

## CONSIDERACIONES SOBRE LA ASIGNACIÓN DE RESPONSABILIDAD PENAL A LA INTELIGENCIA ARTIFICIAL<sup>57</sup>

Federico Daniel Arrué<sup>58</sup>

### RESUMEN:

*El trabajo procura ordenar algunas ideas sobre la hipotética asignación de responsabilidad penal a sistemas de inteligencia artificial en cuanto a argumentos a favor de su establecimiento, compatibilidad con las distintas teorías de la finalidad de la pena e implicancias para el paradigma penal del ser humano como único sujeto penalmente responsable.*

*Se encuadra en una invitación a un debate serio y continuo sobre la asimilación de la inteligencia artificial a los seres humanos en las distintas áreas del derecho.*

**Palabras Clave:** Inteligencia artificial – Responsabilidad penal- Reproche penal – Pena – Teorías de la pena.

### INTRODUCCIÓN Y ESQUEMA DEL TRABAJO

Desde estas acotadísimas líneas procuraremos ordenar algunas ideas sobre las posibles ventajas de asignar responsabilidad penal a sistemas de inteligencia artificial y las dificultades que

esto conllevaría.

Primeramente daremos cuenta de posibles argumentos a favor del establecimiento de tal responsabilidad. En segundo lugar, referiremos a su viabilidad desde el punto de vista de la finalidad de la pena. Seguidamente resaltaremos que

---

57 Enlace al video <https://youtu.be/7JogBIUPVJs>

58 Abogado, Especialista en Derecho Penal. Magister en Derecho. Investigador y Docente de las cátedras de Filosofía del Derecho y de Ciencia Política en la Universidad Nacional del Sur. [federicoarrue@hotmail.com](mailto:federicoarrue@hotmail.com)

el paradigma clásico del derecho penal que consagra al ser humano como único sujeto penalmente responsable ya ha sido roto en el derecho argentino, por lo cual la hipotética irrupción de la responsabilidad penal de la inteligencia artificial sería menos abrupta de lo que puede pensarse en un primer momento. Por último, cerraremos con unas palabras finales.

Por razones de espacio no haremos referencia a qué puede entenderse por inteligencia artificial ni a la compatibilidad de las penas previstas en el Código Penal argentino con ella, ni tampoco a posibles nuevas penas que pueden llegar a establecerse específicamente para sancionar a una inteligencia artificial.

## ARGUMENTOS A FAVOR DEL ESTABLECIMIENTO DE LA RESPONSABILIDAD PENAL DE LA INTELIGENCIA ARTIFICIAL.

### ***Nuevo incentivo a los productores.***

Para la creación de inteligencias artificiales que guarden el máximo nivel de seguridad, los productores tienen diversos incentivos tales como evitar su posible propia responsabilidad por no haber realizado los controles necesarios al producto antes de su puesta en el mercado o por no haber tomado las

medidas necesarias si la falla surgiera con posterioridad. E, incluso libres de responsabilidad penal –si actuaron con total diligencia-, es esperable también que procuren que ninguna inteligencia artificial que lleve su firma se vea envuelta en un hecho dañoso que le sea a ésta reprochable.

En este sentido, la eventual responsabilidad penal de la inteligencia artificial podría conllevar que los productores realicen un seguimiento aún más exhaustivo a fin de evitar la publicidad negativa superlativa que implicaría que uno de sus productos fuera condenado penalmente.

### ***Extensión de las posibilidades de actuación del sistema penal.***

El establecimiento de la responsabilidad penal de la inteligencia artificial implicaría también que, en los hechos dañosos donde ésta esté involucrada como sujeto activo de un presunto delito, tome intervención –de oficio, en la mayoría de los casos, según el Código Penal argentino- el órgano Fiscal. En sus manos estaría, en primer término, el esclarecimiento de lo ocurrido y la determinación de una eventual responsabilidad penal, como paso previo al juzgamiento en tal sede. De no ser así,

la determinación judicial de los hechos dañosos estaría librada a la existencia de una acción civil por parte de la víctima, puesto que la Fiscalía sólo intervendría de resultar plausible la responsabilidad penal del productor de la inteligencia artificial, de un tercero que ilegítimamente haya alterado a ésta o del usuario humano.

La consagración de la responsabilidad penal en cabeza de la inteligencia artificial, entonces, en determinados casos puede habilitar las posibilidades de actuación de la justicia penal.

