



Fotograma de la película *La naranja mecánica* (1971). Stanley Kubrick

Inteligencia artificial y autonomía: reflexiones sobre dilemas (bio)éticos a partir de un caso de ficción... “La naranja mecánica”

Artificial Intelligence and Autonomy: Considerations on Ethical Dilemmas in the Novelated Fictional Case... “A Clockwork Orange”



Ailin Irina Gurfein

Facultad Latinoamericana de Ciencias Sociales

agurfein@hotmail.com.ar

 <https://orcid.org/0000-0002-8543-1074>



Resumen

El siguiente trabajo pone en discusión avances terapéuticos que implican modificaciones en cuerpos humanos mediante inteligencia artificial (IA). Una de las preguntas principales incluye ensayos clínicos, sujetos de investigación y autonomía. Se analizarán aspectos ético-comunicacionales en el vínculo entre investigador/a y sujeto de investigación, así como aspectos de la relación médico-paciente en la implementación de IA para tratamientos terapéuticos. ¿Debe el profesional de la salud transparentar su aplicación a los pacientes?

A su vez, se analizará el recorte específico de ensayos clínicos en personas privadas de la libertad, tomando como analogía el caso de ficción *La naranja mecánica* para reflexionar sobre la decisión autónoma en circunstancias extremas desde una perspectiva bioética a la vez que intentaremos responder a la pregunta, ¿es la modificación artificial del ser una forma ética de resolver problemas sociales?

Abstract

The following work discusses therapeutic advances that involve modifications in human bodies. One of the main questions includes clinical trials, research subjects and the ability to decide. Ethical-communicational aspects in the relationship between researcher and research subject and doctor-patient will be analyzed. Should the possibility of an innovative treatment in which technologies that modify their organism be put into dialogue between researcher and research subject, or between doctor and patient?

In turn, the specific cut of clinical trials in people deprived of liberty will be analyzed, taking as an analogy the fictional case of A Clockwork Orange to reflect on the autonomous decision in extreme circumstances from a bioethical perspective while we will try to respond to the question: Is the artificial modification of being an ethical way of solving social problems?

Key words

Bioética; comunicación; relación médico-paciente; inteligencia artificial; autonomía.
Bioethics; communication; doctor-patient relationship; artificial intelligence; autonomy.

Fechas

Recibido: 18/08/2023. Aceptado: 13/05/2024



1. Introducción

La explosión de la inteligencia artificial (IA) al servicio de la medicina evoca reflexiones necesarias para que la misma obre como herramienta que aporte al progreso de la salud sin poner en riesgo aspectos de la esencia humana.

Desde este enfoque, una de las dificultades que puede plantearse es la de la asimetría de saber. ¿Puede una persona consentir plenamente cuando desconoce los alcances de un tratamiento?

Desde la bioética podemos preguntarnos qué caminos existen para lograr que los avances científico-tecnológicos se encuentren en comunión con los derechos individuales de todo ser humano; entre ellos, y como punto de partida, el derecho a decidir de manera autónoma.

De aquí surgen preguntas en busca de un equilibrio que supere los extremos de la tecnofilia y la tecnofobia.

Desde nuestra perspectiva, consideramos que los aspectos comunicacionales constituyen un factor fundamental para lograrlo.

Partiendo de la concepción sobre la relación médico-paciente (RMP) colaborativa, en la que la toma de decisiones es compartida, debemos reparar en uno de los elementos clave que hace a la consecución de tal vínculo de forma efectiva: el principio de autonomía.

Desde este enfoque, una de las dificultades que puede plantearse es la de la asimetría de saber. ¿Puede una persona consentir plenamente cuando desconoce los alcances de un tratamiento?

Si bien la modificación del cuerpo humano a través de la medicina (operaciones reconstituyentes, prótesis, cirugías estéticas, entre otras) lleva décadas de práctica ininterrumpida, a la vez que las prótesis y tecnologías terapéuticas han evolucionado mejorando notablemente la calidad de vida de los individuos, creemos que la IA aplicada a seres humanos merece un nuevo capítulo.

Consta aclarar que este trabajo no rechaza el empleo de IA para tratamientos terapéuticos o preventivos que hayan superado la fase experimental con los ensayos clínicos previos, éticamente practicados; por ello mismo tiene por objetivo reflexionar sobre los alcances éticos y morales que atañen a la influencia de estas nuevas posibilidades, y la delicadeza con que debieran ser abordados.

Con este objeto, tomaremos como caso paradigmático la famosa novela de ficción de Anthony Burgess *La naranja mecánica*, ya que plantea un doble dilema: 1. El de la decisión *versus* la coerción para participar de un ensayo clínico como sujeto de investigación. 2. La modificación del sujeto a partir de un tratamiento que coarta la capacidad de decidir.



