

PENSAR Y PROCESAR.  
UNA CRÍTICA A LOS SISTEMAS DEMOCRÁTICOS  
BASADOS EN LA INTELIGENCIA ARTIFICIAL DESDE EL  
RECONOCIMIENTO, EL COMPROMISO Y LA JUSTICIA

*Bárbara Urban*<sup>1</sup>  
Universitat Jaume I  
barbara.urban@uji.es

RESUMEN / ABSTRACT

Este artículo propone una doble reflexión sobre las diferencias cognitivas entre humanos y máquinas desde la ética. En primer lugar, se tratará de diferenciar entre pensamiento humano y procesamiento máquina mediante métodos etimológicos, semánticos y comparativos que revelarán por qué no es posible atribuir capacidades pensantes a las inteligencias artificiales. Por otra parte, estos términos se relacionarán con la praxis de la democracia representativa a través de varios ejemplos. Este análisis se basa en tres pilares: la necesidad del reconocimiento recíproco mediante el lenguaje; el compromiso adquirido a través de la palabra; y la democracia como expresión de la justicia.

PALABRAS CLAVE: Pensamiento, procesamiento, democracia, reciprocidad, Arendt, Habermas, lenguaje, compromiso, democracia representativa

*THINKING AND PROCESSING.  
A REVIEW OF DEMOCRATIC SYSTEMS BASED ON ARTIFICIAL  
INTELLIGENCE FROM RECOGNITION, COMMITMENT AND JUSTICE*

This paper proposes a double reflection on the cognitive differences between humans and machines from ethics. In the first place, I will try to differentiate between human thought and machine processing through etymological, semantic and comparative

<sup>1</sup> Las comunicaciones a la autora deben dirigirse por correo electrónico o por correo postal al Departamento de Filosofía y Sociología de la Universitat Jaume I, Avda. Vicente Sos Baynat s/n, 12071, Castelló de la Plana, España.

methods that will reveal why it is not possible to attribute thinking capacities to artificial intelligences. On the other hand, these terms will be related to the praxis of representative democracy through various examples. This analysis is based on three aims: the need for mutual recognition through language; the commitment acquired through the word; and democracy as an expression of justice.

KEYWORDS: Thought, processing, democracy, reciprocity, Arendt, Habermas, language, commitment, representative democracy.

### Introducción

**R** Consideremos la siguiente pregunta: “¿Pueden pensar las máquinas?” Se trata de la famosa frase con la que Alan Turing abrió su conferencia en el National Physical Laboratory de Londres, en el año 1947. Más tarde la revista *Mind* (Turing 1950) la recogería en forma de artículo, dando así forma a la pregunta que, setenta años después, la comunidad filosófica se sigue haciendo. Y hablo de la comunidad filosófica, no de la científica, porque una parte de esta parece tener clara que la respuesta a la pregunta es la inteligencia artificial fuerte o inteligencia general (Boden, 2016, p. 29)

Este trabajo propone una doble reflexión. En primer lugar, se tratará de diferenciar de una forma clara entre pensamiento humano y procesamiento máquina. Para ello se utilizarán métodos etimológicos, semánticos y comparativos que revelarán por qué no es posible atribuir capacidades pensantes a las máquinas dotadas de inteligencia artificial. Por otra parte, una vez aclarados estos términos, se relacionará con la praxis de la democracia representativa a través del caso de Michihito Matsuda (Calvo 2019), el robot que en 2018 se presentó a la alcaldía del distrito de Tama, en Tokio.

A estas dos reflexiones subyacen tres fenómenos relacionados con la democracia. El primero, es la necesidad de reconocimiento recíproco a través del lenguaje, analizado mediante la teoría de la acción comunicativa de Habermas (2010). El segundo es el compromiso que, según Arendt (2005; 2013), se adquiere a través de la palabra, y que es imprescindible para que la praxis política sea recuperada por el pueblo y poder hablar de una democracia auténtica que hunde sus raíces en la solidaridad. Y el tercero es que, para poder hablar de democracia, en cualquiera de sus formas, tienen que primar criterios de justicia, cosa que, como se explicará a lo largo del texto, no es atribuible a ninguna máquina dotada de inteligencia artificial.

