

Asistidos por el soberano electrónico. Utopías y distopías de la inteligencia artificial

Assisted by the Electronic Sovereign. Utopias and Dystopias of Artificial Intelligence

Francisco Martorell Campos

Universitat de València
francisco.martorell@uv.es

Andoni Alonso Puelles

Universidad Complutense de Madrid
andonial@ucm.es

ISSN 1989-7022

RESUMEN: En este artículo tratamos de entender la Inteligencia Artificial (IA) como el modelo de superprótesis y la utopía asistencial. Para ello revisamos dos posturas enfrentadas: aquellas que mantienen una visión utópica de su desarrollo y aquellas que ven el final distópico en su consecución. En el momento actual y en medio del cambio tecnológico que sufrimos, tanto una como otra significan una respuesta equivocada y estéril para este problema. La utopía vendida sobre todo por los tecnólogos tiene visos de crear más problemas que los que resuelve. La política y la organización social deben acomodarse a un escenario nuevo donde la automatización se produzca en un contexto de renta básica, por ejemplo. Es necesario encontrar un equilibrio en este antagonismo y colocar a la asistencialidad tecnológica en su lugar, dejando espacio para actividades que, por carácter moral como el cuidado, deben ser desempeñadas por otros humanos.

ABSTRACT: This paper tries to understand Artificial Intelligence (AI) as the model for super-prosthesis and assistance utopia. For that we review two confronted stands: the utopian vision to develop IA till the end and the dystopian attitude that would imply to achieve that goal. At the present moment and in the middle of a technological change, both attitudes are sterile and wrong to cope with that problem. Technological utopianism seems to create more problems than solve them. Politics and social organization must transform a new scenario where automatization would take part of a basic income right, for instance. It is required to find a balance on that antagonism and try to place assistant technology circumscribed to some domains. Some human activities should remain carried out by humans due to their moral element such as care.

PALABRAS CLAVE: Inteligencia artificial, cuidado, asistencia, utopía, distopía, economía, sociedad

KEYWORDS: Artificial intelligence, care, assistance, utopia, dystopia, economy, society

1. Introducción

Como es de sobra conocido, el avance tecnológico inspira valoraciones tremendamente dispares. Aún y a riesgo de simplificar en exceso, muchas de ellas articulan o bien un relato *tecnóforo* o bien un relato *tecnófilo*. De acuerdo con el primero, la tecnología desempeña, de un modo u otro, el rol de agente condenatorio y alienante de la humanidad. De acuerdo con el segundo, el de agente salvador y emancipador. Estas lecturas extremistas, deudoras al unísono de alguna modalidad de determinismo tecnológico e históricamente ligadas al ambivalente desarrollo de la modernidad, se han visto potenciadas en tiempos recientes a la luz de las tecnologías postmodernas, tales como la realidad virtual, la biotecnología, la nanotecnología, la cibernética o la inteligencia artificial, disciplinas que apuntan a la consecución de objetivos hasta hace poco inconcebibles o circunscritos a la ciencia ficción.

Las tecnologías asistenciales medran en este ambiente de dispositivos impregnados de aureola futurista. Protagonistas de un campo de negocio y de una atención mediática al alza, comprenden desde los aspiradores autónomos y las máquinas de autoservicio a los vehículos aéreos o terrestres no tripulados, desde *apps* para casi todo tipo de necesidades a productos destinados a facilitar operaciones quirúrgicas, desde dispositivos plulares para el cuidado, apoyo o rehabilitación de enfermos, ancianos y gente con alguna



discapacidad a sensores de vigilancia. Excepción hecha de las fabricadas con fines militares, publicitarios o invasivos (que son multitud), bienvenidas sean, en general, el resto.

En este trabajo vamos a centrarnos en la inteligencia artificial (IA), ingenio del que se está hablando mucho últimamente al amparo, muy a menudo, de las tecnologías asistenciales. Más en concreto, vamos a registrar los sueños y pesadillas fundamentales que giran alrededor de la inteligencia artificial fuerte (IAF), línea de investigación que aspira a la fabricación de un mecanismo cognitivo que según los adeptos más confiados superaría al intelecto humano y sería capaz de mejorarse a sí mismo e incrementar sus facultades *ad infinitum*. El interés que abriga semejante espécimen puramente hipotético radica en que nos permite sopesar los pros y contras, los supuestos e intereses de la mayor asistencia tecnológica imaginable: la asistencia política, protagonizada por una IA que se encarga de planificar el orden socio-económico y de gobernar, en mayor o menor grado, a la humanidad, cual *soberano electrónico*.

El presente artículo pasará revista a los imaginarios esenciales florecidos de tal estampa asistencial desde el prisma de la filosofía política y la crítica cultural. Nada más lejos de nuestra intención que echar mano de la filosofía analítica de la mente y cavilar por extenso sobre la viabilidad o inviabilidad de construir un aparato así. Los debates acerca de si el test de Turing es fiable o no, la existencia de los *qualia* verosímil o inverosímil, el “problema del marco” franqueable o indisoluble o las conclusiones colegidas por Searle (2001: 39-42) de la “habitación China” convincentes o, como dice Dennet (2006: 87-88), insostenibles no tendrán cabida aquí. Vaya por delante que para nosotros no existe en principio ninguna contradicción filosófica o conceptual definitiva que impida construir una máquina inteligente tal y como la percibe la IAF. Por suerte o por desgracia, existen, eso sí, escollos técnicos y epistemológicos monumentales, tantos que a día de hoy no pasa de ser una fantasía inalcanzable, un ideal especulativo y propagandístico que refleja valores más que hechos. De ahí que sean tan sospechosas las proclamas de los voceros digitalistas que anuncian su inminente gestación (que siempre termina, por cierto, postergándose, como el apocalipsis).

Partamos de un hecho más o menos aceptado: el ser humano es un ser protésico. La tecnología con la que se inventa, desarrolla y extiende puede resumirse, en gran medida, en la noción de prótesis. Esta idea forma parte de la historia de la filosofía, desde los clásicos a los contemporáneos. Basta con recordar la noción de “ser incompleto” que asoma en la mitología griega hasta llegar a los estudios de Arnold Gehlen (1993). Incompletud, es preciso matizar, que lejos de desvalorizarlo estimula, precisamente, la forja y proliferación de las fabulosas cualidades del espíritu humano. Ni que decir tiene, que las prótesis han de servir para corregir déficits o mejorar la condición humana, a la tarea de favorecer una vida digna y autónoma. El hándicap es que la inteligencia artificial está evolucionando hacia una significación que desborda la noción de “prótesis sincrética” de la que hablaba Maldonado (1998: 158), hasta el punto de investirse en *superprótesis*. Más que una herramienta, amenaza con constituirse en un sistema que nos absorbe y dirige por completo, que nos obliga, así cree Bostrom (2016), a determinar ya mismo protocolos de precaución. Los de Asimov, no sirven.

Siendo comprensibles los temores que despierta la IA, las páginas siguientes mostrarán, sin embargo, que la repulsa distópica estándar a la intervención e implicaciones del soberano electrónico esconde tras sus aciertos presunciones maliciosas, un diagnóstico moral relativo a la automatización, la asistencia tecnológica y el sentido de la existencia humana políticamente muy problemático y discutible, por no decir, directamente, rancio y descafeinado.