**Facilitación de la acción civil. Equiparación de situaciones.** La intervención del sistema penal en casos donde no existe una persona física imputada tiene también un beneficio ligado a la compensación económica por el daño causado. Esto es así porque -dejando de lado el problema de la concreta ejecución de una condena penal sobre una inteligencia artificial-, su sola existencia es relevante, no sólo por su rol simbólico social sino porque facilita el ejercicio de una eventual acción civil por parte de la víctima o sus familiares. Basta pensar en la práctica habitual del fuero penal, donde muchas veces el interés principal del damnificado en la condena pasa por ver reforzadas sus eventuales

aspiraciones económicas dado que, frente a un veredicto de culpabilidad, en sede civil ya no puede discutirse hecho ni autoría.

En este sentido, la víctima de un hecho producido por una inteligencia artificial puede verse en una desventaja procesal en relación a la víctima de un hecho similar provocado por una persona física.

### ***Mantenimiento de la confianza en el sistema penal.***

Por último, diremos que en una presuntamente no tan lejana sociedad donde la inteligencia artificial pasaría a realizar una cantidad variada y substancial de actividades cotidianas, su responsabilidad penal puede ser conveniente para la reafirmación de la vigencia y eficacia del sistema. Es decir que, al ser el comportamiento y funciones de las inteligencias artificiales cada vez más semejantes al de los seres humanos, su inimputabilidad resquebrajaría la confianza ciudadana en las normas prohibitivas y punitivas de conductas. Así, para parte de la población, podría tornarse complejo entender y aceptar que una inteligencia que realiza habitualmente las mismas conductas que un humano no sea juzgada penalmente por un hecho que, de ser cometido por un humano, conduciría

sin dudas a su procesamiento.

El ciudadano podría concebir, de modo más o menos consciente, a las inteligencias artificiales como entidades que *optan* por delinquir –pese a que no esté claro que pueda ser sin más expresado de ese modo- y que, pese a eso, si no fuera reputada penalmente responsable, quedaría impune.

Esta función punitiva simbólica no podría ser reemplazada por medidas de prevención administrativas, pues éstas operan por fuera de lo penal.

## LOS PROPÓSITOS DE LA PENA A LA INTELIGENCIA ARTIFICIAL DESDE LAS DISTINTAS TEORÍAS

La *concepción retributiva de la pena* entiende que es justo o correcto que el delincuente reciba un castigo en atención a su acto, independientemente de cualquier debate sobre la utilidad o beneficio social o personal de éste. Desde esta perspectiva, corresponde punir a la inteligencia artificial si puede considerársela moralmente responsable, lo cual presumiblemente esté ligado a que haya desarrollado la capacidad de comprender –en un sentido semejante al humano- las nociones básicas de bien y de mal, y, además, que haya desarrollado

también la capacidad de receptor sentimientos no deseados, dolorosos – también en un sentido similar al humano- asociados a la pena. Es decir, que esta línea de pensamiento sobre la función de la pena tendrá sentido si la inteligencia artificial puede, de alguna forma, concebir el mal: tanto el que ha realizado como el que habría de recibir en justa retribución.

En contrapartida, las cuatro teorías de la función de la sanción llamadas *preventivas* se centran en atender a la utilidad ésta.

La finalidad de *prevención especial negativa* de la pena implica que ésta debe aplicarse para neutralizar la potencialidad delictiva del delincuente concreto que la sufre. Esto sería plenamente aplicable a la inteligencia artificial en cuanto ésta pase a no poder o tener dificultada la comisión de nuevos delitos mientras dure la sanción que se le haya impuesto.

La finalidad de *prevención general negativa* considera que la pena debe imponerse para desalentar al conjunto de potenciales delincuentes de cometer hechos delictivos. La pena es vista como un ejemplo atemorizante.

Entendemos que esta visión también resulta aplicable a la punición de la inteligencia artificial, la cual puede servir para amedrentar a personas físicas. El

mensaje de amenaza del sistema penal a la sociedad será de una entidad tal que ni siquiera algunos programas informáticos estarán al margen de él.

Por otra parte, incluso es posible que la sanción impuesta a una inteligencia artificial pueda influir en el comportamiento de otra inteligencia artificial, si esta pudiera receptar el dato y tenerlo en cuenta para su propio accionar.