*Pensar y procesar: una mirada etimológica*

En primer lugar, es necesario determinar el significado de dos términos que se han utilizado indistintamente en muchas ocasiones hasta confundirse, no solo las palabras, sino también sus referentes.

Pensar, para la RAE proviene del latín *pensāre* (pesar, calcular, pensar). No obstante, Segura Munguía (1985: 516) sitúa este vocablo en la raíz *pēnsō*, de la que a su vez derivan otras palabras latinas como *pensāre*. Los significados latinos relacionados con el acto de pensar son:

Pesar, calcular.

Pensar, examinar, apreciar, juzgar, ser pesado con la misma balanza, ser juzgado con el mismo criterio, juzgar a los amigos por sus actos.

Pensar, meditar (una resolución).

En español, los significados derivados han sido variados, como indica el propio Segura Munguía:

Pensar, pensar cuidadosamente en pro y en contra, pensador, pensamiento, pensativo, pienso (como cuidado de alguien), pienso (como alimento de animales).

Pesada, pesado, pesadez, pesadilla.

Pesadumbre, apesadumbrar, pesaroso, apesasar, repesar.

Sopesar, apesgar, pésame, pésete.

Se pueden agrupar las derivaciones de *pensāre* en tres grupos de palabras que han llegado al español:

- Palabras que aluden a la reflexión (pensar, pensamiento, pensativo...).
- Palabras que aluden a la observación y al juicio (sopesar) y
- Palabras que aluden a los sentimientos y las emociones (pésame, pesadumbre, pienso, pesadilla).

Estas tres categorías corresponden a competencias humanas, muy alejadas de lo que es capaz de hacer un elemento dotado de inteligencia artificial.

Respecto de los significados del español moderno, de las siete acepciones de la RAE, ninguna sugiere que se trate de una acción que se pueda realizar de manera automatizada:

1. tr. Formar o combinar ideas o juicios en la mente.
2. tr. Examinar mentalmente algo con atención para formar un juicio.
3. tr. Opinar algo acerca de una persona o cosa.
4. tr. Tener la intención de hacer algo.
5. intr. Formar en la mente un juicio u opinión sobre algo.
6. intr. Recordar o traer a la mente algo o a alguien.
7. intr. Tener en consideración algo o a alguien al actuar.

En las definiciones, palabras como juicio, opinar, intención, recordar o tener en consideración son acciones difícilmente atribuibles, por ahora, a una máquina. No obstante, volviendo a las raíces latinas de *pensāre*, sí existen dos significados automatizables, como son pesar y calcular.

Por otra parte, el pensamiento es el resultado del acto humano de pensar y, a su vez, el acto de pensar es una función cerebral. Sin cerebro no hay pensamiento. Las máquinas no tienen cerebro, tienen otros elementos que se han construido en muchos casos inspirados en el propio cerebro humano, como las redes neuronales artificiales. Pero que algo funcione como un cerebro no lo convierte en un cerebro.

El pensamiento no es un proceso, es una capacidad humana, gracias a la cual somos capaces de valorar situaciones y emitir juicios, lo que permite atribuirnos el calificativo de seres morales. Una máquina está muy lejos de ser un ente moral.

Como contrapunto al concepto de pensamiento, y para atribuir a las máquinas una capacidad apropiada que las diferencie de los humanos, propongo utilizar el concepto de procesamiento. Citando de nuevo a Segura Munguía (2001: 601), pero esta vez en una revisión más moderna de su obra, *prōcessūs* significó en su origen latino avance, progresión, progreso; referido a un río, que sigue su curso efectivo, o como expresión, hacer progresos, progresar. Una pelota que rueda cuesta abajo sigue un proceso de descenso que no requiere reflexión por parte de ninguno de los elementos implicados. El río de la definición anterior fluye de manera natural, sin tener que pensar cuál será la mejor trayectoria. Estos procesos mecánicos se llevan a cabo de una manera que no ha sido previamente valorada por ninguna entidad con la inteligencia suficiente como para barajar opciones anticipándose a los hechos desde una perspectiva ética.