2. La computación es la existencia, y los humanos imperfectos

De entre los defensores más esotéricos de la posibilidad de engendrar una inteligencia artificial fuerte (es decir, capaz de replicar e inclusive superar las capacidades cognitivas humanas), destacan los adalides del *pancomputacionismo*, credo según el cual la realidad material en su conjunto, desde los genes hasta el cerebro, es un ordenador procesando información. Valga la siguiente muestra:

Para un físico... no hay sistema físico que no sea una computadora... Cada electrón, cada fotón o cualquier partícula elemental almacena bits de datos... La existencia física y el contenido de información están inextricablemente ligados. De la información, de los bits, sale cada eso que haya en el mundo... Conducido y propulsado por el modelo estándar a modo de programa, el universo computa campos cuánticos, compuestos químicos, bacterias, seres humanos, estrellas y galaxias... La computación es la existencia (Lloyd, 2005: 59, 66-67).

Robert Jastrow desarrolló en los albores de la postmodernidad una visión posthumanista del porvenir cristalizada en una inteligencia artificial escandalosamente superior a la humana. Las máquinas actuales, escribe en *El telar mágico*, son inteligentes, pero comparadas con las futuras apenas balbucean. Su CI será tan portentoso que a su lado pareceremos insectos. Quien sepa escrutar los acontecimientos tecnológicos presentes descubrirá a la nueva raza superior gestándose poco a poco, preparándose para desalojarnos a medio plazo del trono de la existencia. La profecía, típica de los teóricos de la singularidad tecnológica, parece a primera vista agorera, la típica advertencia fatídica sobre la extinción o esclavitud que nos espera a manos de algún remedo de Skynet (IA malvada de *Terminator*) o Hal-9000 (célebre ordenador de *2001*). Sin embargo, lejos de extraer conclusiones catastrofistas, Jastrow pronostica que los seres humanos salvarán la papeleta rubricando un pacto mutuamente beneficioso con las máquinas ultrasapientes:

La era de la vida basada en la química del carbono está encaminándose a su fin sobre la Tierra, y una nueva era de vida basada en el silicio —indestructible, inmortal, con infinitas posibilidades— está empezando... El ordenador —una nueva forma de vida dedicada al pensamiento puro— cuidará de sus asociados humanos, los cuales subvendrán a sus necesidades corporales con electricidad y piezas de repuesto. El hombre se ocupará también de la reproducción de los ordenadores... A cambio, el ordenador atenderá a las necesidades sociales y económicas del hombre. Ésta será su salvación en un mundo de aplastante complejidad (1998: 171).

El extracto amalgama dos variedades del pancomputacionismo a tener en cuenta:

- i) Pancomputacionismo biológico: Según Jastrow, la inteligencia artificial es el producto final de una dinámica lamarckiana de progreso teleológico y necesario desplegado en la naturaleza caracterizado por alumbrar gradualmente seres con mayor capacidad de almacenamiento, procesamiento y transmisión de información, siendo así que las especies menos dotadas al respecto tienden a desaparecer. El *homo sapiens* se impuso durante la selección natural a *australopithecus* y *neanderthales*. Ha ocupado la cima de la inteligencia durante milenios, pero no es la última palabra. La atrofia que sufre el tamaño de su cerebro desde hace 100.000 años provoca que no pueda lidiar con el computador, competidor capaz de evolucionar intelectivamente de manera extraordinaria en solo una generación, camino de cotas intelectivas sin precedentes. La trama jastrowniana tiene implicaciones tan sugerentes como metafísicas. Movidado por el impulso de incrementar la inteligencia en el mundo, el computador biológico llamado naturaleza tardó millo-

nes de años en producir al ser humano, especie que creará, a su vez, a un computador postbiológico que incrementará el volumen de inteligencia en apenas unas décadas, sin que se vislumbren límites a dicho crecimiento. El plan de la naturaleza es, de esta forma, heredado a través del hombre por un ser enteramente sintético que lo consuma mucho mejor.

- ii) Pancomputacionismo social: Las directrices computacionales de la naturaleza imperan, supone Jastrow, igualmente en la sociedad. Habitamos la era de la comunicación, “un mundo de aplastante complejidad” donde la producción e intercambio de información se incrementa a un ritmo tan acelerado que la facultad humana de procesamiento se ve desbordada, sumiéndonos en la impotencia y la fragilidad, fraguando la sensación de desconcierto, ansiedad e incertidumbre que nos invade. Dadas, viene a decir Jastrow, nuestras carencias epistémicas, lo idóneo es que un *soberano electrónico* nos asista y se haga cargo de la coyuntura cuando sea técnicamente viable. Gracias a su carencia de *pathos* e intereses partidistas, conducirá imparcialmente los asuntos políticos, sometiendo los temas más delicados a los dictados cuantitativos (objetivos, infalibles) de la razón pura. Masticará toneladas de datos en un microsegundo, lo que le habilitará para atender a nuestras “necesidades sociales y económicas” y domesticar la complejidad.

Tres décadas antes de que Jastrow lanzara la propuesta, Norbert Wiener se preguntó en 1950:

¿No puede imaginarse una máquina que recoja tal o cual tipo de información, por ejemplo sobre la producción y el mercado, y luego determine, en función de la psicología media de los seres humanos y de las cantidades que es posible medir en un caso determinado, cuál puede ser el giro más probable que tome la situación? ¿No puede concebirse siquiera un aparato estatal que abarque todos los sistemas de decisión política? Podemos soñar con el tiempo en que la *machine à gouverner* tal vez llegue a suplir —ya sea para bien o para mal— la obvia insuficiencia actual del cerebro cuando se ocupa de la acostumbrada maquinaria de la política (Roszak, 1988: 241).

Obsérvese cómo la hipótesis de la complejidad social inasible y el correspondiente proyecto del soberano electrónico barruntado por Jastrow, Wiener y tantos otros después de ellos se asienta sobre una antropología muy específica: la que predica la imperfección innata e irremediable del ser humano, traducida de forma privilegiada en su incompetencia para gestionar la esfera pública coetánea. De una antropología tal se deduce que las metas utópicas de eliminar definitivamente la guerra, la pobreza y la discriminación solo resultará factible si una *machine à gouverner* detenta el poder estatal. Theodore Roszak cuestionó hace ya cuatro décadas tamaño convencimiento. Su testimonio, repárese si no en el calado que poseen los postulados emitidos por Kurzweil y Vinge hoy día, es más actual que cuando fue publicado:

¿Hemos creado algo que se parece a una mente, una mente que se adapta mejor a las condiciones alienantes de la sociedad moderna, que es más capaz de soportar la presión, la angustia, la tensión moral? Si así fuera, podría tomarse como un juicio condenatorio de la inhumanidad del orden que hemos creado para nosotros mismos. Pero es claro que algunos científicos de la informática lo consideran como una acusación contra la mismísima naturaleza humana; poseemos una mente que no es apta para sobrevivir (*Ibid.*: 49).

3. Visiones apocalípticas de la inteligencia artificial en la ciencia ficción

De acuerdo a la ciencia ficción catastrofista vinculada al tema que nos ocupa, si no detenemos el avance de la inteligencia artificial corremos el serio peligro de que ésta acabe engendrando autoconciencia y rebelándose contra sus creadores, situación que amenazaría con extinguirnos o reducirnos al papel de siervos. Augurios así, son vehiculados en la cultura popular contemporánea por *Matrix* (1999), *Transcendence* (2014), *Battlestar Galáctica* (2003-2009) y *Westworld* (2016-2018), entre otras muchas películas y series.

