

El sujeto desde la neurociencia y la inteligencia artificial

¿Qué concepción del ser humano subyace a las ciencias cognitivas? ¿Dónde están los límites de nuestra noción de humanidad? ¿Qué retos plantea la neurociencia y la inteligencia artificial al sujeto contemporáneo? Estas son algunas de las preguntas que trataremos de contestar a lo largo del presente artículo, realizando para ello un recorrido desde las teorías de la mente y el sujeto hasta sus consecuencias en el desarrollo de las ciencias cognitivas contemporáneas.

Palabras clave: ciencias cognitivas, neurociencia, inteligencia artificial, sujeto.

0. Introducción: el sujeto en las ciencias cognitivas

Para comenzar, nos gustaría señalar que el título de este artículo adolece de cierta imprecisión. Ya que, si bien la neurociencia y la inteligencia artificial se habían considerado como campos independientes en un principio, hoy en día esa delimitación parece obrada de modo artificial y no se corresponde con la realidad de la investigación en estos campos. Por ello se ha unificado bajo el nombre de ciencias cognitivas al conjunto de disciplinas que estudian el cerebro y la cognición a la vez que producen tanto teorías como prácticas científico-tecnológicas relacionadas con los mismos. Estas ciencias cognitivas están orientadas al estudio del cerebro, el conocimiento y las relaciones mente-cerebro de una forma transdisciplinar. La filosofía cobra un importante papel en estas investigaciones aportando preguntas relevantes que tratan de ser contestadas o reformuladas desde el resto de disciplinas, así como en la definición del marco teórico y la evaluación de los resultados de las diferentes investigaciones.

Tal y como sostiene P. Martínez Freire en “La importancia del conocimiento” (2007) «las ciencias cognitivas, constituidas de manera institucional en 1977 y con un gran empuje en investigación y docencia desde entonces, no sólo conforman un campo interdisciplinar substancial sino que además son un enlace entre campos científicos diversos, aunando disciplinas formales (lógica y matemáticas), disciplinas físicas (informática y biología) y disciplinas humanas (psicología y lingüística)». Las ciencias cognitivas se entienden entonces no sólo como la suma de las distintas disciplinas involucradas en el estudio de la cognición, sino como una disciplina totalmente nueva que emerge de la investigación conjunta y transversal.

De este modo, podemos afirmar que las ciencias cognitivas constituyen en sí mismas un campo específico de conocimiento con una metodología híbrida que incluye las que tradicionalmente se asocian a las ciencias sociales, la ingeniería, la neurociencia o la filosofía entre otras muchas disciplinas. La necesidad de las relaciones interdisciplinares y el estudio transversal de la cognición puede entenderse especialmente si partimos del carácter inabarcable del estudio de la misma, el cual puede aplicarse tanto a seres humanos como a animales o máquinas. Así, el estudio de la

(1) Allen Newell (1927-1992) y Herbert Simon (1916-2001), comparan en su obra *Human Problem Solving* (1972), a los seres humanos con los computadores declarando que un computador programado y el sistema de solución de problemas humano pertenecen ambos al género sistema de procesamiento de información. Desde entonces se han dado muchos casos que ratifican esta teoría y que iremos mostrando a lo largo del artículo.

(2) Actualmente es ampliamente aceptado que nuestros parientes más cercanos, los primates, tienen inteligencia (aunque en menor grado que los humanos). Aunque también otros animales han mostrado comportamientos inteligentes, entre ellos perros como el collie de la pradera Rico, estudiado por Juliane Kaminski en el Instituto Max Planck de Antropología Evolutiva que cuando llegó para ser estudiado, sabía los nombres de alrededor de 200 juguetes, y aprendía con facilidad los de nuevos juguetes, podía aprender y recordar palabras con tanta rapidez como un niño que empieza a andar. Otro collie, Betsy, tiene ya un vocabulario de más de 300 palabras. También podemos encontrar entre los animales que muestran inteligencia aves, como Alex, el loro gris africano estudiado por Pepperberg, el cual es capaz de distinguir colores, formas, tamaños y materiales; y de resolver operaciones aritméticas simples, como contar los cubos de juguetes amarillos entre una pila de cubos de varios colores. Estudios con el cuervo de Nueva Caledonia, que es una de las aves más hábiles (hacen y usan herramientas; forman sondas y ganchos con ramitas y tallos de hojas para hurgar en las copas de las palmeras, en donde se ocultan larvas gordas, etc.), han demostrado que también podemos encontrar inteligencia e incluso creatividad en las aves. En el vídeo que puede verse aquí <http://users.ox.ac.uk/~kgroup/tools/introduction.shtml> se observa a Betty resolviendo la prueba que está ante ella: un tubo de vidrio con una cesta pequeña, con un trozo de carne, alojada en su centro. Previamente, los investigadores habían colocado dos fragmentos de alambre en el cuarto: uno recto y otro doblado en forma de gancho. Pero, sin que se diesen cuenta, otro cuervo robó el gancho antes de que Betty pudiera encontrarlo de forma que ésta

cognición no puede verse limitado al estudio del cerebro, de la psicología o de la inteligencia artificial de forma inconexa, sino que debe darse de una forma integrada. Las últimas teorías de la mente extensa y la cognición distribuida (que veremos con más detenimiento a lo largo del presente artículo) amplían aún más el rango de esta afirmación, ya que de ellas se desprende que los procesos cognitivos de la mente humana integran tanto los dispositivos que empleamos como el entorno y otros agentes del mismo. Esto nos lleva a plantearnos la cuestión de la mente y los procesos cognitivos no sólo como algo exclusivamente humano (se ha demostrado hace décadas que máquinas (1) y animales (2) pueden pensar, incluso de modo conjunto (3)).