2. Primera parte o “un diálogo sobre la moral”

—Señor, he hecho todo lo posible, ¿verdad? (...) Me he esforzado, ¿verdad?
 —Creo —dijo el chaplino— que en general te has portado bien, 6655321. Colaboraste, y creo que has mostrado verdaderos deseos de reformarte. Si sigues así, conseguirás fácilmente que te reduzcan la pena.
 —Pero, señor —lo interrumpí—, ¿qué hay de ese nuevo tratamiento que permite salir enseguida y garantiza que uno nunca vuelve?
 (...)

 —Supongo que te refieres a la técnica de Ludovico.
 —Ignoro cómo la llaman, señor —dije—. Solo sé que a uno lo saca rápidamente de aquí, y aseguran contra toda posible vuelta.
 —Así es —dijo el chaplino, mirándome y frunciendo el ceño—. Así es, 6655321. Por supuesto, no ha pasado la etapa experimental. Es algo muy sencillo, pero muy drástico.
 —¿Pero no la están usando aquí, señor?— pregunté. Esos edificios blancos en la pared sur. Vimos cómo los construían mientras hacíamos gimnasia.
 —Todavía no se la ha aplicado— dijo el chaplino, por lo menos en esta prisión, 6655321. Él mismo tiene 0 dudas acerca del asunto, y he de confesar que yo las comparto. El problema es saber si esta técnica puede hacer realmente bueno a un hombre. La bondad viene de adentro, 6655321. La bondad es algo que uno elige. Cuando un hombre no puede elegir, deja de ser hombre. (Anthony Burgess, p. 86)

El diálogo citado corresponde a la famosa novela de ciencia ficción de Anthony Burgess *La naranja mecánica*. De marcada controversia (su primera edición es de 1962), fue traspasada a la pantalla grande en 1971.

La trama narra cómo, un joven de quince años, se ofrece en uno de sus ingresos a la correccional a ser sujeto de un experimento que, según ha escuchado en los pasillos, acortaría su condena

La trama narra cómo, un joven de quince años que pasa sus noches cometiendo crímenes y actos de vandalismo como violaciones, robos, golpizas brutales, se ofrece en uno de sus ingresos a la correccional a ser sujeto de un experimento que, según ha escuchado en los pasillos, acortaría su condena. Lo único que Alex, el protagonista, sabe del mismo es que, de esa manera, se encontrará rápidamente en libertad y se asegurará el nunca volver a prisión.

Según el “chaplino” (capellán de la prisión, apodado así por el protagonista), se trata de algo sencillo pero drástico, que no ha pasado la etapa experimental y aún no se ha aplicado, por lo que no se encuentra convencido de sugerirle que se someta a formar parte. Y, principalmente y desde su visión moral, expresa que se trata de una técnica cuyo objetivo es hacer bueno al hombre, pero se cuestiona: ¿es realmente posible “hacer bueno a un hombre”?, y concluye que la bondad es una elección, que la bondad viene de adentro y, sobre todo: que “cuando un hombre no puede elegir, deja de ser hombre”.



El protagonista hace todo lo posible por que lo acepten como sujeto de experimentación, sin saber que el mismo lo sometería a semanas de sufrimiento extremo, y a la pérdida del goce de lo que más valora: la música clásica.

Entonces, se le inyecta una aguja hipodérmica y se le expone a un experimento que podríamos denominar conductista, cuyo fin es la reforma de la conducta a través de asociar cuerpo y psiquis con una fuerte sensación de desagrado al ver o pensar en escenas de violencia. Como efecto paralelo, la música de Beethoven que acompañan las imágenes del horror que se proyectan al personaje hasta el hartazgo, queda asociada a la sensación de repulsión, arruinando su mayor felicidad.

3. The moral machine: una analogía con *La naranja mecánica*

En el artículo *The Moral Machine Experiment* (Awad et al., 2018), se evalúan las maneras en que una máquina debería/podría ser codificada para responder según las leyes morales de los seres humanos, planteando la problemática de que la moral no es única e inamovible, sino que responde a patrones culturales e —incluso— dentro de cada cultura, etnia, civilización, a momentos históricos. Si bien pueden existir ciertas bases compartidas que hacen a la definición del “bien” y el “mal”, estas nunca conformarían una homogeneidad tal que permitiese codificar a un robot de manera unánime sin implicar un desafío y desacuerdos.

Dice Salas (2022) que “una duda aún no resuelta es si el profesional médico debe o no revelar a su paciente que se está usando IA y en qué ámbitos del quehacer clínico”

Coincidimos con que el uso de tecnologías aplicadas al campo médico ha demostrado avances inusitados y de colaboración para las competencias humanas —siguiendo a Salas “ya existe suficiente evidencia que muestra que sistemas de IA son más eficientes en realizar diversos diagnósticos médicos que los mejores expertos en las diversas áreas clínicas” (2022, párr.1)—. Sin embargo, cabe dilucidar cuestiones que atañen a las decisiones de las personas (pacientes o sujetos de investigación) desde la bioética en relación con la aplicación de IA en los diagnósticos, tratamientos, ensayos clínicos.