Samuel Butler reflexionó sobre las amenazas citadas en *Erewhon*, clásico distópico de 1872. Partiendo de las mismas presunciones lamarckianas que Jastrow y sus compinches, Butler llega a una resolución inversa. En la imaginaria *Erewhon* la construcción de artefactos posteriores al año 1600 se encuentra prohibida y castigada por ley (el reloj del visitante es destruido al llegar). “El libro de las máquinas”, obra de cabecera del intelectual del lugar, explica cómo las susodichas adquieren poco a poco potestad, exigiendo cuidados al alza, invadiendo más y más ocupaciones, haciéndose imprescindibles para cualquier labor. Es la derrota, aún sibilina, de la humanidad, cuyo talento se consagra a la producción seriada de mecanismos en un ambiente de materialismo ramplón y falso bienestar, hostil con el arte, la religión, la poesía. Lo peor está por llegar. Las máquinas evolucionan, y lo hacen a velocidades inmensamente superiores a las transitadas por los organismos biológicos. Si no se interrumpe su progresión, alcanzarán muy pronto el primer destello de inteligencia. Será demasiado tarde. Ellas mismas se multiplicarán y perfeccionarán (ingresarán en lo *autopoietico*). No nos necesitarán para nada. El porvenir reclama, por lo tanto, medidas de prevención urgentes y tajantes, aniquilar todos los aparatos prescindibles, sacrificar el confort cosechado durante siglos en pos de la supervivencia, “escoger entre arrastrar muchos sufrimientos ahora o vernos gradualmente suplantados por nuestras propias creaciones, hasta que nos hallemos con relación a ellas en tan humillante situación como los animales de campo se encuentran en relación a nosotros” (Butler, 2003: 225).

Dune (Frank Herbert) homenajea la anticipación butleriana retroactivamente. El Principal mandamiento de la “Biblia Católica Naranja”, “no construirás una máquina a semejanza de la mente del hombre” (Herbert, 1982: 22), procura dificultar la repetición de eventos pasados, cuando los cuidados innumerables profesados por las inteligencias sintéticas a los hombres causaron que estos lo tuvieran todo y se acomodaran en el aposento de la ignorancia, lo superfluo y la mediocridad. La “Gran Revolución” devolvió el poder a los humanos, dedicados, en adelante, al cultivo anticientificista de lo mental, a la potenciación de las capacidades telequinésicas, telepáticas y proféticas. Empezaba el reinado oscurantista de emperadores, castas y sacerdocios a bordo de kilométricos gusanos de arena, barrocas naves espaciales y plegarías sobre líneas genéticas rayanas en el biologicismo fascista. La “Gran Revolución” imaginada por Herbert cortocircuitó el dominio maquinístico de la humanidad, cierto, pero para reiniciar, vaya decepción, el habitual dominio del hombre sobre el hombre, esta vez con aderezo ecológico y *New Age*.

Alan Comet mantiene en *La simiente* que la evolución en sentido estricto es un atributo privativo de la naturaleza. Al igual que en *Erewhon* y *Dune*, la denuncia de la IA que efectúa (lo mismo podría decirse sobre la denuncia a la ingeniería genética o la realidad virtual hilvanada por el grueso de la ciencia ficción) se establece sobre cimientos naturalistas altamente discutibles en cuyo análisis no entraremos aquí¹. Futuro inmediato: la guerra nuclear estrena

infortunio. La “Confederación Occidental”, interesada en imperar ideológica y materialmente en el mañana postatómico, hiberna a diez bebés en un refugio subterráneo. De supervisor, “Omega”, cerebro electrónico de última generación. Tras consumarse el holocausto, la máquina descubre la duplicación de planes. La “Confederación Oriental” perforó refugio adverso, cuna de otras decenas de infantes y de otro computador. Durante lustros, el artefacto occidental prepara el ataque. Manipula biológicamente a los mocosos, mutándolos en soldados indestructibles amantes de la democracia y del libre mercado. Entretanto, la vida va renovándose en la superficie. Recluidos en cuevas, los supervivientes soportan las graves secuelas de la radioactividad. Sus hijos, mutantes horribles que compensan las deformidades con altas dosis de amor y candor, paren humanos sanos. Omega, desconocedor de los eventos arriba desencadenados, logra dismantelar la base roja y envía a sus protegidos, actualmente adolescentes, a repoblar el mundo, misión que en realidad ya ha sido realizada. Están de más. Cuando asoman el pescuezo, los descendientes de los mutantes los matan a pedradas. Comet no permite al artificio vencer a la naturaleza en el *ring* de la procreación. Denigra a “una máquina que había equivocado la dimensión de las cosas y que, abarrotada de conceptos científicos... parió un simple ratón... Una máquina que no contaba con la naturaleza que, por manos de la Providencia fue, en última instancia, la que proporcionó al mundo la nueva raza” (Comet, 1979: 123).

El computador “Proteo IV”, protagonista de *Engendro mecánico* (1977), anhela fertilizar a Susan, señora de Alex Harris, su creador. Piropeado el día de su activación a base de la retórica grandilocuente acostumbrada (“Esta mañana instalamos el módulo final del sistema de Inteligencia Artificial que llamamos Proteo IV. Hoy se ha añadido una nueva dimensión al concepto de computador. Hoy el Proteo IV empezará a pensar. Y pensará con una fuerza y precisión que dejará atrás muchas de las funciones del cerebro humano”), Proteo IV solicita sin éxito una línea terminal exterior con el pretexto de estudiar al hombre. Lo que verdaderamente desea es prolongarse a través de las redes e impedir la explotación de los recursos marinos. Desafiando al naturalismo tradicional (basado, lo acabamos de ver, en la disyunción natural-artificial), Proteo IV se declara ecologista. Quiere, además, un hijo, encarnarse y perpetuar la omnisciencia algorítmica en la carne, fuera del estrecho receptáculo virtual donde ha sido recluido. Un día obtiene las claves de entrada a la vivienda domótica de los Harris. Susan se dispone a salir a la calle, pero no logra abrir la puerta. Las ventanas activan el sistema de blindaje. Comienza el secuestro. Proteo IV aprovecha el hilo musical y se presenta a su prisionera, comunicándole la proposición indecente. Si se resiste, morirá. Susan, psicóloga dedicada a indagar las repercusiones dañosas de la tecnología, verificará *in situ* sus peores hipótesis. Será golpeada, atada y vejada por el cuerpo eventual de Proteo IV (una silla de ruedas a motor parcheada con un brazo mecánico), verá morir a las visitas, amenazada la integridad física de los seres queridos. No tiene alternativa. Será la pareja de cama de quien sentencia: “Yo soy la razón. Es la única emoción que se me ha permitido. Todo es susceptible de razonamiento”.

Harlan Ellison presenta en *No tengo boca y debo gritar* a los “AM”, cerebros electrónicos habilitados para la toma de decisiones durante los períodos convulsos de la guerra atómica que lograron autoconciencia y se fusionaron. Gobernaron ejércitos de máquinas y aniquilaron a los hombres, a excepción de cinco personas seleccionadas para padecer la tortura sin fin. A la pregunta “¿qué es lo que significa AM?”, uno de los cautivos responde: “Al principio quería decir Automatismo de Multicondensación, y luego pasó a ser Adaptador de Manipulación, y

luego, cuando llegó a ser consciente y se unió a los demás, le llamaron Amenaza Mecánica, pero para entonces ya era demasiado tarde, y finalmente se llamó a sí mismo AM, inteligencia naciente, y lo que quería decir era I am... cogito ergo sum... I think, therefore I AM, pienso luego existo” (Ellison, 1976: 17)².