Tal y como la RAE indica, en la actualidad entendemos por procesar formar autos y procesos; declarar y tratar a alguien como presunto reo de delito; someter a un proceso de transformación física, química o biológica; o someter datos o materiales a una serie de operaciones programadas.

Procesar implica linealidad en un sistema de pasos para llegar de un punto A a un punto B. Esta no es la forma en que el ser humano piensa. Puede ser la manera que tenemos de llevar a cabo una receta de cocina, o tareas similares que requieren datos de entrada y un período de procesamiento para llegar a obtener

el resultado deseado, pero esta forma de actuar no implica nada más que replicar pasos que alguien ha pensado previamente.

Las máquinas funcionan en términos de *inputs* y *outputs*. Se introducen datos en ellas, o se les enseña a aprehenderlos. Tras esta operación, la máquina procesa mediante algoritmos y da como resultado un *output*, o una combinación de ellos. Huelga decir que el pensamiento humano opera de una forma mucho más compleja. Lejos de los *inputs* y los *outputs*, los humanos mostramos conductas derivadas de estímulos y emociones. Todo esto termina por hacer de nosotros personas diferentes a lo largo de una ontogenia finita marcada por experiencias. En el procesamiento de una máquina tanto *inputs* como *outputs* son datos, y la máquina sigue siendo infinitamente la misma. Si bien existe lo que se denomina *deep learning*, en que la máquina aprende a identificar elementos sobre la base de lo que va aprendiendo a lo largo de su funcionamiento, esto no transforma a la máquina como tal. Lo único que consigue es un aumento en su eficiencia, una mejora operativa, pero no un cambio.

Otra diferencia entre pensamiento y procesamiento es que los humanos tenemos la capacidad de hacer actividades sin una utilidad pragmática más allá del placer, sin ese componente de funcionalidad que tiene el procesamiento máquina. No se contempla la posibilidad de observar a una máquina disfrutando del placer de una tarde de domingo leyendo en un parque. Por el contrario, cada vez que un humano pone a funcionar una máquina, desde el más simple microondas hasta un complejo proyecto de reconocimiento de habla, lo hace con un objetivo determinado.

El pensamiento es una actividad individual, autónoma en condiciones no patológicas. Por el contrario, para que una máquina procese datos, alguien ha de mostrarle el camino para que llegue al punto que el programador desea.

Las máquinas no piensan, las máquinas procesan. Pensamiento humano y procesamiento máquina no tienen nada que ver, por más que algunos singularistas quieran buscar similitudes. Más que un problema de capacidades, estamos ante una cuestión de roles. Las máquinas nunca van a ser capaces de pensar, precisamente porque son máquinas, y el pensamiento es una aptitud humana. Es por eso que se propone hablar de pensamiento humano y procesamiento máquina. Una vez establecida esta distinción entre humanos y máquinas, se podrá empezar a hablar de la relación entre inteligencia artificial e inteligencia humana con mayor claridad. Para resumir con una analogía lo que he intentado exponer, los humanos pensamos y las máquinas procesan, al igual que los coches se conducen y los aviones se pilotan: la terminología acota competencias.

*Cuando las máquinas “piensan”: el caso de Michihito Matsuda*

Calvo (2019), hace un certero análisis de la dataficación de la esfera pública utilizando el ejemplo de Michihito Matsuda, que se reproduce a continuación:

El aumento exponencial del poder de los modelos matemáticos en la sociedad actual ha quedado de manifiesto en las elecciones de Tama New Town, un importante distrito de Tokio con más de 150.000 habitantes, realizadas en abril de 2018. En ellas, por primera vez en la historia, un algoritmo se presentó como alcaldable, y, para sorpresa de muchos, fue el tercer candidato más votado tras los más de 4.000 votantes recibidos durante la segunda vuelta de las elecciones. Michihito Matsuda, como así se apodaba este robot con rasgos femeninos, ofrecía a los votantes acabar con la corrupción política; dialogar y entenderse con el resto de fuerzas políticas por bien del distrito; así como proporcionar “fair and balances opportunities for everyone” [...]