Por eso, este trabajo comprende que la IA, aplicada al acto médico, incluye diversos dilemas: a. El respeto por el principio de autonomía; b. La pregunta por el ser; c. Los límites éticos en su aplicación.

3.1. Inteligencia artificial y autonomía: el derecho a decidir o “cuando un hombre no puede elegir dejar de ser hombre”

Dice Salas (2022) que “una duda aún no resuelta es si el profesional médico debe o no revelar a su paciente que se está usando IA y en qué ámbitos del quehacer clínico” (Párr. 2). Aquí es donde nos detendremos para el análisis.

Siguiendo a Beauchamp y Childress (2019, p. 99), uno de los cuatro principios éticos a seguir es el del respeto por la autonomía. Según los autores “The autonomous individual



acts freely in accordance with a self-chosen plan (...)” siempre que no se encuentre coaccionado por las circunstancias y sus competencias no se encuentren disminuidas. En el caso de Alex, su autonomía se encuentra disminuida, en primer lugar, porque está encarcelado. Si bien podríamos decir que salir de la cárcel más rápido al someterse al experimento se correspondería con su plan de vida, por lo que estaría actuando de forma autónoma, aún podemos interpretar que las circunstancias desfavorables lo impulsan a tomar tal decisión. Incluso es él quien se ofrece voluntariamente, aún sin que se le haya propuesto participar; de todas formas, esto no exime de responsabilidad a los investigadores que le inyectan una sustancia (podríamos compararlo con aplicar IA en un tratamiento) que genera un cambio en su manera de percibir el mundo, y de actuar sobre él, sin explicar detenidamente los alcances y efectos de la misma.

Si bien que Alex esté en prisión nos enfrenta a un dilema extra, el hecho de insertar cierta tecnología para modificar quien es él implica una conducta totalmente antiética. ¿Desea Alex realmente modificar su esencia, o simplemente desea salir de prisión?

Analizado desde la perspectiva de los autores, retomaremos el problema de la autonomía desde una teoría “no ideal” o un enfoque de “non-ideal conditions” con su propuesta de la “three conditioning theory” (2019, p. 102) que tiene en cuenta los siguientes aspectos: (1) intencionalidad, (2) entendimiento e (3) influencias externas.

Desde esta mirada, la autonomía de Alex se corrompe en varios niveles: el primero, al no brindársele la información necesaria para que decida de forma comprendida (entendimiento); el segundo, al no explicársele el proceso por el que deberá pasar para la consecución del experimento, lo que le quita intencionalidad —según los autores la intencionalidad corresponde a un plan propuesto en forma de eventos, que pueden o no resultar como esperado pero cuyo proceso, al margen de los resultados, ha sido transparentado—; y el tercero, al encontrarse influido por las circunstancias externas (estar encarcelado).

Si bien que Alex esté en prisión nos enfrenta a un dilema extra, el hecho de insertar cierta tecnología para modificar quién es él implica una conducta totalmente antiética. ¿Desea Alex realmente modificar su esencia o simplemente desea salir de prisión?

En esta línea de reflexión podemos preguntarnos qué desea un paciente que accede a un tratamiento o terapia en el que, por ejemplo, se le incruste un chip o una prótesis inteligente, convirtiéndolo en un hombre “postorgánico” o “transhumano” (Sibilia, 2005; Hans, 1995), al trascender su condición de humano para convertirse en un ensamble con algo más (*cyborg*), porque desea mejorar su vida o no perder aspectos significativos de sí mismo, aunque no desee combinar su organismo con IA.

En casos más extremos, ¿qué sucedería si la vida del paciente dependiera de esta modificación, pero sus valores rechazaran la implantación de IA en su organismo, impidiendo salvar su vida? ¿Debería el profesional actuar de forma paternalista? Nos encontraríamos frente a un dilema similar al del testigo de Jehová que rechaza una transfusión de sangre.

Entrando en conflicto el principio de autonomía con el de beneficencia en pos del mejor interés para el paciente, ¿estaría justificada la mentira y la ruptura con la obligación a la veracidad? (Beauchamp & Childress, 2019).



Dilemas de esta magnitud pueden hallarse en temáticas como las de enfermedades hereditarias, en que el profesional se encuentra entre el cumplimiento del secreto profesional y la confidencialidad *versus* beneficiar a terceros al informarlos de posibles daños que pudieran ocurrirles en caso de no tratar la enfermedad (Dove et. al, 2019).

Tanto en el caso de las transfusiones como en el de enfermedades hereditarias, el conflicto se contextualiza en el entramado de la RMP, y las contradicciones entre paternalismo y principio de autonomía, que también se hacen visibles en los avances con IA.