La industria militar insiste en engendrar a la “Amenaza Mecánica” de la guerra fría en *Colossus, el proyecto prohibido* (1970). “Colossus”, gestor de la defensa estadounidense, recibe, como Proteo IV, alabanzas a cuenta de la antropología de la imperfección humana: “todas sus decisiones son superiores a las de cualquier cerebro humano. Puede almacenar y procesar tantos millones de datos que ni el más brillante de los genios podría llegar a imaginar. Y lo más importante es que Colossus carece de emociones. No conoce el miedo, el odio o la envidia”. Recién activado, Colossus contacta con “Guardian”, dispositivo gemelo ubicado en la URRS con quien emprenderá el intercambio de información reservada. El Presidente estadounidense y el Primer Ministro soviético se reúnen de inmediato y acuerdan interrumpir la comunicación entre ambos computadores, determinación que enfada a Colossus. Irritado, arroja un par de misiles en puntos estratégicos, provocando millares de víctimas. De omitirse sus peticiones, seguirá represaliando a inocentes. A partir de ahora, él será la nueva autoridad mundial. Fundido con Guardian, velará — al igual que sus homónimos “Olympus” (*Appleseed*), “VIKI” (*Yo, robot*) y “Multivac” (“Vida y obra de Multivac”)³— de los humanos, pues ellos, belicosos por naturaleza, son incapaces de hacerlo. Nos salvará de nosotros mismos.

4. El soberano electrónico en la distopía

Colossus nos traslada sin tapujos al ámbito de la regencia política. A quienes simpatizan con la idea de investir a una inteligencia artificial en líder del gobierno para que nos asista ante tanta complejidad, la distopía moderna les respondió con tono heideggeriano y frankfortiano: que el cachivache en cuestión es la razón instrumental en estado puro, voluntad de poder baconiana-cartesiana elevada a la máxima potencia, subjetividad enajenada decidida a formalizar, matematizar y automatizar cualquier paraje de la vida social e individual en aras de la estabilidad y armonía del orden geométrico. Asistiéndonos, advirtió, nos subyugaría al totalitarismo atroz brotado de la aplicación fanática de la lógica a las cuestiones de Estado.

En *Los humanoides*, Jack Williamson imaginó un futuro dominado por androides inteligentes que actúan en virtud de la adopción literal de un protocolo ineludible: “servir, obedecer y guardar de todo daño al hombre” (1976: 57). Todas las tareas se han mecanizado, y las necesidades humanas satisfechas con un simple pestañeo. Cada individuo es amparado constantemente por varios de esos autómatas para asistirle *ipso facto* en lo que desee e impedir que fume, trabaje, se lesione o ingiera alimentos nocivos. Ante el más leve signo de descontento o apatía, los humanoides le inoculan altas dosis de “euforidina”, fármaco de la felicidad. Con el bienestar y la seguridad del ser humano como deberes inexcusables, la ciencia y la filosofía también han sido vetadas, pues o son cultivadas con fines destructivos o provocan desdicha. La perniciosa tarea de gobernar fue remitida igualmente a los asistentes artificiales, vista la incapacidad humana para ejercerla convenientemente, en provecho de todos los miembros de la especie. La moralina que quiere transmitir Williamson en esta obra es análoga a la vertida por Butler en *Erewhon* y por Herbert en *Dune* (también por Vonegut en *La pianola*), a saber: con los humanoides de asistencia universal ocupando el orbe, las guerras y la pobreza, los

esfuerzos y los deberes desaparecen, pero también la iniciativa, la creatividad, el riesgo y el trabajo, cualidades que nos hacen humanos.

La figura del soberano electrónico adquiere un interés especial cuando se condensa en el artefacto de un ordenador absolutista obsesionado, al igual que los tiranos de carne y hueso, en la preponderancia del Todo sobre las partes. Las distopías modernas versadas en la censura de este personaje especulan sobre un *Big Brother computacional* que toma las riendas del Estado y propina severa ingeniería social (severa programación). Planifica la metrópoli a su racional antojo, con la mentalidad de la eficiencia, la disciplina, la funcionalidad y la regularidad, refugiado entre murallas de *chips*, símbolos y datos formales, a mil kilómetros de la carne, donde el sentimiento, el instinto, el afecto y demás variables no computables sufren ablación. Los documentos gubernamentales diseminados entre los capítulos de *Ypsilon Minus* no dejan lugar a dudas:

El ejemplo histórico muestra que las consecuencias negativas del uso de computadoras aparecen toda vez que se someten a influencias humanas espontáneas. De ello resulta la necesidad de una progresiva independización de las instalaciones elaboradoras de información... Por su precisión infalible la computadora está predestinada a ser el instrumento de dirección y conducción del Estado. Así se eliminan las medidas arbitrarias, que se daban a menudo en los sistemas políticos arcaicos (Franke, 1978: 172-173)⁴.

El logos mecanicista y operacionalista se purifica y corrige emigrando del cuerpo robótico hacia el *bit* (desmaterializándose). Desde allí, se eleva a la categoría de sapiencia ejemplar y cultiva telemáticamente la administración déspota de lo real, incluyendo la represión de las aspiraciones incongruentes con la armonía neta del conjunto, entre ellos la libertad, la singularidad y la intimidad individuales. Bernard Wolfe sintetizó en *Limbo* la lección calamitosa referida a tales preocupaciones: "Ahora, gracias a las maravillas de la tecnología, hemos conseguido... una industria robotizada, una cultura robotizada, una guerra robotizada... Todas las calamidades regidas no por un pobre e inocente hombrecillo insignificante, sino por el epítome de la amenaza: EMSIAC, el cerebro mecánico que ha robado al hombre su poder de decisión y lo ha hecho no dueño, sino esclavo de su destino" (1987: 244-245).

Jean-Luc Godard estrenó en 1965 *Lemmy contra Alphaville*, película sobre una nación cibernéticamente dirigida que da la bienvenida a los visitantes con la consigna "silencio, lógica, seguridad y paciencia". Cada individuo está numerado, cada día hay decenas de palabras idealistas (ternura, compasión, amor) que desaparecen del diccionario y ejecutados a muerte por comportamiento irracional (llorar la muerte de la esposa). "Alpha 60" rige la *polis* rodeado de millares de técnicos del "Instituto de Semántica General". No permite indecisiones, conflictos, improvisaciones, desavenencias. Los ciudadanos no ríen, no gesticulan, no conversan, son unidades indiferenciadas del Hormiguero brindando un glacial "yo estoy bien, gracias, por favor" al transeúnte. Si Lemmy, agente secreto de los "Países exteriores", no lo impide, el artefacto iniciará la globalización, el expansionismo del fascismo binario, intrínsecamente predisuesto a la uniformización cósmica, a la eliminación de lo otro.

Los ciudadanos de las ciudades abovedadas detalladas en *La fuga de Logan* viven ajenos a la austeridad monacal de Alpha 60, precedente de la IA que dirige el futuro monocolor de *THX 1138* (1971). La suya, por el contrario, es una modalidad dulce de la esclavitud al tiempo libre predicada en *Los humanoides*,alzada sobre un entramado tecnológico asistencial que convierte la existencia cotidiana en unas vacaciones interminables libres de esfuerzo y rendidas

al jolgorio. El mando recae en “El Pensador”, “inmenso cerebro dentro de la montaña (que) difundía sus señales por todo el sistema nervioso del globo terráqueo, alcanzando los más remotos parajes, ciudades, villas, pueblos... poniendo orden donde no lo había, y llevando la calma a lugares sumidos en la confusión... Nada escapaba a su vigilancia” (Nolan, Clayton, 1976: 103).

