



**LA TEORÍA DE LA VIRTUD EN ARISTÓTELES: DILEMAS EN TORNO AL PROBLEMA  
DE LA INTELIGENCIA ARTIFICIAL**

**Vettulo Maier, Rocío**

*Universidad Nacional del Litoral - Facultad de Humanidades y Ciencias*

Director: Berrón, Manuel

Área: Humanidades

Palabras claves: virtud, phronesis, inteligencia artificial

Título del proyecto: La teoría de la virtud en Aristóteles: dilemas en torno al problema de la Inteligencia Artificial.

Instrumento: Adscripción en Investigación, cátedra de Filosofía Antigua.

Año de convocatoria: 2021

Organismo: UNL

Director: Berrón, Manuel

## INTRODUCCIÓN

El presente trabajo tiene por objeto realizar un exhaustivo análisis de los dilemas que atraviesa la inteligencia artificial (IA) en lo que concierne al ámbito ético, clarificando los rasgos concretos que la diferencian del pensamiento humano y evaluando cuan posible es lograr una IA que posea la capacidad humana de seguir un comportamiento ético. La perspectiva ética sobre la cual se analizará la calidad y posibilidad de la IA de otorgar decisiones y soluciones prudentes será la teoría de la virtud presentada en la obra de la *Ética a Nicómaco* del filósofo griego Aristóteles.

De este modo, ubicándonos en el paradigma actual y el avance que supone la inteligencia artificial para la humanidad en diferentes ámbitos y campos no solo del conocimiento sino también de la praxis, examinaremos como el carácter multidisciplinar que brinda la IA como instrumento al alcance de la mano nos exige una reflexión filosófica-ética que ponga en tela de juicio la calidad moral del desarrollo de sistema de IA y las respuestas que estos brindan ante las consultas y necesidades de los usuarios, pues no son pocas las decisiones que parecen recaer bajo la responsabilidad de los algoritmos.

De este modo, la ética de las virtudes, por sobre otras éticas como la deontológica o la utilitarista, se destaca por su objeto protagonista: la virtud. En la tradición griega, la *areté* (virtud) era aquello que llevaba a término la buena disposición de su objeto, por tanto, es que Aristóteles decreta que al hombre *qua* hombre también le es inherente una virtud que se adecua a su capacidad racional y se asocia a la vida social-política donde, posteriormente, alcanza el pleno desarrollo de sus capacidades morales e intelectuales.

El ser humano entonces, siguiendo lo propuesto por el filósofo estagirita, en tanto *zoon logistikón* está provisto de herramientas que puede usar o no para la sensatez y la virtud. En consecuencia, en tanto que posee logos, el ser humano, tiene la posibilidad de llevar a término el perfeccionamiento de su propia naturaleza, esto es, no solo conocer su fin sino de elegirlo y lograrlo. Este codiciado fin humano del que habla este libro es la *eudaimonía*, en su traducción más aceptada: felicidad o vida plena, que para lograrse requiere de una vida virtuosa en la cual por medio de un cierto aprendizaje (*mátheis*) o ejercicio (*áskesis*) y a partir del conocimiento de lo correcto, el ser humano, en su modo de actuar y ser, aprende a discernir lo correcto, tornando sus decisiones y actos hacia la prudencia (*phronesis*). La prudencia se presenta entonces como el justo medio entre dos extremos y al cual tiende el hombre de carácter virtuoso en sus actos y decisiones.

La virtud aristotélica engloba entonces los siguientes rasgos: justo medio (*mesotes*), deliberación (*boúleusis*), elección (*proaíresis*), educación (*paideia*), conocimiento. Requiere de una vida entera, es decir, no meramente de intervalos virtuosos y de decisiones que no solo reconozcan, sino también que se redireccionen al justo medio.

Sin embargo, si nos posicionamos en el contexto socio-cultural actual, la inteligencia artificial ha escalado a espacios impensables. La promesa futurista del *big data* como conocimiento absoluto y, por ello, facilitador de la vida en varios aspectos es hoy en día una realidad. Todo conocimiento, todo objeto y toda responsabilidad y elección de la vida cotidiana parece delegarse a los algoritmos por lo que cabe la siguiente pregunta: ¿existe algo que un conjunto de caracteres no pueda solucionar? la respuesta es dudosa. No obstante, entre tanta tecnología e innovaciones en el campo digital, la verdad es que a día de hoy la IA no supone una amenaza para el pensamiento del ser humano. De hecho, se trata de una forma de saber bastante primitiva en el que todo es calculable, predecible y controlable. La *data mining* o minería de datos descubre las correlaciones y el saber absoluto del *big data* se queda estancado precisamente en estas correlaciones y reconocimiento de patrones en los que, empero, nada se comprende.

La inteligencia artificial no puede reproducir al ser humano en muchos aspectos, sus límites se encuentran en la capacidad afectiva. La IA no tiene mundo, no se encuentra en él. No es capaz de generar un pensamiento propio, ni tampoco de interrumpir el curso histórico- contextual con un pensamiento nuevo.