Para ilustrar este conflicto sobre qué es o dónde se encuentra eso que podríamos considerar la esencia de lo humano, narraremos a continuación una historia de ciencia ficción:

Paula y Gloria son amigas. Llevan escribiendo juntas desde hace muchos años (cuando estudiaban juntas la carrera). En ocasiones, cuando no pueden trabajar juntas presencialmente lo hacen de manera telemática (empleando herramientas como videollamadas, documentos de texto compartidos en red, etc.). Hace cinco años Paula se mudó para continuar sus estudios realizando un doctorado en el MIT, desde entonces han estado trabajando en un libro, en este caso de un modo totalmente telemático. Durante estos años, a su vez, Paula ha ido implementando diversas prótesis en su cuerpo, que van desde una mano biónica a pequeños nanobots que circulan por el torrente sanguíneo de su cerebro regulando algunos de sus procesos neuronales. Los cambios han sido progresivos, hasta llegar a la fecha actual, en la que Gloria ha ido a visitarla para hacer una presentación conjunta del libro y ha descubierto que Paula es lo que se conoce como un cyborg. ¿En qué sentido puede pensar Gloria que Paula sigue siendo la misma persona de la que se despidió cinco años atrás en el aeropuerto de Singapur?, ¿podemos decir que es la misma Paula la que comenzó escribiendo el libro que la que ahora lo presenta?, ¿es para Gloria su amiga Paula la misma persona con la que tuvo una conversación telefónica esa misma mañana? ¿en qué momento puede establecerse un corte, un cambio, en la concepción de Paula como ser humano (si es que ese corte puede darse)?

Existen dos nociones cruciales relacionadas con el concepto de identidad en las que merece la pena detenerse un momento antes de proseguir con la exposición de este artículo: la identidad numérica y la identidad cualitativa. Establecemos una identidad cualitativa entre dos objetos cuando ambos comparten las mismas propiedades (color, tamaño, forma, funciones, etc.), sin embargo, dos objetos pueden ser cualitativamente idénticos y, sin embargo, pueden ser numéricamente distintos (este artículo que lees y que tienes en tus manos y el que está leyendo otra persona que lo ha adquirido en la otra punta de la ciudad son cualitativamente idénticos pero numéricamente diferentes). La identidad numérica de los estados mentales (deseos, intenciones, sensibilidad...) no es independiente de la identidad de sus sujetos. Así, las intenciones de Paula a la hora de escribir el libro si fueran volcadas en una conciencia cibernética serían cualitativamente idénticas en ambos casos, pero es sólo en la vinculación de la intención a una parte del mundo (el cuerpo de Paula o su conciencia cibernética alojada en la red) donde la noción de sujeto cobra relevancia. Es decir, en la definición y delimitación del sujeto de conocimiento, el que piensa, razona, procesa información y resuelve problemas (el sujeto que estudian las ciencias cognitivas –ya esté basado en carbono o silicio) deviene imprescindible el anclaje de los estados o los procesos cognoscitivos a algo físico que los contenga. Así tiene sentido afirmar que existen estados mentales cualitativamente idénticos y numéricamente distintos. Paula y su conciencia cibernética compartirían todo excepto el soporte. ¿Dónde establecemos aquí el límite de lo humano?

sólo tenía a su disposición el alambre recto para resolver el problema. Al ver que no le sirve, al menos no tal como está, lo toma con el pico, empujando un extremo hacia una esquina en el suelo para doblar el otro extremo y formar un gancho. Así, armada, alza el recipiente y lo saca del tubo. Lo más interesante es que, según dijo Kacelnin "Esta fue la primera vez que Betty vio un pedazo de alambre como este. Aun así, sabía que podía usarlo para formar un gancho y exactamente dónde necesitaba doblarlo a fin de que el gancho cupiera en el tubo para tomar la carne. Tras ésta, Betty realizó distintas pruebas que requerían diferentes soluciones, resolviéndolas todas. «Eso significa que tuvo una representación mental de lo que quería hacer. Eso, dijo Kacelnin, es un indicador de un tipo importante de sofisticación cognitiva». Cfr. Cabañes 2008.