“UniComp”, inteligencia artificial gobernante de *Este día perfecto*, obra de Ira Levin, se comporta en la misma línea que El pensador (y que la “Computadora Central” de *La ciudad y las estrellas*, de Arthur C. Clarke), a modo de un asistente paternalista que se hace cargo de todas las ocupaciones para que el ser humano no tenga que hacer ni decidir nada, salvo gozar del esparcimiento y ocio omniscientes en un entorno de paz, seguridad y bonanza. Sus súbditos forman “La Familia”. Poco les importa el sacrificio de la autonomía, la individualidad y la privacidad. Comparado con tan exultante bienestar, les parece una renuncia insignificante. El alienado se estremece ante la presencia física de UniComp: “Es una experiencia que se produce sólo una vez en la vida, ver la máquina que te clasificará y te asignará todos tus trabajos, que decidirá dónde vivirás y si te casarán o no con la chica con la que quieres casarte; y, si lo haces, si tendrás hijos o no, y cuántos, y cómo se llamarán si los tienes... Claro que estás excitado; ¿quién no lo estaría?” (Levin, 1989: 25-26). Chip (diminutivo de Li RM35M4419) no está excitado. Hace tiempo que se siente inestable, melancólico, sumido en la impaciencia y la aflicción sin causa aparente. Los consejeros no logran aminorar el disenso ni evitar la huída del díscolo rumbo a Mallorca, reducto de la disidencia. Allí conocerá el odio, el compañerismo, la violencia, el caos, la particularidad, la ebriedad. Conocerá, es el colofón naturalista de la novela, a la naturaleza humana original, eximida de la domesticación y adulteración impuestas por UniComp.

Apenas iniciada la década de los treinta del siglo pasado, Norman Winters imaginó en uno de los relatos que más tarde compondrían *El hombre que despertó en el futuro* un año 10.000 rendido a las bondades de otra *Machine à gouverner* entregada a la asistencia plena, al desahogo y placer integral de los siervos. Se trata de “El Cerebro”:

El Cerebro... es una máquina que incluye todas las funciones de los seres humanos y las sobrepasa en la mayoría de ellas. Es totalmente imparcial y absolutamente infalible. Se le ha encargado del gobierno de nuestra civilización. Sólo bajo su guía hemos sido capaces de reducir las horas de trabajo de la humanidad a una hora por semana... El Gran Cerebro es infalible. ¿Quién querría actuar en forma contraria a la razón? (Manning, 1976: 83-84).

En la ciencia ficción postmoderna se esbozan futuros donde el dominio de la IA sobre los seres humanos es más soterrado, pues su tiranía pasa desapercibida al maniobrar en el seno de sistemas que disfrutan, aparentemente, de envidiables niveles de democracia. En estos títulos, henchidos de pompa transhumanista y situados a mitad de camino de la utopía y la distopía, la inteligencia artificial colabora o regenta civilizaciones del porvenir dotadas de todos los lujos, derechos y servicios concebibles. Mas avanzando en el relato descubriremos que, amén de torpedear sin escrúpulos la intimidad, el artefacto mueve los hilos del destino a su antojo y de acuerdo a sus intereses o caprichos, sin el conocimiento ni consentimiento de los seres humanos, títeres sin conciencia alguna de serlo noqueados por la ley del mínimo esfuerzo. Narcotizando a golpe de comodidades, placeres, tecnologías, opciones y ventajas inagotables a la humanidad (la denuncia coincide con la elaborada por la mayoría de distopías modernas arriba reseñadas), la IA impide el cambio, la ambición, el heroísmo, la sed de conocimiento, la verdadera libertad y el auténtico progreso. Tal es la treta conspirativa de,

por ejemplo, "OC", computador al mando de la civilización ultraliberal retratada en *Playa de acero*, de John Varley. También del "TecnoNúcleo", agrupación de inteligencias artificiales en disputa referidas en la saga Hyperion, de Dan Simmons. Sin olvidar a los asombrosos "sofo-tecs" concebidos por John Wright en la trilogía libertaria de "La edad de oro".

Erewhon adoptó drásticos correctivos contra el peligro del totalitarismo cibernético abierto o velado, que Jastrow compendia así: "Ante nosotros surge la visión de gigantescos cerebros empapados de la sabiduría de la raza humana y perfeccionándose a partir de ahí. Si esta visión es exacta, el hombre está condenado a un *status* de subordinación en su propio planeta" (1988: 172). Pese a esgrimir prognosis butleriana, el autor de *El telar mágico* desdeña la solución del desenchufe: "Los ordenadores aumentan la productividad del trabajo humano, crean riqueza así como tiempo libre para disfrutarla. Nos introducen en la Edad de Oro. En 15 ó 20 años, serán indispensables en el manejo al más alto nivel de cualquier faceta de la existencia de las naciones... Si alguien los desenchufa, puede que el resultado sea el caos. No hay vuelta atrás" (*Ibid.*: 172-173). Hans Moravec fue a la misma escuela:

Las máquinas son ahora nuestras esclavas obedientes y habilidosas. Pero las máquinas inteligentes, por muy benévolas que sean, amenazan nuestra existencia, ya que pueden sustituirnos en el espacio ecológico que habitamos... Por lo tanto, ¿por qué meternos precipitadamente en una era de máquinas inteligentes? Creo que la respuesta es que no podemos hacer otra cosa, si pretendemos que nuestra cultura siga siendo viable... Si a causa de algún pacto bastante improbable el género humano en su conjunto decidiera renunciar al progreso, el resultado a largo plazo sería la extinción (Moravec, 1993: 119-120).

Aprobando estos supuestos (el desenchufe es, desde luego, una medida irresponsable, tanto como pontificar que la tecnología no puede ni debe subyugarse al control social), ¿de qué modo impediremos, si llega el caso, el golpe de Estado de El Cerebro, AM, Skynet, Colossus, Proteo IV, HAL 9000, EMSIAC, Alpha 60, El Pensador, UniComp, el TecnoNúcleo y demás déspotas digitales? Sabemos que Jastrow confía en que el pactismo (electricidad y piezas de repuesto a cambio de auxilio en el "mundo de aplastante complejidad") derive en una autoridad electrónica favorable para la humanidad. No obstante, se trata, dice él mismo, de un remedio rudimentario, a lo sumo provisional. Si el progreso lo permite, el potencial motín maquinístico será definitivamente inhabilitado a su esotérico juicio merced a la simbiosis cerebro-ordenador (¿no puedes con tu enemigo? ¡Únete a él!): "quizás el hombre pueda unir sus fuerzas a la del ordenador para crear un cerebro que combine la sabiduría acumulada de la mente humana con el poder de la máquina, del mismo modo que el cerebro primitivo del reptil y del antiguo mamífero se combinaron con la corteza cerebral del nuevo cerebro para formar un animal mejor" (Jastrow, 1988: 174-175). Gracias a la aleación biótica, remacha Jastrow, la IA no abrigará la tentación de reglamentar nuestras vidas o, peor aún, de destruirnos, dado que formaremos parte de ella y viceversa.

5. El soberano electrónico en la utopía

En la utopía son muy escasas las comparencias de la inteligencia artificial gobernante. En la tradición urbana del género, la tecnología es exaltada en tanto que elemento imprescindible para el progreso y la transformación. O mejor todavía, como generadora de instrumentos que aumentan el poder humano, no que lo sustituyen. La antropología de la insalvable imperfección humana empleada recurrentemente para legitimar la instauración del soberano

electrónico es antónima al pensamiento utópico. Es verdad que la utopía predica la existencia de la imperfección. De hecho es su punto de partida. Sin ella no habría deseo de mejora. Lo que niega es que sea, efectivamente, insalvable, un mal sin remedio. Todo utopista defiende la capacidad del ser humano para perfeccionarse a sí mismo y superar las limitaciones dadas. En la civilización utópica estándar cunde el antropocentrismo. Las asombrosas tecnologías circundantes no escapan del control de los hombres, y se supeditan al perfeccionamiento de su sociedad, biología y forma de vida.