El pathos, la pasión, es el comienzo del pensamiento. La inteligencia artificial es apática, no siente. Su límite es el cálculo dentro de horizontes estrictamente marcados.

Pero, pese a estas diferencias abismales y esta incapacidad afectiva que supone más de un inconveniente, ¿por qué se considera que algunas decisiones, incluso las que deben ser tomadas con más cuidado y prudencia, pueden encomendarlas a la IA? ¿es acaso es posible que la IA sugiera una decisión prudente? Si la respuesta llegara a ser afirmativa cabría destacar que la prudencia de la elección que sugiriera la IA no radicaría más que en el sujeto que la interpreta.

El pensamiento necesita de la totalidad del mundo para conceptualizarlo, la inteligencia artificial no necesita un mundo porque se limita a procesar hechos predeterminados y antiguos, no puede darse a sí misma nuevos hechos, no puede entenderlos de forma diferente a la que ya tiene. El pensamiento humano es, en definitiva, más que cálculo y resolución de problemas.

## OBJETIVOS

- Dilucidar las diferencias entre la inteligencia artificial y el pensamiento humano.
- Examinar la posibilidad de concebir una ética para la inteligencia artificial partiendo de lo que entiende Aristóteles como 'decisiones prudentes' desde su teoría ética de la virtud.

## METODOLOGÍA

Para alcanzar los objetivos propuestos se implementó una metodología dividida en tres etapas específicas en la que el avance exitoso de cada etapa brinda el fundamento para la siguiente. El eje sobre el cual giraba cada etapa fue preestablecido al comienzo de la investigación, no obstante, estuvo sujeto a modificaciones.

La primera etapa consistió en un análisis profundo y detallado de los dos pilares sobre los cuales se apoya la investigación: la ética de las virtudes y la IA. Para lograr esto se rastreó en la bibliografía primaria todo aquello que pudiera ser objeto de consideración para el presente trabajo. Se plantearon los primeros interrogantes y se visualizaron posibles senderos por los cuales avanzar. Por otro lado, para lograr un avance certero y guiado en el campo de la IA se recurrió a la experiencia y conocimientos de expertos e investigadores del Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional (SINC-i) por medio de sucesivas entrevistas y consultas.

En la segunda etapa, ya habiendo recopilado lecturas y entrevistas, se procedió a analizar las consecuencias y límites que plantea la inteligencia artificial en su vínculo con teorías éticas como la de la virtud. Con un enfoque en lo intrínsecamente humano, se persiguió el logro del primer objetivo: identificar los elementos que distinguen a la inteligencia humana de la inteligencia artificial. A partir de la respuesta de este primer objetivo, se plantearon nuevas problemáticas y posibles respuestas para cumplir el segundo objetivo.

Finalmente, la última instancia se basó en hallar un lineamiento general y transversal que tocara a cada uno de los interrogantes planteados. Así, ya habiéndose analizado en detalle los límites y consecuencias que supone el vínculo entre la inteligencia artificial y la ética de la virtud se logró arribar a una conclusión que, si bien no era la esperada, era, no obstante, satisfactoria.

## CONCLUSIÓN

Si bien la inteligencia artificial puede reparar al pensamiento humano en su capacidad conceptual, no puede, sin embargo, generar decisiones prudentes, ni virtuosas. La inteligencia artificial no conoce un término medio, no conoce los extremos sobre los cuales se debaten y deliberan algunas elecciones. La máquina no constituye aún una amenaza para el pensamiento humano, pues no trae consigo ninguna novedad, es una forma de saber aún primitivo. Así, al no manejarse en el mundo cabe la pregunta de si es posible concebir una ética de la inteligencia artificial, porque, en última instancia, su uso y repercusión en la vida humana no depende más que del usuario.

## BIBLIOGRAFÍA

- **Aristóteles.** *Ética a Nicómaco*. Madrid: Instituto de Estudios Políticos. 1959. Introd., trad. y notas de Julián Marías y María Araujo.
- **Monasterio, A.** 2017: "*Ética algorítmica: Implicaciones éticas de una sociedad cadavez más gobernada por algoritmos*", *Dilemata*, 185-217.
- **Moreau J.** 1962: *Aristóteles y su escuela*, traducción de Marino Ayerra, Buenos Aires, Editorial Universitaria de Buenas Aires.
- **Orts, A.** 2019: *Ética de la inteligencia artificial*. Valencia. (Proyecto de Investigación Científica y Desarrollo Tecnológico).
- **Rorty Amélie Oksenberg.** 1981: *Essays on Aristotle's Ethics* (Univ. California Press).
- **Terrones, A.** 2018: *Inteligencia artificial y ética de la responsabilidad*. Ecuador. *Cuestiones de Filosofía*, 4 (22), 141-170.
- **Byung-Chul Han.** 2021: *No-cosas, quiebras del mundo de hoy*. Taurus. Trad. Joaquin Chamorro Mielke.