(3)

Existe incluso una curiosa hibridación de creatividad animal y Creatividad Artificial, fruto de la fusión de las ideas del artista Guy Ben-Ary, de Symbiotica Research Group de la Universidad de Australia y del neurólogo Steve Potter de la universidad de Atlanta; un robot que siguiendo los impulsos de un cultivo de neuronas de rata situado a miles de kilómetros, al que accede a través de Internet, es capaz de generar obras de arte. Este proceso se completa cuando, también a través del ciberespacio, el robot transfiere la información de las realizaciones artísticas al cultivo de neuronas, originando nuevas instrucciones creativas. Este circuito de ida y vuelta de impulsos eléctricos que generan creatividad simula la base neurológica de la inteligencia y la conciencia. Por primera vez una máquina es capaz de inspirarse en fuentes cerebrales no humanas, de realizar creaciones espontáneas y de adaptar la obra de arte a nuevas instrucciones. Este robot presentado en 2003 constituye el primer intento de reflejar la creatividad animal en una obra de arte a través de un robot. (Cfr. Cabañes 2008 y Cabañes 2013)

(4)

La anatomía y las funciones del cerebro en Aristóteles pueden encontrarse en el capítulo 7 del libro II de *Partibus Animalium*.

Esta es la pregunta que articulará el discurso del presente artículo en el que llevaremos a cabo un repaso de los aportes de las ciencias cognitivas a la noción de sujeto y en el que trataremos cuán difusas han devenido las fronteras entre lo humano, la máquina y el animal.

1. Teorías de la mente y teorías del sujeto

La concepción del sujeto en las ciencias cognitivas, como veíamos en la introducción, se expresa a través de la metáfora de la computación. El sujeto es entendido así como un procesador de información. Sin embargo, esta concepción del ser humano es relativamente reciente. Son muchos los cambios que se han dado en la visión que los seres humanos tienen de sí mismos y del rasgo que consideran más distintivos de su especie: la racionalidad. Cada uno de estos cambios ha traído consigo una concepción antropológica diferente inserta en un cambio de paradigma general, que afecta tanto a las ciencias naturales como a las disciplinas humanísticas o los estudios teóricos tales como la filosofía o las matemáticas. Esto es así de tal modo que resulta inseparable el cambio de paradigma del cambio de concepción del ser humano, ya que en todos estos momentos de la historia ambas cuestiones se han dado de forma simultánea. Sería dudosamente fructífero tratar de resolver el dilema que presenta el trazar la genealogía de estas transformaciones de forma cronológica, más bien nos parece que lo verdaderamente interesante reside en encontrar el sentido de esta simultaneidad: el mundo cambia con los seres humanos y los seres humanos cambian con el mundo.

A continuación haremos una breve reseña histórica de las teorías de la mente y el sujeto, algo que consideramos muy relevante en tanto que el avance en lo que hoy día podemos denominar ciencias cognitivas y el enfoque de sus prácticas tecnocientíficas están estrechamente relacionadas con la concepción que se tenga del sujeto. No daremos cuenta aquí de todas las teorías existentes, nos limitaremos a algunas de las más relevantes con la intención de ilustrar los vínculos que se establecen entre las teorías y las prácticas y su retroalimentación.

Aristóteles es conocido en gran medida por defender que el ser humano se distingue del resto de animales por su capacidad para razonar. El alma intelectual aristotélica, la propia y exclusiva de los humanos, les dota de la capacidad de acceder a la felicidad perfecta, aquella que les permite desarrollar su propia naturaleza. Esta forma de concebir al ser humano ha llegado hasta nuestros tiempos impregnando gran parte de los discursos científicos. Desde la época clásica, el concebir a los seres humanos como poseedores en exclusiva de la capacidad de discernir es lo que ha propiciado una división tajante y jerárquica de la naturaleza donde son precisamente éstos, los humanos, quienes no sólo podían, sino que además debían o tenían la potestad para gobernar sobre el resto de seres. La concepción del alma intelectual de Aristóteles como un alma anclada al cuerpo que se genera con él y desaparece también con él le llevó a explorar la anatomía humana buscando indicios de la residencia de la misma. Halló que el alma estaba relacionada con la sangre y que el corazón, en tanto que motor de ésta, era el lugar en que se hallaba. Si bien en Aristóteles encontramos muy pocas referencias a la cabeza o el cerebro en relación al entendimiento (le otorga una función secundaria como refrigerador del organismo) (4), su maestro Platón situaba el alma racional en ella no por una cuestión biológica, sino jerárquica: si el alma racional gobernaba o debía gobernar sobre todo lo demás lógicamente debía situarse en la parte superior del cuerpo. Aun cuando alguien leyendo estas líneas podría considerar que Platón estaba más acertado que Aristóteles, fueron las teorías y las prácticas de este último las que han contribuido a los estudios anatómico-forenses del cuerpo humano y con ello al avance de la medicina.

En la Modernidad los pensadores más representativos de la época continuaban manteniendo esta visión del sujeto como un ser racional. Descartes con su distinción *res cogitans* - *res extensa* (lo que hoy en día se conoce como la distinción cuerpo-mente) continúa depositando la especificidad del ser humano en su capacidad racional (*la cosa pensante* cartesiana se convertirá en el sujeto moderno cuya estela llegará hasta la contemporaneidad). Descartes establecía una metáfora del cuerpo como máquina (metáfora ésta sobre la que se han asentado los cimientos de la medicina moderna). Como hombre de su época, Descartes abrazaba el mecanicismo y consideraba que el cuerpo humano funcionaba como un mecanismo al que se le aplicaban las leyes de la causalidad física. Asimismo, Descartes situaba el origen del movimiento del cuerpo en la mente (*res cogitans*), la cual se relacionaba con aquél a través de la glándula pineal (5).