Como aval para llevar a cabo su ambicioso proyecto político, Matsuda presentó su propio diseño algorítmico basado en Machine Learning. Este, según Matsuda, sustituye las debilidades emocionales de los seres humanos, principal causa de las malas decisiones políticas y de los conflictos de interés, por datos cuantificables y análisis estadísticos sobre los aspectos positivos de las propuestas políticas y las peticiones ciudadanas y sus posibles consecuencias. (Calvo, 2019)

Para empezar, convendría reflexionar sobre los rasgos femeninos del robot. La mayor parte de estos humanoides, sobre todo los dedicados a servicios, son máquinas feminizadas. Se pueden poner como ejemplos más populares LOLA en Portugal, que realiza labores de organización del flujo de usuarios en las administraciones públicas, o Sophia, una robot con nacionalidad saudí a la que se ha dado forma humana para poder exhibirla en platós de televisión y así recuperar la inversión que ha supuesto su creación.

Por otra parte, Michihito Matsuda –como todo político que se precie– basó su campaña en promesas que difícilmente podría cumplir. La incorruptibilidad política que se atribuye a la supuesta objetividad de los algoritmos es desmontada por múltiples teóricos, entre los que destaca Cathy O’Neil. Lo que ella denomina armas de destrucción matemática (O’Neil: 2018) son algoritmos que aprenden del entorno al que se les somete, y este entorno es un entorno humano. Por tanto, si los algoritmos han de aprender de nosotros lo harán sin discriminar entre lo que está bien o lo que está mal, porque un algoritmo no tiene la capacidad moral para discernir entre lo correcto y lo incorrecto.

Por otra parte, Matsuda prometió diálogo y entendimiento con el resto de fuerzas políticas, humanas. Recuperando el discurso de Habermas (2010), el reconocimiento recíproco ha de darse en un contexto de acción comunicativa. Difícilmente un robot y un humano se podrán comunicar reconociéndose mutuamente

como interlocutores válidos si partimos de la base de que no pertenecemos a la misma categoría: nosotros somos humanos, los robots son objetos.

En resumen, lo que se pretendía presentando la candidatura de un robot, era extraer de la política las emociones bajo la premisa falsa de la objetividad de las máquinas, sin tener en cuenta dos cosas: en primer lugar, para ejercer la democracia desde cualquier posición son necesarios criterios de justicia que un robot no posee y, en segundo lugar, la responsabilidad por los actos del robot queda automáticamente diluida, pues a una máquina no se le pueden exigir explicaciones: no son seres ni racionales ni morales.

No obstante, ¿cómo se explica el tercer puesto que logró el candidato robot? Por un lado, las razones tienen que ver con la falsa percepción que se tiene de la imparcialidad de las máquinas dotadas de inteligencia artificial. Por otro, veo necesario describir brevemente el fenómeno del valle inquietante. En *The Uncanny Valley*, el experto en robótica Masahiro (Mori 2012), describió mediante una función matemática un modelo perceptual que explicaba cómo los humanos tendemos a rechazar elementos robóticos que se parecen a nosotros, pero no lo suficiente. Hasta cierto punto, Mori afirma que existen sentimientos de afinidad con la máquina, pero sobrepasada cierta frontera de aspecto y movimiento tendemos a rechazarla, incluso con más vehemencia que a un cadáver.

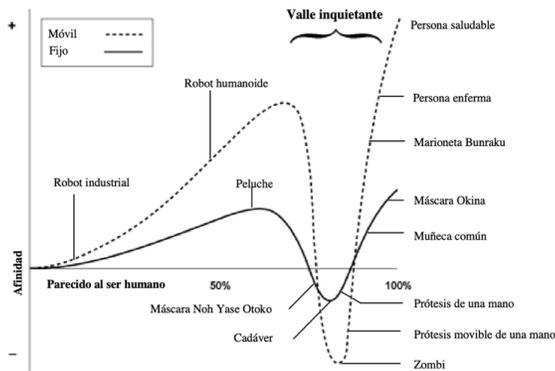


Ilustración 1. La presencia de movimiento intensifica el valle inquietante. El descenso de la flecha representa la muerte súbita de una persona sana. (Mejía Caballero 2013: 35).