3.2. Reflexiones preliminares. Parte 1

Si bien la tecnología y la IA pueden resolver cuestiones médicas e incluso hacer al progreso de la medicina y el tratamiento o diagnóstico de enfermedades, existen ciertos límites que corresponden al ámbito de la humanidad y que no deberían ser resueltos por inteligencias artificiales.

Pretender que la moral sea programada mediante algoritmos es como pretender que una persona modifique su esencia y actúe por conductismo.

Pretender que la moral sea programada mediante algoritmos es como pretender que una persona modifique su esencia y actúe por conductismo

La moral corresponde al ámbito humano, natural y no artificial, y por ello consideramos que las decisiones tomadas con respecto a la utilización de IA en diagnósticos y tratamientos médicos que afecten a los cuerpos de los pacientes deben ser explicadas, conversadas, comprendidas y consensuadas, sin ocultarse información de ningún tipo, ni ejercerse coerción alguna aunque cumplir con la veracidad pusiera en peligro la concreción de un tratamiento y la vida de la persona. La autonomía debe ser respetada en primer lugar porque “cuando uno no puede elegir, deja de ser hombre”¹.

3.3. Reflexiones preliminares. Parte 2: investigación con seres humanos privados de la libertad desde un caso de ficción *La naranja mecánica*

En esta segunda parte profundizaremos en el tema de la investigación con seres humanos en condiciones de encierro desde el caso de ficción mencionado en el apartado 1 *La naranja mecánica*.

Nos preguntamos por la posibilidad de realizar ensayos clínicos en sujetos privados de la libertad para modificar la conducta mediante IA, atravesado —a su vez— por el interrogante sobre la moral: ¿es el ser humano pasible de ser condicionado para actuar virtuosamente?

1 Respetamos la frase original, si bien entendemos que se encuentra desfasada del contexto actual, en el que el uso de la palabra “hombre” como representante de “personas” o “seres humanos” denota una connotación que no contempla la perspectiva de género.



Durante la novela se narra la desesperación del protagonista por salir del encierro y recuperar su libertad, previamente asociada a la satisfacción de sus pulsiones tanáticas

Como disparador de la reflexión, se retoma el caso ficticio de Alex De Large, protagonista de la novela citada, en que el personaje es el líder de un grupo vandálico, a la vez que aficionado a la obra de Ludwig Van Beethoven; este dato no es menor ya que, en su habilidad para apreciar obras artísticas maestras, encuentra tanto goce como en el acto de destruir, artificio que será empleado en el experimento para alcanzar la reforma psicológica de manera conductista —mediante la inhibición de sus impulsos naturales—.

Dice Burgess que:

[...] por definición, el ser humano está dotado de libre albedrío, y puede elegir entre el bien y el mal. Si sólo puede actuar bien o sólo puede actuar mal, no será más que una naranja mecánica, lo que quiere decir que en apariencia será un hermoso organismo con color y zumo, pero que de hecho no será más que un juguete mecánico al que Dios o el Diablo (o el Todopoderoso Estado, ya que está sustituyéndolos a los dos) le darán cuerda. Es tan inhumano ser totalmente bueno como totalmente malvado. Lo importante es la elección moral. La maldad tiene que existir junto a la bondad para que pueda darse esa elección moral. La vida se sostiene gracias a la enconada oposición de entidades morales. (2005)

Sobre el título aclara:

Las naranjas mecánicas no existen, excepto en el habla de los viejos londinenses. La imagen era extraña, siempre aplicada a cosas extrañas. “Ser más raro que una naranja mecánica” quiere decir que se es extraño hasta el límite de lo extraño. [...] Yo la uso para referirme a la aplicación de una moralidad mecánica a un organismo vivo que rebosa de jugo y dulzura. (2005)

3.4. Reflexiones preliminares. Parte 3: sujeto de experimentación y consentimiento informado

Comencemos por el sujeto de experimentación a quien llamamos por su nombre novelado: Alex.

Alex, en uno de sus actos vandálicos, asesina a una mujer y es encarcelado.

Durante la novela se narra la desesperación del protagonista por salir del encierro y recuperar su libertad, previamente asociada a la satisfacción de sus pulsiones tanáticas. Entre pasillos escucha sobre un tratamiento llamado “Ludovico”, en fase experimental, que podría reformar aquellos impulsos destructivos.

Sin embargo, lo que motiva realmente a Alex a someterse al mismo es su ansia de libertad, el salir de prisión. El personaje no expresa en momento alguno una genuina reflexión sobre su propia moral, ni la conciencia del deber como noción de virtud moral



en función de reformar sus impulsos para a futuro actuar de acuerdo con el deber, más allá de sus inclinaciones.