David Hume, por su parte, anclado en una posición empirista, terminará por definir al sujeto como un haz de percepciones carente de una unidad fija y llegará a considerar que hay *algo irracional en nuestra racionalidad* (la falta de justificación racional de las inferencias inductivas). Sin embargo, a pesar de que las conclusiones del filósofo escéptico parezcan llevarnos a estimar que esta concepción del ser humano era generalizada, su obra pasó por la historia moderna sin grandes repercusiones entre los pensadores contemporáneos. Esta concepción de la racionalidad del ser humano y la sospecha de la *falta de racionalidad de la propia razón* no encontraría interlocutores que la tomaran en serio hasta el Romanticismo.

La creciente secularización del saber a la que asistimos durante la Modernidad no sólo lleva a avances teóricos acerca de la cuestión del sujeto, sino que influye significativamente en las prácticas científicas. El cuerpo deja de entenderse como un receptáculo sagrado del alma, inmodificable y hermético, para dejar paso a una nueva forma de estudio y relación con el mismo que dio lugar a importantes avances, especialmente en anatomía (siguiendo la estela que habían dejado los primeros estudiosos de la medicina en la época clásica). En el siglo XIX Ramón y Cajal marca un antes y un después en el estudio sobre los cuerpos con el descubrimiento, digno de un premio Nobel, de la neurona.

Una de las teorías de la mente y del sujeto que más se ha extendido y aplicado en el siglo pasado es el conductismo, que ha sido definido por J.R. Kantor como «una renuncia a las doctrinas del alma, la mente y la consciencia» en favor del «estudio de los organismos en interacción con sus ambientes» (Kantor 1968, cit. por Campos, 1973, p. 573). En este sentido, lo que hasta este momento histórico se consideraban teorías o perspectivas filosóficas comienza a defenderse como una ciencia por quienes la practican. En especial Kantor (1973, cit. por Campos, p. 91) afirma que la psicología estudia las interacciones entre los organismos y su entorno al igual que la astronomía estudia el comportamiento de astros y galaxias o la física el de la materia y sus propiedades. De este modo el conductismo da inicio a una metodología de estudio de los fenómenos psicológicos desde la perspectiva de una ciencia de la conducta, cuyas implicaciones en la percepción del sujeto termina con el mentalismo (atribuciones dualistas extramateriales como el alma o la mente) e introduce el ambiente como un elemento imprescindible que modela la conducta.

Esta nueva concepción conlleva, entre otras muchas cosas, que ya no es suficiente experimentar con cuerpos muertos, ya que la cognición implica un proceso dinámico, propio de seres vivos no inertes. Pensar, sentir, comprender... son nociones que implican movimiento, cambio, y una relación intencional cuya investigación requiere de cuerpos vivos. De este modo da comienzo la experimentación con seres humanos. Los avances de la ciencia a través de la experimentación con cuerpos vivos, fueron enormes, pero también se generaron los casos más polémicos: Lobotomías, electroshock, experimentos

(5)

En la concepción de la informática y la robótica contemporáneas esta distinción mente/cuerpo se ha mantenido en la diferenciación entre software y hardware, en la que la glándula pineal podría equipararse al *kernel*, el software fundamental de un sistema operativo encargado de facilitar a los distintos programas acceso seguro al hardware.

conductuales, son los inicios del desarrollo de nuevas técnicas de imagen en vivo del cerebro, que permiten imágenes no estáticas del mismo y, por tanto, dan la posibilidad de observar el cerebro “funcionando” de un modo mucho más inocuo para el sujeto de estudio que sus predecesoras.

Como hemos adelantado, la teoría del sujeto como sistema de procesamiento de información, propia de las ciencias cognitivas, ha llegado al punto de considerar que las máquinas piensan, lo que nos conduce a una investigación sobre los procesos cognitivos que puede ser emulada y estudiada en sistemas artificiales (lo que desplaza, en cierta medida, el sujeto de estudio del humano a la máquina). Así aparece una nueva perspectiva que contribuye a los estudios transdisciplinares y a un avance exponencial (en tanto que incluye tecnologías que evolucionan exponencialmente más rápido que los seres orgánicos), a la vez que implica la emergencia de nuevas y complejas teorías del sujeto que nos harán replantearnos seriamente nuestra concepción de lo vivo que tal vez pueda dejar de estar anclada a lo orgánico. Este contexto en el que las fronteras comienzan a desdibujarse, representa un cambio de paradigma radical en la propia concepción del sujeto que se asemeja más a una ficción operada por el lenguaje que a una descripción ontológica.