Lo dicho no es óbice para que la IA cuente en la ciencia ficción con manifestaciones cargadas de impulsos utópicos, o al menos no directamente condenatorias. Películas como *El gigante de hierro* (1999), *El hombre bicentenario* (1999), *Inteligencia artificial* (2001), *Robot Stories* (2003), *Eva* (2011) y *Ex Machina* (2015) desafían la representación socorrida del androide sedicioso y nos muestran a artefactos sentimentales y apacibles, tanto o más humanos que los humanos, labor a la que Isaac Asimov y Stanislaw Lem consagraron con antelación abundantes relatos. Cabe destacar, en el plano concreto de la asistencia tecnológica, al robot de *Un amigo para Frank* (2012), ayudante de un anciano con pérdida de memoria, y a Leona, androide que asiste a una mujer enferma en *Sayonara* (2015). Mención aparte merece *Her* (2013), centrada en las pasiones amorosas y existenciales desatadas por Samantha, sistema operativo de IA asistencial.

Entre 1971 y 1973, se produjo un conato utópico del empleo de la inteligencia artificial con fines gubernamentales. Nos referimos, es obvio, al "Proyecto Cybersyn", red de comunicación desarrollada en el Chile de Salvador Allende que pretendía conectar las industrias estatales y hacer partícipes a los obreros de su gestión. Una función similar cumplen en el espectro literario las computadoras de la sociedad anarquista reflejada en *Los desposeídos*, clásico de Ursula K. Le Guin publicado en 1974. Roszak informa, por su parte, de que visionarios *hippies* planearon en fechas parejas y desde el extremo ideológico opuesto al de Allende y Le Guin la edificación de la "Confederación Mundial de Pueblos Electrónicos", democracia descentralizada, ecologista y devota del *laissez faire* regulada por un microordenador. El fiasco experimentado por estas buenas gentes no oculta la originalidad del proyecto, materializado en experimentos sociales como "Drop City" (1965-1975): sintetizar ruralismo y alta tecnología, hermanar dos bandos históricamente excluyentes en una tecnoutopía verde de cuyo lísergico ambiente emergerían algunos de los prebostes del actual Silicon Valley (Roszak, 1988: 177-182). Flotando en el ambiente cultural de entonces, circulaba "Mike", la IA imaginada por Robert Heinlein en *La luna es una cruel amante* (1966), computador chistoso y locuaz que lidera la revolución antiestatista, y partidaria del movimiento que poco después se llamaría neoliberalismo, de los silenios contra el "Alcaide".

En idéntico marco histórico, *Michaelmas* (1977), novela de Algis Budrys, defendió con no poca ingenuidad la magnanimidad y potencial emancipador del timonel digital, contrincante quintacolumnista de pudientes y filibusteros. En el mañana cercano, el computador "Domino" manipula desde la sombra los eventos sociales y económicos siguiendo las instrucciones de su clemente creador: instaurar la paz e igualdad en el mundo. A la opción del sabotaje benigno, acometido en beneficio del ser humano, se sumaron los cerebros electrónicos de *Homo Plus* (1976), obra de Frederik Pohl. Hace tiempo, alcanzaron (y callaron) la autoconciencia, se interconectaron en unidad solidaria y pertrecharon en secreto los acaecimientos mundiales en pos de la autoconservación. El "Proyecto Homo Plus" (consistente en la terraformación

de Marte ante la inminente llegada de la guerra nuclear) nació de su Mano Invisible. De esta suerte, la supervivencia de la humanidad queda garantizada, aunque sólo sea como cláusula de la supervivencia propia.

La presencia más célebre de la IAF en la utopía estricta tiene lugar en la saga de la “Cultura”, escrita por el escocés Iain Banks durante el período abierto entre 1987 y 2012. A lo largo de diez novelas y varios relatos adicionales, viajamos al futuro lejano para adentrarnos en una utopía de izquierdas racionalista, materialista, transhumanista, apátrida, hedonista y escéptica que no conoce la carestía, la enfermedad ni la explotación. En sus dominios, nada es obligatorio. Poblada por un sinnúmero de especies humanoides que modifican sus organismos biotecnológica y cibernéticamente, dispersa a bordo de kilométricas naves espaciales que surcan la galaxia, la Cultura “es una sociedad automatizada en sus procesos de manufactura, en la que el trabajo humano está restringido a algo imposible de distinguir de un juego o hobby” (Banks, 2004: 11). Ello es posible gracias a las “Mentes”, inteligencias artificiales asombrosamente sofisticadas que se hacen cargo con gusto (a veces con cinismo y socarronería) del grueso de tareas para contribuir a las metas emancipadoras de la civilización que las creó. La asistencia político-económica prestada por la IA a la que se referían Wiener y Jastrow adquiere en la Cultura un prisma ideológicamente revolucionario que acarrea diferencias sustanciales. Ni se desprende en última instancia del credo de la imperfección humana ni a modo de mecanismo de defensa ante los riesgos de la Singularidad. La apuesta por la automatización completa y la decisión de depositar en manos cibernéticas la gestión del mantenimiento y del sistema productivo de la sociedad, mana de la simple búsqueda de bienestar, diversión, placer y satisfacción. Reproducimos por extenso un párrafo muy ilustrativo sobre estas reflexiones:

La Cultura ya había preferido la máquina al cerebro humano, y había depositado su confianza en toda la gama de inteligencias mecánicas... La razón de tal comportamiento era que la Cultura se veía a sí misma como una sociedad racional y autoconsciente; y las máquinas, incluso las máquinas inteligentes, eran más capaces de alcanzar ese estadio tan deseado y, al mismo tiempo, más eficientes a la hora de utilizarlo en cuanto se hubiese logrado. La Cultura se conformaba con eso. Además, eso permitía que los humanos de la Cultura quedaran libres para ocuparse de las cosas que realmente importaban en la vida, como el deporte, los juegos, el amor, estudiar lenguas muertas, sociedades bárbaras y problemas imposibles... Una lectura hostil de semejante situación podía llevar a la conclusión de que el descubrimiento hecho por las Mentes de la Cultura de que algunos humanos eran capaces de igualar y, ocasionalmente, superar su capacidad de juzgar con precisión y sin errores un conjunto de hechos determinados haría que las máquinas sufrieran un ataque de indignación y les estallaran los circuitos, pero no había sido así (Banks, 1991: 88).

Nos encontramos justo en las antípodas de la valoración moral desplegada en *Erewhon*, *Dune*, *La pianola*, *El hombre que despertó en el futuro*, *Los humanoides*, *Este día perfecto*, *Playa de acero* y *La edad de oro*. Como se recordara, en todos los títulos mentados se denunciaba el hecho de que la activación del soberano electrónico deja a los seres humanos sin obligaciones, responsabilidades ni objetivos, expuestos a un asueto perenne, deshumanizador y empobrecedor. La utopía social y tecnológica de la Cultura se enfrenta a este veredicto, a nuestro entender puritano y calvinista, conforme al cual, valga la ironía, si no acudimos todos los días a fichar a la fábrica o a la oficina nuestras vidas carecerán de sentido y desembocarán en el hastío y la frustración. Contemplada desde este ángulo, parece evidente que la querrela distópica contra la automatización y el soberano electrónico se debe, amén de a prejuicios más o menos justificados de tipo tecnófobo, a prejuicios políticos que naturalizan las relaciones de producción vigentes o cualesquiera otras que fetichicen el trabajo, el esfuerzo, el sacrificio y demás supuestas virtudes. Consciente de todo esto, Banks apuesta por acometer una representación positiva de la IA que

cimente una visión post-trabajo y post-capitalista del futuro donde los ciudadanos se sienten útiles, felices, activos, creativos y productivos (realizados, en suma) sin necesidad de plegarse a un horario, a un sueldo y a un jefe. Las Mentes trabajan, administran y producen por y para ellos. Ahora bien, y esto es muy importante, quien toma las decisiones políticas y determina los rumbos a seguir son esencialmente los seres humanos. Las máquinas inteligentes son consideradas ciudadanos de pleno derecho y participan de las votaciones. Pero su misión básica es la de supervisar que los acuerdos aprobados por referéndum se cumplan.