Durante la conversación, el capellán advierte a Alex que se trata de un experimento en revisión, que no ha sido aún puesto a prueba, al que no considera apropiado

Una vez en oídos de Alex, sucede una conversación con el capellán de la institución, quien ha acompañado siempre al protagonista motivando su redención (ver parte 1). Sin embargo, Alex nunca busca la redención, sino la ventaja. Se acerca al capellán como informante con el interés de obtener beneficios a cambio.

Durante la conversación, el capellán advierte a Alex que se trata de un experimento en revisión, que no ha sido aún puesto a prueba, al que no considera apropiado.

Más allá de las advertencias, Alex decide someterse al tratamiento, incluso sin saber en qué consiste. La única información en su poder es que convierte a la persona en un ser “de bien”, evitando que vuelva a prisión, pero no se explica mediante qué herramientas.

Más avanzada la obra, se revela el procedimiento: un experimento de corte conductista en el que la psiquis es reformada a través de la tortura. Las imágenes tortuosas (“ultravioletas”) que le son impuestas y que le obligan a ver, aun cuando pide que se detengan, son acompañadas de su música predilecta, lo que unirá por asociación la repulsión por la violencia a la música clásica.

4. De las reflexiones parte 1, 2 y 3: aspectos éticos en la investigación con seres humanos y Declaración de Helsinki

En lo anteriormente reflexionado, aparecen diversas situaciones de carácter antiético.

Por un lado, no existe consentimiento informado (CI) apropiado ni comprendido.

En términos de De Ortúzar (2021) el CI tiene por objeto y responsabilidad ética el respeto de la persona como agente moral independiente, teniendo en cuenta que la misma sea una persona autónoma y capaz, y que el proceso esté acompañado de la transmisión de información plena, deliberada, en el que los profesionales se cercioren de la comprensión del experimento a la plena conciencia por parte del sujeto sobre la naturaleza del mismo, las consecuencias y los riesgos.

Beauchamp y Childress (2019) sostienen que algunas teorías se basan en la evaluación de “capacidades” (autogobierno, entendimiento, deliberación, manejo, elección independiente) para la definición de la autonomía, pero enfatizan su preferencia por un enfoque en “elecciones autónomas” más que capacidades generalizables, objetando que, incluso personas en condiciones saludables pueden encontrar obstáculos hacia la concreción de la elección más beneficiosa para sí misma o en concordancia con sus deseos reales y plan de vida a consecuencia de situaciones temporales como el padecimiento de una enfermedad, depresión, ignorancia, y cualquier condición que limite su juicio/opiniones. Entre estas, y a las que a este análisis concierne, mencionan la situación de coerción.



La privación de la libertad actuaría como coerción para un presidiario al que se le asegura la libertad inmediata a cambio de su participación en un ensayo clínico experimental.

Como dicta la Declaración de Helsinki:

El propósito principal de la investigación médica en seres humanos es comprender las causas, evolución y efectos de las enfermedades y mejorar las intervenciones preventivas, diagnósticas y terapéuticas (métodos, procedimientos y tratamientos). Incluso, las mejores intervenciones probadas deben ser evaluadas continuamente a través de la investigación para que sean seguras, eficaces, efectivas, accesibles y de calidad. (AMM, 2013)

La privación de la libertad actuaría como coerción para un presidiario al que se le asegura la libertad inmediata a cambio de su participación en un ensayo clínico experimental

Tomando este documento como referencia, investigar las causas por las que un sujeto tiene impulsos dañinos o perversos hacia otros sujetos sería quizás un objeto pertinente de estudio, mas no así experimentar el desarrollo de un método que modifique la moral (el actuar, en realidad) de las personas. Los sujetos competentes no deben ser manipulados en su actuar mediante la imposición de un tratamiento, sino que deben gozar libremente de la responsabilidad que implica ser un sujeto social, emocional, racional y moral, y responder por las consecuencias de sus actos.

5. Análisis del procedimiento desde las teorías kantiana y utilitarista

John Stuart Mill sostiene que

El credo que acepta la Utilidad o Principio de la Mayor Felicidad como fundamento de la moral, sostiene que las acciones son justas en la proporción con que tienden a promover la felicidad; e injustas en cuanto tienden a producir lo contrario de la felicidad. Se entiende por felicidad el placer, y la ausencia de dolor; por infelicidad, el dolor y la ausencia de placer... (2004, cap. II)

Desde el utilitarismo, experimentar con presos, si esto les permitiera reformarse y alcanzar la libertad inmediata, maximizaría la felicidad de los sujetos en condiciones de encierro al acceder al placer de la libertad a cambio; sin embargo, una técnica como "Ludovico", sería a la vez injusta, ya que la misma se acompaña de sufrimiento y dolor, así como la limitación del placer de acciones saludables como escuchar música.

A la vez sería dilemática ya que rompería el contrato social de cumplir una condena por cometer actos en perjuicio de terceros, lo que generaría descontento e infelicidad general para la sociedad, contradiciendo el principio de la mayor felicidad para el mayor número de personas.