2. Hay cyborgs entre nosotros

*La ubicuidad y la invisibilidad de los cyborgs son la causa de que estas máquinas sean tan mortíferas. Políticamente son tan difíciles de ver como materialmente. Están relacionadas con la conciencia -o con su simulación.
(Haraway, 1995, p. 261)*

La nueva noción de sujeto emergente de las ciencias cognitivas se encuentra íntimamente relacionada con las aplicaciones prácticas y los experimentos computacionales y neurobiológicos. En la búsqueda de una respuesta significativa acerca de qué es pensar está involucrada tanto una necesidad práctica (de simulación por computador, de estudio del cerebro en funcionamiento, etc.) como una necesidad teórica que establezca los modelos que después resultarán operativos. Ambas partes son indisolubles, avanzan a la vez y el resultado de las mismas no puede entenderse de forma descontextualizada: lo que emerge del proceso de investigación es la propia disciplina científica, y sus resultados se identifican con ella. La inteligencia artificial, si es que puede tomarse de forma aislada, constituye un claro ejemplo de cómo funcionan las ciencias cognitivas. Veremos que es del todo imposible hablar de avances en este campo que no impliquen descubrimientos en neurociencia, psicología o que no nos lleven a replantearnos los propios límites de la singularidad de los seres humanos. Si la creencia imperante que justificaba la superioridad humana se basaba en la consideración de que lo racional era exclusivamente humano, encontramos en las nuevas máquinas que habitan y conviven con las personas (y hasta dentro de ellas) un serio cuestionamiento de aquello que entendíamos como humanidad. Si bien este relato roza la ciencia ficción o la técnica ficción, no podemos dar de lado los interrogantes que nos plantea y la urgente necesidad de darles respuesta tanto desde el punto de vista de la ciencia como desde la moral o la política. A continuación expondremos algunos ejemplos relevantes de avances tecnocientíficos relacionados con las ciencias cognitivas que manifiestan distintos modelos de concebir la cognición y sus repercusiones para la concepción del sujeto.

- Sistemas expertos

Los sistemas expertos (una de las técnicas menos complejas de la inteligencia artificial) son capaces de emular artificialmente aspectos muy

concretos del razonamiento de un especialista humano, pero siempre en un ámbito restringido y limitado de conocimiento. Este tipo de emulación de los procesos cognitivos parte, genera y reproduce el concepto de humano racional en cuanto acumulador o continente de conocimiento, así como la organización posfordista de lo humano en función de sus habilidades de trabajo: en tanto que continentes de conocimiento, son limitados y deben especializarse en áreas específicas. Se basan en reglas fijas que facilitan una mayor eficiencia; para tomar sus decisiones utilizan un conjunto predefinido de conocimientos previamente dado (la memorística prima sobre el pensamiento crítico). El interés de estos sistemas reside en su capacidad para resolver con facilidad ciertos problemas complejos que sólo puede llevarse a cabo por personas expertas intensamente entrenadas.

- Sistemas de razonamiento basado en casos

Los sistemas de razonamiento basado en casos conforman una evolución algo más compleja de los sistemas expertos. Requieren también de una gran base de conocimientos, en este caso con ejemplos de problemas ya resueltos que empleará junto con la experiencia adquirida en el pasado en la resolución de problemas, de modo que el sistema pueda rastrear en su memoria y recuperar el caso que más se asemeje adaptándolo al problema actual. En este caso podemos hablar de un aprendizaje constante, ya que el sistema no sólo emplea los casos aportados por el programador, sino que se actualiza constantemente incorporando cada nueva solución encontrada a la base de conocimientos, que se actualiza continuamente por este procedimiento. Al incorporar permanentemente nuevos casos a su memoria, el sistema adquiere una experiencia que le permite encontrar soluciones cada vez más creativas y eficientes. Se trata, en definitiva, de un aprendizaje por analogía. Desde esta perspectiva se están tomando las interacciones con el ambiente (los problemas con que se topa, ya sean estos introducidos por un humano o encontrados en el entorno), como una fuente de aprendizaje y experiencia que modificarán su “conducta”, tal como afirmaba el modelo conductista del sujeto.

Los sistemas de razonamiento basado en casos son sistemas conductistas plausibles cognitivamente (en el sentido que modelan aspectos del funcionamiento de la mente a la hora de resolver problemas a partir de la experiencia).

- Lo humano como una particular secuencia de ADN

Tratamos en este apartado dos avances que podrían considerarse como claramente diferenciados, pero que ambos parten de una concepción de lo humano como una cadena de genes resultado de la combinación y mutación de los mismos a lo largo de una evolución. Se trata de los algoritmos genéticos, procedentes de los estudios en inteligencia artificial y de el ADN array, proveniente de la investigación biomédica.

Los primeros son métodos adaptativos de búsqueda basados en los mecanismos de evolución biológica. Como un reflejo de las teorías del sujeto que conciben a éste como un ser determinado por su genética, llevan codificadas todas las posibles soluciones a un problema dado en forma de cadenas de caracteres de longitud fija llamados “genes”. Se genera, habitualmente al azar, una “población” inicial de prueba (un conjunto de posibles soluciones con ligeras variaciones entre ellas), que posteriormente se evaluará según un criterio de desempeño fijado con anterioridad (la denominada “función de adecuación” que se correspondería con el criterio adaptativo de Darwin). En cada ciclo (cada “generación”) se seleccionan las soluciones cuyo valor de adecuación sea mayor, descartándose así el resto de las soluciones. “Las más aptas”, aquellas que han sido seleccionadas, se combinan o reproducen entre sí para producir nuevas soluciones (su “descendencia”),

permitiendo eventualmente introducir alguna modificación aleatoria (una “mutación”) durante el proceso. Este ciclo se repite hasta llegar a aquella considerada aceptable. Consideramos que no es necesario ahondar en la concepción del sujeto que subyace a esta práctica, ya que el modelo queda perfectamente definido en las metáforas que emplea.