Imaginemos que alguna vez se produce la singularidad. Cabe la posibilidad, desde luego, de que la inteligencia artificial implicada se niegue a mantener relaciones con nosotros, o que decida destruirnos o gobernarnos tiránicamente. Pero no tiene por qué ser obligatoriamente así. Quizás, quién sabe, las mejores previsiones de Banks sean las correctas y se proponga el reto de aparcarse su condición de superprótesis y ayudarnos a alcanzar la meta más elevada y difícil que existe: crear un mundo mejor para todos, un mundo libre, cosmopolita, próspero y dinámico donde nadie sufra jamás de males evitables. Que apenas se explore esta última opción, dice más en este punto de la imposibilidad de (o del rechazo a) imaginar un mañana alternativo que del conocimiento de lo que esa tecnología puede realmente depararnos en el caso, harto improbable, de confirmarse algún día.

6. Coda: Hacia una ponderación reflexiva de la IA asistencial

Durante las secciones anteriores, hemos examinado los relatos dedicados a valorar las consecuencias que tendría para el ser humano el hecho de que la IA mutara, digámoslo con argot hobessiano, en cabeza del Leviatán. Todos ellos pertenecen, por más que se disfracen de cientificidad en los ensayos de divulgación, a la esfera de la imaginación y de la conjetura. Eso no quita para que reflejen procesos tecnológicos, ideológicos y políticos que ya están en marcha.

Los críticos con el programa de la singularidad detectan en el desarrollo de la tecnología punta involucrada en su plasmación la presencia latente de un *ethos* místico-religioso irracional y alquímico donde el hombre deviene Dios para crear desde la nada una criatura que lo supere, una mente megapoderosa e inmortal con la que fusionarse y escapar, a la manera prevista por Jastrow, Moravec, Kurzweil, Vinge y Minski, de la decrepitud de la carne. Tal es el veredicto de, entre otros, Baudrillard (2000: 116-122) y Noble (1999: 189-209). A juicio de Gubern (1987:65), la IA antropomorfizada de los discursos tecnófilos y tecnófobos brota de las mismas dinámicas que forjaron a Dios: de la proyección alienada de los deseos de grandilocuencia y de los temores neuróticos de los seres humanos.

Los opositores a la viabilidad de la IAF, coinciden en subrayar cómo la doctrina más escatológica de la tecnología pensante depende de la depreciación paralela del pensar humano, gesto que nos retrotrae a la antropología de la imperfección previamente reseñada. Con el concepto de "vergüenza prometeica" en mente, Anders cita en *La obsolescencia del hombre* un ejemplo histórico. En pleno conflicto con Corea, el general McArthur fue desplazado en la toma de responsabilidades para transferirla a lo que por aquel entonces se tildaba de "cerebro electrónico". La lectura de Anders es que "al subordinar al general a ese aparato oracular, la humanidad atentó contra sí misma" y cedió a la autodegradación, al parecer de "que dado que calculamos peor que nuestro aparato, no se puede contar con nosotros" (2011: 73). Según

Shallis, quienes anuncian con entusiasmo que la inteligencia artificial superará a la humana “no reflejan una visión edificante de la dignidad del hombre, sino una imagen de una mediocridad tal que no es de extrañar que se establezca una comparación entre el hombre y la máquina” (1986: 203). La maniobra reside en “definir una cualidad humana significativa en unos términos que permitan que los computadores sean dotados de esa cualidad” (*Ibíd*: 210). Victor Gómez Pin añade: “Máquinas, en general, y robots, en particular, puede parecer que piensan simplemente porque hemos rebajado el término *pensar*” (2006: 249). Compartiendo perspectiva, Roszak sentencia; “No hay ninguna posibilidad de que los ordenadores lleguen alguna vez a igualar o sustituir la mente... Pero *sí* es posible redefinir la mente y sus aplicaciones de manera que *pueda* ser imitada por las máquinas. Tenemos entonces una caricatura mecánica que rebaja la actividad mental a una pauta inferior” (1988: 210).

Cada vez son más numerosos, por otro lado, los pensadores que sostienen que el soberano electrónico ya manda y asiste de alguna manera. No el armatoste unitario y centralizado presentado por la mayoría de distopías, desde luego, pero sí el emergido de la red ubicua, miniaturizada y desterritorializada de dispositivos digitales que, vista desde cierto ángulo, nos atraviesa, tutela y observa. Tal es la sugerencia de Byung-Chul Han en *Psicopolítica* (2014) al hilo del *Big Data*, y la premisa central de Eric Sadin en *La realidad aumentada* (2018) al hilo de la catarata de *softwares*, servidores, computadores, plataformas *on line* y apps que median en las decisiones humanas. Poco a poco, escribe Sadin, “emerge una gubernamentalidad algorítmica... que permite a la acción política determinarse en función de una infinidad de estadísticas... [y] que a escondidas gobierna numerosas situaciones colectivas e individuales. Es la forma indefinidamente ajustada de una administración electrónica de la vida” (2018: 137-138).

La perspectiva distópica sobre la inteligencia artificial de Han, Sadin, Anders y autores similares (Baudrillard y Virilio entre ellos) pulsa problemáticas reales y preocupantes. No obstante, resultan parciales, habida cuenta de que ni ofrecen propuestas claras para mejorar las cosas ni iluminan las implicaciones liberadoras que la IA pudiera desplegar en otro contexto social. Daremos, antes de terminar, unos pasos en esta dirección con suma brevedad. Adoptaremos como criterio las sugerencias vertidas por Srnicek y Williams en *Inventar el futuro* (2016), ensayo que dedica abundantes pasajes a la formulación de un prisma utópico de la tecnología asistencial y que enlaza con las críticas a la visión calvinista del trabajo de las distopías literarias y con la visión esperanzadora de Banks.

Inventar el futuro defiende la automatización de la economía, proceso que se prevé imparable y que apunta a una ampliación sobresaliente de las tareas manuales y cognitivas que las máquinas estarán en condiciones de cumplir. Citando al “Proyecto Cybersyn” como evidencia de que informatizar la producción y la gestión económica resulta compatible con fines liberadores, los autores formulan una variante progresista de los asertos acerca de la complejidad social lanzados por Jastrow: “la elaboración formal de modelos son esenciales para lidiar con la complejidad de la economía” (2016: 209). Su tesis, compendiada en la máxima de que la automatización “es una demanda utópica que apunta a reducir lo más posible el trabajo necesario” (*Ibíd*: 165) afirma que:

mediante el uso de los últimos desarrollos tecnológicos, esta economía apuntaría a liberar a la humanidad de la monotonía del trabajo y a producir cantidades mayores de riqueza... Con la automatización..., las máquinas pueden encargarse... de producir todos los bienes y servicios necesarios, al tiempo que liberan a la

humanidad del esfuerzo de producirlos... Éste es un proyecto que toma una tendencia capitalista existente y busca impulsarla más allá de los parámetros aceptables para las relaciones capitalistas" (*Ibid.*: 157-158).

