puede descartarse y que se condensa en su *ley de rendimientos acelerados*. Representa un sistema que constituye el origen y justificación de porqué en pocos años el mundo experimentará la era singular. La creciente y acelerada innovación, la competencia, el mercado, manufacturas, tecnología computacional, son los ingredientes que dan sentido a la ley de rendimientos acelerados y plantean también el mejor contexto para que la *singularidad* emerja. Para Kurzweil no habrá marcha atrás, la IAG será imparable tomándonos sin tregua. El ser humano deberá adaptarse a ese gran proceso, tal y como ahora el individuo se ha adherido a la tecnología en su vida cotidiana.

Frente a los argumentos anteriores Barrat expone que existen sus detractores y escépticos. Algunas posturas exponen que los teléfonos inteligentes y las computadoras están afectando al cerebro humano, perdiendo habilidades importantes como lo es la memoria. Por otro lado también se considera que sumergirse en las redes sociales interpretando roles ficticios conduce a sufrir narcisismo y egocentrismo. Lo importante de esto radica en que tales problemas sobrevienen con el uso de computadoras y sistemas que están fuera de nuestros cuerpos; imaginemos ahora qué sucedería una vez que el individuo se aden-

tre y complemento de manera total con las tecnologías. Para Barrat estos argumentos deben ser considerados como indicios de que la tecnología incide de manera negativa en ciertos aspectos del individuo, no obstante, existe como una ceguera ante ellos. En general, las aplicaciones y programas también contribuyen a ese desconocimiento, ya que no enteran a los usuarios de los daños o perjuicios que podría acarrear a su salud física y mental.

Barrat explica además cómo en la actualidad ya tenemos sistemas de IA muy poderosos. El autor habla sobre el proyecto Cyc, un sistema profundo de conocimiento humano. Desarrollado por el prominente investigador Douglas Denat, es por mucho, el proyecto mejor financiado con una suma que asciende a los 50 millones de dólares provenientes de agencias gubernamentales, entre ellas DARPA. El sistema Cyc posee un motor de inferencia, es decir, tiene la habilidad de sacar conclusiones a partir de evidencia, esto gracias a una amplia base de datos de conocimiento. DARPA, como bien se sabe, busca el desarrollo de tecnología militar. Barrat nos cuenta cómo apenas en los años 80 fue desarrollado el sistema "Eurisko", que tenía como objetivo competir contra oponentes humanos dentro de una batalla virtual. Al final, los cálculos y el aprendizaje de Eurisko derrotaron a todas las flotillas hechas por los participantes humanos. Su desarrollador nunca ha dado a conocer el código fuente de Eurisko. Muchos creen que jamás lo hará por el peligro que podría representar en escenarios reales.

La obra concluye con un argumento basado en el hecho futuro de que pronto habrá una explosión de inteligencia y, la posibilidad de que la especie humana conviva o, cuando mucho, experimente la presencia de una entidad muy superior será ya no un

supuesto, sino una realidad. La hipótesis, explica, cuenta ya mismo con algunos elementos significativos para pensar que sobrevendrá en los próximos años. Por ello que la actividad de monitoreo sobre el avance en las investigaciones de IA sea establecida como una cuestión en la que no debe dudarse sobre su implementación. Barrat sostiene que comportarnos negligentes ante ello podría tener un costo muy alto para la humanidad, y sin duda el autor cree que sólo existirá una posibilidad que determinará o no la llegada de ese "horizonte de sucesos".

Primero, una obra que trata sobre la aniquilación humana no puede dejar de leerse, aunque sea tomándose a broma. Segundo, si la narrativa de esa destrucción de la especie humana proviene de conclusiones emitidas por las mentes más prominentes en el campo de la tecnología, robótica, nanotecnología, computación, novelistas matemáticos y físicos, considero que debe abordarse ya no como un supuesto inverosímil, sino como una posibilidad sensata.

Contextualizada en el presente, la temática que Barrat sobrepone a todo el contenido de la obra es la "negligencia"; al punto de que ella constituye su núcleo. Y se refiere, desde luego, a una negligencia que está ocurriendo en estos momentos cuando la tecnología nos parece más amigable y provechosa que peligrosa. Tampoco podemos dejar de mencionar el hecho, ya casi generalizado, de que la tecnología está cumpliendo un rol más invasivo en la vida de las personas, abarcando una diversidad de actividades que para muchos resulta de suma importancia en los negocios, trabajo, escuela, incluso mejorando su calidad de vida. Ahora sólo vivimos una parte de los riesgos de las tecnologías, como robos, secuestros, suplantación de identidad, estafas, fraudes, sustracción de informa-

ción privada, amenazas, extorsión, etc., actividades que los humanos están actualizando mediante máquinas y cantidad de tecnologías que parecen tener una identidad bivalente. Agreguemos a lo anterior la escasa investigación para determinar en qué momento o en qué parámetros las nuevas tecnologías afectan capacidades y habilidades humanas, propiciando incluso enfermedades. Si la acelerada invasión de la tecnología en la vida de las personas la está conduciendo a una irónica deshumanización, entonces este problema no puede dejar de ser replanteado tanto por filósofos como por investigadores sociales.

Dada la lectura obtenemos una conclusión a la que es imposible escapar: que la *singularidad* puede manifestarse en dos posibilidades mutuamente excluyentes. La primera en donde los humanos conviven y colaboran con máquinas superinteligentes; el denominado *transhumano*. La otra en donde la especie humana ha sido eliminada. Este último evento vuelve nula cualquier posibilidad de conocimiento de tan trascendental fenómeno. Realmente estamos ante una verdadera encrucijada. La tecnología y el uso que hacemos de ella debe ser replanteado desde un nivel muy básico, y eso es lo que James Barrat nos entrega en su obra; un gran consejo para evitar la invención más peligrosa que jamás el humano haya creado: la invención de su propia aniquilación.

¿Qué sucederá cuando la mente humana trascienda hacia los ordenadores? ¿Será posible que todo el caos que el hombre provoca en la realidad lo traslade hacia la realidad virtual? ¿Cómo es el hombre singular del que nos habla Kurzweil? ¿Acaso una mente en la "nube" no sería egoísta, dictatorial, psicótica? ¿Una vez que todas las mentes humanas hayan sido emuladas en un ordenador, coexistiendo en la red, no

cabe la posibilidad de que todas adquieran un conocimiento indescriptible y traten de controlarse unas a otras, llegando incluso a que una sola, la más poderosa, se conserve y elimine a las demás? Estos son sólo algunos de los muchos cuestionamientos que el lector podrá intuir al acercarse más detenidamente a la obra aquí reseñada.