En términos de Kant: “Ni en el mundo, ni, en general, tampoco fuera del mundo, es posible pensar nada que pueda considerarse como bueno sin restricción, a no ser tan sólo una buena voluntad”

Por otro lado, coincidimos con la ética kantiana en la que ningún sujeto es un medio para un fin, sino que un fin en sí mismo, por lo que nadie debe ser expuesto a una práctica de experimentación bajo condiciones en que la autonomía se encuentra vulnerada, así como se brinda información incompleta y no se sigue un adecuado procedimiento para la comprensión y aceptación del CI.

Incluso si el experimento diera resultados efectivos, la implementación de la técnica sería en sí misma antiética; nos lleva a cuestionarnos por la verdadera intención y el valor del ser: la posible reforma de conciencia de ese sujeto es aniquilada al so-

meterlo a un experimento de corte conductista en el que su moral no es conscientemente reformulada, sino que el cambio en sus acciones se da en respuesta a coartar su libertad, su autonomía, mediante estímulos negativos (dolor).

En términos de Kant: “Ni en el mundo, ni, en general, tampoco fuera del mundo, es posible pensar nada que pueda considerarse como bueno sin restricción, a no ser tan sólo una buena voluntad”:

La buena voluntad no es buena por lo que efectúe o realice, no es buena por su adecuación para alcanzar algún fin que nos hayamos propuesto; es buena sólo por el querer, es decir, es buena en sí misma. Considerada por sí misma, es, sin comparación, muchísimo más valiosa que todo lo que por medio de ella pudiéramos verificar en provecho o gracia de alguna inclinación y, si se quiere, de la suma de todas las inclinaciones. Aun cuando, por particulares enconos del azar o por la mezquindad de una naturaleza madrastra, le faltase por completo a esa voluntad la facultad de sacar adelante su propósito; si, a pesar de sus mayores esfuerzos, no pudiera llevar a cabo nada y sólo quedase la buena voluntad —no desde luego como un mero deseo, sino como el acopio de todos los medios que están en nuestro poder—, sería esa buena voluntad como una joya brillante por sí misma, como algo que en sí mismo posee su pleno valor. La utilidad o la esterilidad no pueden ni añadir ni quitar nada a ese valor. (1999, Cap. I)

Desde esta concepción, el sujeto en cuestión estaría perdiendo su libertad como ser humano, su autonomía y su dignidad.

Incluso, aunque pensáramos desde el utilitarismo que someter a los sujetos que cometen delitos y acciones inmorales a experimentos o tratamientos como este podría resultar en el aumento del bienestar general porque las calles y la sociedad serían más seguras, maximizando el nivel general de felicidad, aún podemos aferrarnos a la pregunta por la máxima kantiana “¿puedes creer que tu máxima se convierta en ley universal?”. Y a esto responder que no porque, aunque sostuviésemos la lógica utilitarista, la misma podría fallar en el caso de que un inocente fuera condenado de forma errada y sometido al tratamiento en cuestión cuando no fuera un criminal y no constituyese una amenaza para la sociedad, derribando así la idea de un mayor bienestar real a través del procedimiento. Y porque el bienestar de una sociedad no



debe basarse en una ética retaliativa en la que se vulnere la dignidad del ser humano como forma de hacer justicia, más allá de sus acciones.

Como sostiene la Comisión Interamericana de Derechos Humanos en el documento *Principios y buenas prácticas sobre la protección de las personas privadas de libertad en las Américas*:

Se les protegerá contra todo tipo de amenazas y actos de tortura, ejecución, desaparición forzada, tratos o penas crueles, inhumanos o degradantes, violencia sexual, castigos corporales, castigos colectivos, intervención forzada o tratamiento coercitivo, métodos que tengan como finalidad anular la personalidad o disminuir la capacidad física o mental de la persona. (Principio I, 2008)

A la vez que reconoce como derecho fundamental de toda persona privada de libertad a ser tratada humanamente, el respeto por y la garantía de su dignidad, su vida, integridad física, psicológica y moral y se explicita el no aplicar tratamientos que anulen la personalidad o disminuyan sus capacidades, que es lo que el tratamiento “Ludovico” haría en Alex.

Y porque el bienestar de una sociedad no debe basarse en una ética retaliativa en la que se vulnere la dignidad del ser humano como forma de hacer justicia, más allá de sus acciones

Además, destaca que las personas privadas de la libertad se encuentran en una situación de vulnerabilidad y que la finalidad esencial es “la reforma, la readaptación social y la rehabilitación personal de los condenados; la resocialización y reintegración familiar; así como la protección de las víctimas y de la sociedad” (Párr. 2).