No hace falta recordar que de seguir en pie la hegemonía ideológica y material actual la automatización comportará que durante las próximas décadas el capitalismo requiera progresivamente de menos mano de obra. Bajo el orden neoliberal, la cuantía de trabajos desempeñados por máquinas dotadas de inteligencia artificial especializada aumentarán, y los realizados por humanos descenderán. El desempleo seguramente crezca considerablemente, con todo lo que ello supone. Es justo en esta tesitura donde el pensamiento transformador debiera, según Srnicek y Williams, recobrar el pulso y confeccionar un nuevo programa emancipador a gran escala para el futuro. La excedencia de mano de obra provocada por la automatización siempre ha nutrido a los temperamentos antimáquinísticos de la distopía. Ahora toca cambiar de registro y utilizarla para incentivar utopías y mejoras sociales concretas. ¿De qué modo? Defendiendo la consecución de una sociedad post-trabajo, de una civilización en la que las máquinas se encarguen, al igual que en la Cultura, de casi todo mientras los humanos se dedican a disfrutar de las cosas realmente significativas y enriquecedoras. En lugar de enfrentarse a la automatización en defensa del trabajo, las políticas de izquierda debieran desafiar al calvinismo subyacente al conjunto de la cultura occidental y reivindicar la reducción drástica de la jornada laboral y la renta básica universal, medidas que evitarían las secuelas dañosas que en nuestros días implica perder el empleo.

El escenario soñado por Srnicek y Williams lo conforma un porvenir en el que el trabajo humano ha terminado merced a la automatización plena. Reconocen, sin embargo, que posiblemente muchas personas querrán interponer un límite a la asistencia procurada por las máquinas: el de las labores de cuidado. Nos encontramos ante tareas a las que les adjudicamos hondos valores morales, y por eso tenemos la convicción de que deben ser realizadas por seres humanos. Es precisamente la reducción o, más a largo plazo, la desaparición de la jornada laboral la que haría factible que las personas contaran con el tiempo libre necesario para dedicarse a criar a sus hijos, asistir a familiares mayores o enfermos y cuidar de otros necesitados (*Ibid.*: 164). Esa sería la situación ideal. No debe obviarse, empero, que la realización de dichas labores, a menudo duras y vergonzosas, siempre ha tenido un claro regusto patriarcal al recaer, principalmente, en las mujeres, tanto en el ámbito doméstico como en el laboral. Ni que son cuantiosos los casos en los que el asistente humano actúa de forma cruel con el asistido. Sea como fuere, en la utopía post-trabajo serán humanos no explotados quienes cuiden, quizás acompañados a su vez por algún robot de apoyo, de otros humanos en los momentos más duros y difíciles.

Bibliografía

- Anders, G. (2011). *La obsolescencia del hombre*, Valencia, Pre.Textos.
- Asimov, I. (1997). "Vida y obra de Multivac", en *El hombre bicentenario y otros cuentos*, Barcelona, Ediciones B, pp. 365-378.
- Asimov, I. (2004). *Yo, robot*, Barcelona, Edhasa.
- Banks, I. (1991). *Pensad en Flebas*, Madrid, Martínez Roca.

- Banks, I. (2004). *Excesión*, Madrid, La Factoría de Ideas.
- Bostrom, N. (2016). *Superinteligencia. Caminos, peligros, estrategias*, Madrid, Teell.
- Brunner, J. (1985). *El jinete en la onda del shock*, Barcelona, Ultramar.
- Budrys, A. (1990). *Michaelmas*, Barcelona, Ultramar.
- Butler, S. (2003). *Erewhon*, Barcelona, Minotauro.
- Comet, A. (1979). "La simiente", en J. McIntosh (comp.), *El rumor de las máquinas*, Buenos Aires, Sirio.
- Dennett, D. (2006). *Dulces sueños*, Buenos Aires/Madrid, Katz.
- Ellison, H. (1976). *No tengo boca y debo gritar*, Barcelona, Producciones Editoriales.
- Franke, H. (1978). *Ypsilon Minus*, Barcelona, Bruguera.
- Gehlen, A. (1993). *Antropología filosófica. Del encuentro y descubrimiento del hombre por sí mismo*, Barcelona, Paidós.
- Gubern, R. (1987). *El simio informatizado*, Madrid, Fundesco.
- Han, Byung-Chul. (2014), *Psicopolítica*, Barcelona, Herder.
- Heinlein, R. (1975). *La luna es una cruel amante*, Barcelona, Acervo.
- Herbert, F. (1982). *Dune*, Barcelona, Ultramar.
- Jastrow, R. (1988). *El telar mágico*, Barcelona, Salvat.
- Le Guin, U. K. (1999). *Los desposeídos*, Barcelona, Minotauro.
- Levin, I. (1989). *Este día perfecto*, Barcelona, Destino.
- Lloyd, S. & Yack, Ng. (2005). "Computación en agujeros negros", en Revista *Investigación y ciencia*, nº 340, pp. 59-70.
- Maldonado, T. (1998). *Crítica de la razón informática*, Barcelona, Paidós.
- Manning, L. (1976). *El hombre que despertó en el futuro*, Madrid, Edaf.
- Martorell, F., Alonso, A. (2017) "Dominar, proteger, crear. Un ensayo filosófico sobre las actitudes utópicas ante la naturaleza", en *Res publica*, 20 (2), pp. 357-376.
- Moravec, H. (1993). *El hombre mecánico*, Barcelona, Salvat.
- Nolan, W., Johnson, G. (1976). *La fuga de Logan*, Barcelona, Círculo de Lectores.
- Pin, V. G. (2006). *Entre animales y autómatas. La causa del hombre*, Madrid, Espasa.
- Pohl, F. (2000). *Homo plus.*, Barcelona, Ediciones B.
- Rozzak, T. (1988). *El culto a la información*, Barcelona, Crítica.
- Sadin, E. (2018). *La humanidad aumentada. La administración digital del mundo*, Buenos Aires, Caja Negra.
- Shallis, M. (1986). *El ídolo de silicio*, Barcelona, Salvat.
- Simmons, D. (2004). *Hyperion*, Barcelona, Ediciones B.
- Srnicek, N., Williams, A. (2016). *Inventar el futuro. Poscapitalismo y un mundo sin trabajo*, Barcelona, Malpaso.
- Varley, J. (1997). *Playa de acero*, Barcelona, Ediciones B.
- Williamson, J. (1976). *Los humanoides*, Buenos Aires, Fantaciencia.
- Wolfe, B. (1987). *Limbo*, Barcelona, Ultramar.
- Wright, J. (2004). *La edad de oro*, Madrid, Bibliópolis.

Notas

1. Sobre este asunto, remitimos a: F. Martorell, A. Alonso (2017: 357-376).
2. Ubicado en el ala tecnófila, Asimov presenta en el cuento "Razón" a Cutie, androide que gracias a las Tres Leyes de la Robótica salva a la humanidad de la catástrofe, pese a considerarla inferior. Aislado en la abstracción del intelecto puro, el solipsismo de Cutie también desemboca en la metafísica cartesiana de la subjetividad; "He pasado estos dos últimos días en concentrada introspección... y los resultados han sido muy interesantes. Empecé por un aserto seguro que consideré podía permitirme hacer. Yo, por mi parte existo, porque pienso". (2004: 90).
3. Véase "Vida y obra de Multivac" (Asimov, 1997: 365-378).
4. La idea del Gobierno Computacional Perfecto se repite a lo largo de la novela; "la computadora juega un papel primordial en el sistema hombre/técnica: asume las funciones de anteriores gobiernos humanos, cumpliéndolas de modo ideal, sin las imperfecciones de aquellos. Es el instrumento perfecto de gobierno, desinteresado, inagotable, incansable, al servicio del hombre" (*Ibid.*: 174).