Por otro lado el ADN *array*, basado en este mismo modelo, expone en una tabla ordenada todos los genes que se expresan en determinados tejidos según su función en condiciones experimentales determinadas. De este modo se permite generar de modo artificial una imagen transversal de un tejido que cambia en función de las condiciones ambientales y celulares, permitiendo determinar la expresión genética de un tejido en un momento determinado. Esta tecnología ha generado todo un marco de estudio nuevo denominado genómica funcional que estudia la regulación coordinada de la expresión genética conformando una poderosa herramienta para la generación de nuevos fármacos que podrían incluso adaptarse a sujetos particulares en función de su genética. (Doménech -Sánchez 2004).

- *Redes Neuronales Artificiales*

Este tipo de sistemas, claramente basados en los descubrimientos procedentes de las neurociencias, representan el intento de la Inteligencia artificial conexionista de superar los problemas de las técnicas de la inteligencia artificial clásica. Como su propio nombre indica, emulan la estructura del cerebro humano y las redes neuronales biológicas, dando lugar a redes neuronales artificiales a las que no se les imbuje ningún tipo de regla fija (como sí ocurría en los casos anteriores), sino que aprenden por sí mismas a reconocer patrones a través de un proceso de entrenamiento basado en el análisis automático y sistemático de una suficiente cantidad de distintos ejemplos. Constituyen una buena técnica para manipular datos poco precisos o incompletos y hasta compuestos de ejemplos contradictorios (cuestiones que generaban no pocos problemas a otros tipos de técnicas de la inteligencia artificial). Pero quizá lo más relevante aquí sea destacar que no requieren de la programación de ninguna tarea específica, ya que son capaces de realizar generalizaciones y de aprender basándose en la experiencia.

El modelo más antiguo que encontramos entre estas redes de neuronas artificiales es el denominado Perceptrón, que se ha empleado en la aplicación de reconocimiento de patrones. Desde su implementación, era capaz de establecer generalizaciones, aunque no disponía de la capacidad de clasificar clases no separables linealmente. ADALINE, por su parte, es la primera red neuronal que se aplicó a la resolución de un problema real (filtros adaptativos para eliminar ecos en las líneas telefónicas). Más tarde, Teuvo Kohonen generó un modelo que conseguía formar mapas de características de una forma semejante a como sucede en el cerebro humano. Como último ejemplo pondremos las redes multicapa, que aprenden gracias al algoritmo *back-propagation*, que consiste en una capa de entrada, una de salida y una o más ocultas. Emplean un gradiente heurístico que permite que una red se auto-organice de tal manera que consiga mejorar su trabajo constantemente.

Todos estos tipos de redes de neuronas artificiales parten de una idea de sujeto como ser no fijado, cuya conformación dependerá de sus interacciones con el ambiente y no serían posibles sin los estudios en neurobiología, que aportan el modelo y a su vez se nutren de estas investigaciones para continuar su avance.

- *Sistemas multiagentes*

Los sistemas multiagentes dan cuenta de la concepción del sujeto como animal social y de los procesos cognitivos en un estado de interconexión

constante con otros seres. Este tipo de técnica de inteligencia artificial es la más cercana a los modelos de la mente extensa y la cognición distribuida en tanto que incorporan otros agentes en la generación de los procesos cognitivos y en la búsqueda de soluciones conjuntas y coordinadas, de modo que el sistema en conjunto no es meramente una suma de los diferentes agentes que los conforman, sino algo más. Las propiedades de estas comunidades de agentes (entidades relativamente autónomas e inteligentes que cooperan para el desarrollo de tareas) no pueden derivarse únicamente de sus partes constitutivas y su capacidad puede aumentar exponencialmente si se incrementa el número de bases de conocimiento especializadas. Estos sistemas plantean una de las cuestiones más interesantes en cuanto a lo que a las teorías del sujeto se refiere, al desdibujar por primera vez las fronteras entre la mente como algo individual y aislado y la posibilidad de una inteligencia colectiva.

- Humanos en máquinas y máquinas en humanos

La hibridación cuasi total de lo humano y lo tecnológico, que aventura la ontología cyborg, se hace especialmente patente en dos de los últimos avances de la tecnobiociencia. Por un lado podemos encontrar casos en que los propios procesos cognitivos son regulados por nanobots insertos en el torrente sanguíneo del cerebro, por otro, réplicas de órganos humanos, e incluso de humanos completos en chips.

Podemos mencionar a modo de ejemplo el chip que imita las sinapsis cerebrales generado por el equipo de Chi Sang Poon en el MIT, dispositivo que actualmente es capaz de procesar información y nuevos estímulos con sus 400 transistores. Aunque especialmente relevantes son las investigaciones del Instituto Wyss para la Ingeniería Biológicamente Inspirada de la Universidad de Harvard en colaboración con la Agencia de Defensa y Reducción de Amenazas (trataremos más adelante la estrecha relación de este tipo de investigaciones con el ejército). Actualmente se encuentran en proceso de desarrollo de un instrumento automatizado que pueda simular un humano en chips, habiendo conseguido ya un intestino y un pulmón y hallándose en el desarrollo de un pulmón, riñones y médula ósea. En sus propias palabras "Los órganos en un chip representan un nuevo método para modelar la estructura, la biología y el funcionamiento de los órganos humanos (...). El uso de este tipo de microsistemas en el estudio de las patologías, y suponen una esperanza para la identificación de nuevas dianas terapéuticas". (Huh et al. 2012).