La finalidad de *prevención especial positiva* plantea que el castigo puede servir para que el delincuente que lo sufre recapacite sobre su conducta y no reincida en el delito. Es la teoría de la rehabilitación y la reinserción social. Trasladado a la inteligencia artificial, en línea con lo que se dijo anteriormente, tendría sentido siempre cuando el programa tuviera la capacidad de internalizar la sanción que él ha sufrido como un dato de la realidad que modifique a futuro sus cálculos con miras a minimizar la posibilidad de que su curso de acción escogido sea un delito.

Por último, en cuanto a la finalidad de *prevención general positiva*, ésta se basa en que la sanción reafirma ante la sociedad la vigencia y efectividad tanto de la norma penal violada en la comisión de delito como del sistema penal en general. Tiene un papel pedagógico, comunicativo

y demostrativo de la superación de un conflicto jurídico.

Creemos aplicable lo dicho en cuanto a la prevención general negativa. Desde esta óptica tendría sentido punir a la inteligencia artificial en tanto muestra –a las personas físicas y eventualmente también a otras inteligencias artificiales- el correcto funcionamiento de la normativa y del sistema penal.

## EVOLUCIÓN SOCIAL Y CAMBIOS DE PARADIGMA

Pretender la responsabilidad penal de la inteligencia artificial conlleva problemáticas que pasan no sólo por la justificación teleológica de la imposición de una sanción penal (tal como se ha apuntado), sino también por el amoldamiento de dichas sanciones a las características propias de los nuevos sujetos punibles (a lo que, tal como anticipamos, no referiremos aquí). A esto se suma la posible ruptura de las concepciones de acción y de culpabilidad –dentro de la teoría estratificada del delito- que aporta la dogmática penal, ligadas a la persona humana como único sujeto penalmente responsable.

En relación a esto último, sin embargo, el cambio ya ha comenzado, puesto que en

el derecho argentino, actualmente, las personas físicas no son los únicos sujetos susceptibles de responsabilidad penal y las penas previstas en el Código Penal no son las únicas sanciones del sistema punitivo. Esto es así desde que la ley 27.401, del año 2017, complementaria del Código Penal, establece la *responsabilidad penal aplicable a las personas jurídicas privadas* por determinados delitos.

La responsabilidad penal de personas jurídicas –y, por analogía, de inteligencias artificiales- no carece tampoco de sustento doctrinario, ya que las *concepciones funcionalistas del Derecho Penal*, al no reconocer límites ónticos, entienden que la opción de sancionar penalmente es pura decisión legal.

Como en esta perspectiva los conceptos jurídicos del Derecho Penal no pueden ser independientes de los fines político criminales –y políticos, en general- del sistema penal, es el legislador quien, a partir de dichos fines, decidirá quién es penalmente sancionable. Así, del mismo modo en que la evolución de la sociedad llevó a establecer la conveniencia de la responsabilidad penal de las personas jurídicas, adaptando la normativa penal a sus especificidades; puede también llegar a evaluar la conveniencia de pautar la responsabilidad penal de la inteligencia

artificial, con las adaptaciones que ello requiera.

## PALABRAS FINALES

Creemos que, a medida que las inteligencias artificiales cobran relevancia en nuestra vida cotidiana y en nuestra sociedad, corresponde dar paso a un debate serio y continuo sobre su asimilación a los seres humanos, en las distintas áreas del derecho, incluido el ámbito penal. Dicho debate, académico, técnico y legislativo, debe estar abierto a la posibilidad de grandes cambios de paradigmas, semejantes a los que las nuevas tecnologías están imprimiendo ya en nuestra cotidianeidad.

## BIBLIOGRAFÍA

- Anitua, Gabriel Ignacio (2002), Historia de los pensamientos criminológicos, Buenos Aires, Editores del Puerto.
- Granero, Horacio R. (Director académico) (2020), Inteligencia Artificial y Derecho, un reto social, Buenos Aires, Albremática.
- Raffo, Francisca Emilia, Nuevas tecnologías y derecho penal (2019), La responsabilidad penal de la inteligencia artificial.



<http://repositorio.udesa.edu.ar/jspui/handle/10908/16303> .

- Zaffaroni, Raúl Eugenio (2018), Manual de Derecho Penal – Parte General, Buenos Aires, Ediar.