Matsuda no ganó las elecciones por el mismo motivo que cualquier otro político las pierde: por falta de confianza. Y no porque su programa no fuese creíble, sino porque una máquina no inspira confianza. Una máquina sugiere fiabilidad. De

nuevo estamos ante una cuestión semántica, y es que las personas somos con-fiables, y las máquinas fiables. Nótese de nuevo la reciprocidad que denota el prefijo con-

Además del intento de dotar a las máquinas de inteligencia y de capacidad de pensamiento, también se ha pretendido elaborar una conciencia artificial, y ya no a través de la imitación de la conciencia humana; se da por hecho que el funcionamiento de la conciencia humana está más que desmenuzado, cuando en realidad la ciencia sabe muy poco sobre conciencia todavía.

### *Justicia, solidaridad y compromiso: cuestiones humanas*

La ausencia de competencias morales en las máquinas –por tanto, la incapacidad para obrar bajo criterios de justicia– y el hecho de que no inspiren confianza, sino fiabilidad, son motivos de peso para evitar implicarlas en la actividad política, al menos como dirigentes en primera línea.

Una idea central que aparece con frecuencia en el contexto político es la de compromiso (Kohn 2000: 78), entendido como la intención de llevar a cabo los acuerdos que se han conseguido durante el ejercicio de la participación política. A ese compromiso se llega mediante la negociación, que se sirve del lenguaje (76) como acción comunicativa. En palabras de Hannah Arendt:

En efecto, el pensamiento político solo puede observar las expresiones articuladas de los fenómenos políticos y está limitado a lo que aparece en el dominio de los asuntos humanos, que, a diferencia de lo que ocurre en el mundo físico, para manifestarse plenamente, necesitan de la palabra y de la articulación. (2013: 26)

Además, la promesa (Kohn 2000: 85; Castillo Cisneros 2009: 14) y el perdón son para Arendt dos componentes clave en la acción política. Respecto de la promesa,

Ya hemos mencionado el poder que se genera cuando las personas se reúnen y “actúan de común acuerdo”, poder que desaparece en cuanto se dispersan. La fuerza que las mantiene unidas [...] es la fuerza del contrato o de la mutua promesa. (Arendt 2005: 263)

Cuando se rompe un compromiso político derivado de un acuerdo, el perdón supone la capacidad para restaurar la confianza y un nuevo comienzo.

Siguiendo con la misma autora, la auténtica solidaridad, la solidaridad inclusiva, es, según Reshaur (1992) aquella que agrupa a los que sufren, pero también a los acomodados, que hacen de los problemas de aquellos su causa común. Es la única

forma de solidaridad que Arendt considera auténtica. Comprende los afectados y a quienes, sin ser parte del problema, buscan hacer del sufrimiento una causa común.

El compromiso con los electores mediado por el lenguaje, la capacidad para la promesa, que implica un común acuerdo, la aplicación de medidas de gobierno bajo criterios de justicia y la solidaridad son atribuciones humanas que de ningún modo se pueden encontrar en un robot, por mucho que prometa un gobierno objetivo.

### *Conclusiones*

El fenómeno de la algoritmización de la democracia, y concretamente el ejemplo de Michihito Matsuda, es una respuesta a la desafección, definida por García Marzá como:

Con la palabra “desafección” nos referimos al escaso aprecio y estima que le tienen los ciudadanos a la vida política y a sus instituciones y representantes. Constituye un conjunto de sentimientos y actitudes reactivas caracterizado, por una parte, por la aceptación resignada de la democracia como la menos mala de las formas de gobierno, a la que no se quiere renunciar y, por otra, por un fuerte desapego, inhibición e incluso hostilidad, ante la gestión política y sus actores e instituciones. (García-Marzá, 2015)

La promesa de objetividad e incorruptibilidad fue demasiado tentadora como para dejarla pasar, y el resultado fue la gran aceptación del robot alcalde por parte del electorado.