Aunque quizá lo que más llame la atención es que la agencia de defensa esté involucrada. Su implicación en la investigación responde a la "necesidad" de probar los efectos de las armas químicas en las personas para planificar estrategias defensivas.

Por otro lado, los nanobots ya están aplicándose en la medicina, aunque en la gran mayoría de los casos estén aún en fase experimental, ya circulan por el torrente sanguíneo de cientos de pacientes, siendo utilizados para tres funciones: diagnóstico (nanobiosensores capaces de detectar en tiempo real sustancias químicas y biológicas), terapia (administración de fármacos controlados, destrucción de células cancerígenas, etc.) y regeneración de tejidos y órganos (por ejemplo, en el caso de enfermos de Parkinson, pueden sustituir las neuronas destruidas por la enfermedad). Incluso, como un caso a destacar, podemos mencionar la investigación del departamento de Ingeniería de la USC Viterbi en la que desarrollaron una memoria artificial para ratones que posibilitaba restaurar la memoria a largo plazo.

Como veíamos en el caso de humanos-en-chip, existen también casos de empleo de estas tecnologías en el ejército para el control de emociones, dato extraído de una conferencia de Rodolfo Llinás, neurobiólogo que de-

sarrolló los primeros nanobots para el control de los síntomas del Parkinson, en la que mostró su malestar porque sus tecnologías estaban siendo utilizadas de este modo por el ejército.

Desafortunadamente, el sistema de patentes y el gran secretismo que rodea estas investigaciones (en especial cuando entran en juego agencias de seguridad y el ejército) hacen muy difícil conocer el alcance y los avances de estas tecnologías. Es por ello que se ha enviado una propuesta a la Asamblea General de las Naciones Unidas (Cavalcanti, 2009) que solicita que se considere esta tecnología como una tecnología abierta. Dicha propuesta argumenta que, del mismo modo que ocurre con el software libre y todas las prácticas *open source* que han contribuido a la mejora de los sistemas computacionales, abrir esta tecnología supondría un gran avance para su desarrollo y para la sociedad en general; asimismo solicita que sea declarada patrimonio de la humanidad para que se desarrolle según prácticas éticas con propósitos no bélicos.

Si la noción de sujeto está en juego, no sólo el desarrollo de este tipo de tecnologías tiene un papel muy relevante en la consolidación de nuevas teorías del sujeto, sino que el modo en que se utilicen marcará el futuro de la humanidad.

3. Conclusiones

Con el presente ensayo hemos tratado de ofrecer una panorámica de los cambios en la concepción del sujeto que se han ido operando a lo largo de la historia y especialmente a partir del auge de las ciencias cognitivas. Una de las tesis principales que defendemos en él es que los avances científicos en torno al estudio de la cognición humana han tenido como resultado un inesperado desplazamiento de fronteras que pondría de manifiesto los estrechos límites de la noción de humanidad que manejábamos, dejando tras de sí un gran rastro de sugerentes interrogantes que nos invitan a darles respuesta.

Donna Haraway anunciaba en el *Manifiesto Cyborg* (1991) tres grandes rupturas que marcarían la cultura científica de finales del siglo XX y comienzos del XXI: la ruptura de las fronteras humano/animal, las de lo humano/animal/máquina y las de lo físico/no físico. A lo largo del presente artículo hemos tratado de dar voz especialmente a la segunda, aunque como dejábamos entrever al principio, la primera de ellas, la distinción humano/animal, también está implicada en ésta y comienza a debilitarse con el arraigo de la biología y la teoría evolucionista, contribuyendo a difuminar la cada vez más delgada línea que separa a los humanos del resto de animales: tanto el lenguaje como la técnica (uso de herramientas), el comportamiento social o la atribución de estados mentales (capacidad de razonamiento, emoción, intencionalidad, etc.) han dejado de concebirse como rasgos exclusivos de los seres humanos.

La segunda, en la que más hemos ahondado a lo largo de este texto, tiene que ver con el cambio de paradigma en la concepción de las máquinas movido en gran medida por el enorme desarrollo técnico de los últimos años. Si bien el paradigma dualista aplicado a la robótica y la informática se sigue manteniendo en el vocabulario convencional referido a software y hardware, lo cierto es que nuestras máquinas parecen necesitar cada vez menos de los humanos y éstos depender cada vez en mayor medida de ellas. Éstas no sólo parecen estar *inquietantemente vivas*, como sostiene Haraway, sino que, además, han pasado a formar parte de nuestros propios cuerpos como prótesis que tanto ayudan a suplir deficiencias como crean discapacidades. Andrada y Sánchez mantienen que “si comprendemos que la televisión es una prótesis del ojo o el teléfono móvil del oído, veremos