Desde esta normativa, queda expreso que se comparte, a nivel internacional, la noción de una reforma y readaptación, así como rehabilitación personal como sinónimo de virtud moral, y no como síntoma de la coerción por inhibición de la capacidad de decidir.

6. Análisis desde la metáfora de las capas y las reglas de Mandela

“Las Reglas de Mandela son recomendaciones que Naciones Unidas hace a los Estados para tratar a los reclusos y administrar las cárceles” (Ministerio de Justicia y Derechos Humanos Argentina, <https://www.argentina.gob.ar/noticias/que-son-las-reglas-mandela>).

Estas reglas tienen particularmente en cuenta las distintas condiciones de los reclusos. No todos los reclusos serán iguales ni tendrán las mismas necesidades, habrá aquellos que, desde la metáfora de las capas (Luna, 2009), acumulen más capas de vulnerabilidad que otros.

Sin embargo, el hecho del encierro comprende, en sí mismo, una capa de vulnerabilidad como base.



La Oficina de las Naciones Unidas contra la Droga y el Delito en el documento Reglas Mínimas de las Naciones Unidas para el Tratamiento de los Reclusos (Reglas Nelson Mandela) expresa que:

La prisión y demás medidas cuyo efecto es separar a una persona del mundo exterior son aflictivas por el hecho mismo de que despojan a esa persona de su derecho a la autodeterminación al privarla de su libertad. Por lo tanto, a excepción de las medidas de separación justificadas y de las que sean necesarias para el mantenimiento de la disciplina, el sistema penitenciario no deberá agravar los sufrimientos inherentes a tal situación. (2015, Regla 3)

Desde esta regla, Alex no debería ser sometido a un ensayo experimental de ningún tipo ya que, encontrarse despojado del derecho a la autodeterminación implica una capa

No todos los reclusos serán iguales ni tendrán las mismas necesidades, habrá aquellos que, desde la metáfora de las capas (Luna, 2009), acumulen más capas de vulnerabilidad que otros

de vulnerabilidad que lo posicionará en una situación de abuso; su poder de consentir se ve coaccionado por el contexto y por la asimetría de poder (Foucault, 2003) en la que se encuentra inmerso a la vez que el ensayo es, en sí mismo, discriminatorio.

Podemos preguntarnos qué sucede, entonces, con el principio de justicia (Beauchamp y Childress, 2019), que indica que ningún grupo debe ser discriminado por sus condiciones particulares y apartado de la posibilidad de participar como sujeto de investigación; sin embargo, teniendo en cuenta que no se trata de una experimentación con fines terapéuticos a extenderse al común de la sociedad sin distinción, sino que, por el contrario, se discrimina específicamente para ser aplicada al grupo de

reclusos con la finalidad de modificar su conducta de forma forzosa, el principio de justicia no aplica.

7. Conclusiones: una pregunta sobre la moral y los derechos humanos

Al reflexionar sobre el principio de autonomía, pensamos que, en el caso reflexionado y analizado, la autodeterminación resulta afectada en los siguientes niveles:

1. Consentimiento informado: está en falta porque no se le entrega una descripción de la metodología a emplear, ni la explicación de la consistencia del tratamiento; solo se le comenta superficialmente que el mismo logrará efectos positivos en su actuar ya que mejorará sus impulsos y por ende su conducta será buena a los fines sociales.
2. Moral y autonomía: el móvil principal es alcanzar la libertad. Alex actuará bajo coerción, lo que restringe su autonomía. A su vez, no decidirá ni desde un móvil virtuoso, ni desde una decisión racional que autolegisle su actuar (Kant, 1999), sino por un fin egoísta: la libertad y evasión de la condena.



3. Moral y dignidad: se verá afectada no solo a los fines de la decisión de participar del ensayo, sino que quedará afectada una vez concluido el mismo, ya que implicará la modificación de sus decisiones como ser humano no por decisión, sino por incapacidad: se le incapacita para cometer acciones criminales, dañinas a terceros, a través de la tortura psicológica. La reforma en la conducta del sujeto no se logra a conciencia, no será una moral autolegislada, sino que responderá a los estímulos negativos insertos de forma cruel en su sistema.

Por lo cual, su autonomía y su condición de ser humano en que decide libremente cómo actuar, su oportunidad de actuar virtuosa y moralmente, le es arrebatada.

Es también importante mencionar que este tipo de procedimientos podría, por pendiente resbaladiza, derivar en la justificación de atrocidades como las sucedidas en campos de concentración durante la Segunda Guerra Mundial en que la ideología nacionalsocialista de una “raza superior” sirvió de argumento para la privación de

la libertad de millones de ciudadanos para ser sometidos a trabajos forzados, inanición y experimentación en ensayos clínicos que condujeron a daños irreversibles y la muerte, y que dieron origen a la protección de los derechos humanos a través del Código de Nuremberg, antecedente de los códigos en ética de la investigación con seres humanos como la Declaración de Helsinki y el Informe Belmont.