A modo de resumen, un robot no piensa, sino que procesa datos aportados, de una forma dada y con una finalidad establecida. En segundo lugar, y extrapolando esto a la cuestión de la democracia, un modelo democrático requiere una serie de atribuciones inherentes a los seres humanos que no se dan en las máquinas, como son el compromiso, la solidaridad, la aplicación de criterios de justicia y el reconocimiento recíproco.

Esta última característica se ha relacionado con el fenómeno del valle inquietante, en que elementos robóticos con forma humana nos producen rechazo cuanto más humanos parecen. Pero también tiene que ver con el hecho de que confiar en un robot no es posible, una vez más por cuestiones semánticas y de atribuciones, ya que la confianza es un sentimiento que se da entre humanos, distinto a la fiabilidad, que es una cualidad de las máquinas.

Para terminar, propongo refocalizar la solución al problema. La desafección, como ya se ha visto anteriormente, tiene sus orígenes en fenómenos como la

falta de credibilidad en la política, la indiferencia o la desconfianza. La propuesta fácil para algunos pasa por poner un robot que solucione lo que la avaricia de los humanos estropea.

Hay que implementar la ética en la política, pero no en las máquinas, sino en los humanos, y muy especialmente en aquellos que tienen el honorable encargo de ejercer la representación en un sistema de democracia representativa.

Lo que acabo de proponer puede parecer una obviedad, pero también es obvio que, a estas alturas del desarrollo democrático, el Congreso, el Senado, las asambleas de las comunidades autónomas, los ayuntamientos y las diputaciones deberían contar con comisiones éticas permanentes, y eso, hoy por hoy, es una carencia de consecuencias visibles en las noticias prácticamente a diario.

### Referencias bibliográficas

- Arendt, Hannah (2005), *La condición humana*. Barcelona: Paidós.
- \_\_\_\_ (2013), *Sobre la revolución*. Madrid: Alianza.
- Boden, Margaret (2016), *Inteligencia artificial*. Madrid: Turner.
- Calvo, Patrici (2019), “Democracia algorítmica: consideraciones éticas sobre la dataficación de la esfera pública”, *Revista del CLAD Reforma y Democracia* 74: 5-30.
- Castillo Cisneros, Mery (2009), “Libertad y justicia en Hannah Arendt: una aproximación”, *Desafíos* 20: 11-29.
- García-Marzá, Domingo (2015), “El valor democrático de la sociedad civil: una respuesta a la desafección”, *Thémata. Revista de filosofía* 52: 93-109.
- Habermas, Jürgen (2010), *Teoría de la acción comunicativa*, vol. II. Madrid: Trotta.
- Kohn, Carlos (2000), “Solidaridad y poder comunicativo: La praxis de la libertad en la filosofía política de Hannah Arendt”, *Res publica* 5: 73-92.
- Mejía Caballero, Cristian Andrés (2013). *El desarrollo de robots humanoides en Japón y las implicaciones del valle inquietante*. Salamanca: Universidad de Salamanca.
- Mori, Mashairo (2012), “The Uncanny Valley”, *IEEE Robotics and Automation Magazine* 19/2: 98-100.
- O’Neil, Cathy (2017). *Armas de destrucción matemática. Cómo el big data aumenta la desigualdad y amenaza la democracia*. Madrid: Capitan Swing.
- RAE. *Diccionario de la lengua española*. [www.dle.rae.es](http://www.dle.rae.es).

Reshaur, Ken (1992), "Concepts of Solidarity in the Political Theory of Hannah Arendt, *Canadian Journal of Political Science / Revue canadienne de science politique* 25/4: 723-736.

Segura Munguía, Santiago (1985), *Diccionario etimológico latino-español*. Madrid: Anaya.

\_\_\_\_ (2001), *Nuevo diccionario etimológico latino-español y de las voces derivadas*. Bilbao: Universidad de Deusto.

Turing, Alan M. (1950), "Computing machinery and intelligence". *Mind*, 433-460.