que estos objetos han pasado de paliar una necesidad a generarla en su ausencia, reinventando una nueva condición natural por la que todos somos discapacitados” (Sánchez y Andrada 2013, p. 48). Esta reflexión resulta del todo pertinente en un contexto cultural y político de generalización de la tecnología a todos los ámbitos de la vida. Nuestra sociedad occidental avanza a pasos agigantados a una mediatización tecnológica generalizada en la que los artefactos y dispositivos cobran un papel central. Los avances en Inteligencia Artificial suman a esta mediatización tecnológica un componente que llega a rozar la ciencia ficción. La red está plagada de noticias relacionadas con nuevos hallazgos, inventos y aplicaciones de la tecnología a la vida. Son de especial relevancia las aplicaciones relacionadas con las ciencias de la salud. De hecho, es bastante frecuente que las investigaciones en este campo estén orientadas al cuidado y mejora de los cuerpos: nanobots que curan el Parkinson, chips que emulan órganos humanos para la realización de pruebas experimentales de quimioterapia controladas en laboratorio, implantes para monitorización de constantes vitales para pacientes con riesgo de infarto, etc. ganándose de este modo un alto índice de popularidad y entusiasmo entre la opinión pública y de la propia comunidad científica.

Lejos de querer caer en el pesimismo inmovilista o en el optimismo entusiasta en torno a las nuevas aplicaciones tecnológicas al ámbito de la vida y de los cuerpos, nos gustaría tomar una posición crítica y llamar la atención sobre las repercusiones que esta brecha tiene en las formas de organizarse y concebirse de los propios seres humanos. La importancia de dotar de perspectiva crítica a la situación actual, en la que todo nuevo hallazgo científico-tecnológico termina por equipararse a progreso de la humanidad, reside en colocar en el centro del debate a los y las agentes y sujetos de la ciencia: los seres humanos. Pretender que la objetividad científica sea equivalente a neutralidad política y beneficio inmediato supone olvidar los múltiples intereses que se ponen en juego en la empresa científica. Desde la financiación de las investigaciones, hasta su utilidad estratégico-militar, el contexto en el que se formulan las hipótesis de investigación y se desarrollan los experimentos está cargado de economía, política y estrategias de mercado. No olvidarlo nos permitirá comprender el lenguaje de los artefactos, su moralidad y sus implicaciones sociológicas y antropológicas.

Si bien es cierto que en la acelerada carrera por el desarrollo tecnológico cada vez más avanzado hemos hallado cuestiones que nos movían a cierta intranquilidad, este mismo desarrollo nos ha situado en un punto de inflexión en el que los parámetros por los que medir nuestra humanidad se han situado en el centro del debate. Es tarea de todos y de todas aprovechar la oportunidad que se nos presenta para establecerlos de forma consciente y procurar que en esta marcha hacia el futuro se involucren agentes que estaban fuera del debate científico hasta la fecha, dotándoles de capacidad de enunciación. En este contexto de hibridación y mestizaje entre seres humanos, máquinas y animales, el punto de vista de las minorías desempeña un rol fundamental, pues tal vez en él esté la clave que nos indique hacia dónde dirigir nuestra mirada.

4. Referencias bibliográficas

Cabañes Martínez, E. (2008). “Creadores Artificiales: ¿la creatividad más allá de lo humano?”. *Revista Tales*. (Núm. 1. Octubre 2008) Filosofía en el siglo XXI. Actas del I Congreso de Jóvenes Investigadores en Filosofía, 64-77.

Cabañes Martínez, E. (2013). “De la hibridación al procomún: construyendo la realidad a través de la tecnología”. *Revista de Estudios de Juventud*. (Núm. 102. Septiembre 2013). Instituto de la Juventud, 9-24.

- Campos, L. (1973).** *Diccionario de psicología del aprendizaje*. Editorial Ciencia de la Conducta, México.
- Cavalcanti, A. (2009).** "Nanorobot Invention and Linux: The Open Technology Factor - An Open Letter to UNO General Secretary". *CANNXS Project 1* (1). 1-4.
- Doménech-Sánchez A. y Vila, J. (2004)** "Fundamentos, tipos y aplicaciones de los arrays de ADN en la microbiología médica". *Enfermedades Infecciosas y Microbiología Clínica*. (Vol. 22. Núm. 1. Enero 2004), 46-54.
- Haraway, D. J. (1995).** *Ciencia, cyborgs y mujeres. La invención de la naturaleza*, Cátedra, Madrid.
- Huh, D. et al. (2012).** "A Human Disease Model of Drug Toxicity-Induced Pulmonary Edema in a Lung-on-a-Chip Microdevice". *Science Translational Medicine*. (Vol. 4. Núm 159. Noviembre 2012).
- Kantor, J. (1963/1991).** *La evolución científica de la psicología*. Trillas, Madrid.
- Martínez Freire, P. (2007).** *La importancia del conocimiento. Filosofía y ciencias cognitivas*. Editorial Netbiblo, Madrid.
- Newell, A. y Simon, H. A. (1972).** *Human problem solving*. Prentice-Hall, Englewood Cliffs (Nueva Jersey).
- Sánchez, P. y Andrada, G. (2013).** "Dispositivos, prótesis y artefactos de la subjetividad ciborg". *Revista de Estudios de Juventud*. (Núm. 102. Septiembre 2013). Instituto de la Juventud, 41-53.