Consideramos que ninguna persona es un medio para un fin (Kant, 1999), y que una legislación ética no concibe que los conflictos sociales sean resueltos mediante la coerción o modificación impuesta a la libertad y decisión de las personas, más allá de su accionar.

Hemos retomado el caso extremo de personas privadas de la libertad, de la falta de moral individual y de la falta de ética profesional en investigación a través de un caso de ficción. El mismo nos permite exponer la responsabilidad de los Estados

de educar a través de la divulgación científica sobre temas de IA aplicada a la salud como un derecho, así como inculcar que decidir de forma crítica y autónoma forma parte de la esencia humana. Por ello, cuando el contexto inhibe o altera la significación y las consecuencias de someterse a un experimento, manipula creencias, desvirtúa definiciones o no da lugar a un CI comprendido, las decisiones dejan de ser genuinamente autónomas, posibilitando que se abuse del poder científico-económico y político en diversos aspectos de la realidad, entre ellos, la “panacea” de un hombre poshumano que —a través de mecanismos artificiales— se convierta en lo que un grupo de personas considere como la “superación” de la naturaleza humana. Comunicar de forma genuina y lograr una elección comprendida de lo propuesto por parte de los sujetos a ser intervenidos, implica un esfuerzo desde cada ámbito involucrado en los procesos de toma de decisiones en salud, para la concreción de una relación médica, científica y política respetuosa de la dignidad humana.

Consideramos que ninguna persona es un medio para un fin (Kant, 1999), y que una legislación ética no concibe que los conflictos sociales sean resueltos mediante la coerción o modificación impuesta a la libertad y decisión de las personas, más allá de su accionar



Referencias

- Asociación Médica Mundial. (2013). *Declaración de Helsinki*. <https://www.wma.net/es/policias-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/>
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J. F. y Rahwan, I. (2018). The Moral Machine experiment. *Nature*, 563, 59-64. <https://doi.org/10.1038/s41586-018-0637-6>
- Beauchamp, T. y Childress, F. (2019). *Principles of Biomedical Ethics* (8.ª ed.). Oxford University Press.
- Burgess, A. (2005). *La naranja mecánica*. Minotauro
- Comisión Interamericana de Derechos Humanos. (2008). *Principios y Buenas Prácticas sobre la Protección de las Personas Privadas de Libertad en las Américas*. <https://www.oas.org/es/CIDH/jsForm/?File=/es/cidh/mandato/basicos/principiosspl.asp>
- Dove, E. S., Chico, V., Fay, M., Laurie, G., Lucassen, A. M. y Postan, E. (2019). Familial genetic risks: how can we better navigate patient confidentiality and appropriate risk disclosure to relatives? *Journal of Medical Ethics*, 45(8), 504-507. <https://doi.org/10.1136/medethics-2018-105229>
- Foucault, M. (2003). *Vigilar y Castigar: Nacimiento de la prisión*. <https://www.ivanillich.org.mx/Foucault-Castigar.pdf>
- Hans, J. (1995). *El principio de responsabilidad. Ensayo de una ética para la civilización tecnológica*. Herder.
- Kant, I. (1999). *Fundamentación de la metafísica de las costumbres*. Biblioteca virtual Miguel de Cervantes. <https://www.cervantesvirtual.com/nd/ark:/59851/bmcg44m4>
- Luna, F. (2009). Elucidating the Concept of Vulnerability: Layers Not Labels. *International Journal of Feminist Approaches to Bioethics*, 2(1), 121-139. <https://doi.org/10.3138/ijfab.2.1.121>
- Mill, J. S. (2004). *El Utilitarismo*. Biblioteca Virtual Antorcha. http://www.antorcha.net/biblioteca_virtual/filosofia/utilitarismo/indice.html
- Ministerio de Justicia y Derechos Humanos Argentina. (2018, 02 de febrero). *Qué son las reglas Mandela*. <https://www.argentina.gob.ar/noticias/que-son-las-reglas-mandela>
- Oficina de las Naciones Unidas contra la Droga y el Delito. (2015). *Reglas mínimas de las naciones unidas para el tratamiento de los reclusos (Reglas Nelson Mandela)*. https://www.unodc.org/documents/justice-and-prison-reform/Nelson_Mandela_Rules-S-ebook.pdf
- Ortúzar, de. (2021). Clase 2. *El concepto de autonomía y su aplicación a la práctica sanitaria a través del "consentimiento informado"*. *Fundamentación teórica y revisión crítica*. Especialización en bioética. FLACSO.
- Salas, S. (2022). *Consideraciones éticas respecto de la inteligencia artificial*. Bioeticar Asociación Civil.
- Sibilia, P. (2005). *El hombre postorgánico. Cuerpo, subjetividad y tecnologías digitales*. Fondo de Cultura Económica de Argentina S. A.