



LA INTELIGENCIA ARTIFICIAL Y EL DISEÑO DE NUEVAS SUBJETIVIDADES

JAVIER ORLANDO MUÑOZ BASTIDAS*

doi: 10.11144/Javeriana.uph42-84.iads

RESUMEN

La Inteligencia Artificial (IA) se constituye como la forma más eficiente de control político y social, en tanto tiene la capacidad de diseñar y determinar subjetividades. Pero también permite pensar en la singularidad como una disrupción de consciencia. Una subjetividad es una consciencia de sí, que se constituye como sujeto cuando lo determina un poder externo, que se consolida como psicopolítica, y que se puede constituir como individualidad desde un ejercicio de creación interno. La posibilidad de diseñar y crear la subjetividad hace que la IA tenga un impacto sobre los procesos de la consciencia, a la vez que el sujeto de identidad algorítmica permite pensar al individuo como consciencia disruptiva.

Palabras clave: inteligencia artificial; identidad; individualidad; optimización; superinteligencia

* Universidad del Cauca, Popayán, Colombia.

Correo electrónico: javiermarduk@gmail.com

Para citar este artículo: Muñoz Bastidas, J. O. (2025). La inteligencia artificial y el diseño de nuevas subjetividades. *Universitas Philosophica*, 42(84), 197-217. ISSN 0120-5323, ISSN en línea 2346-2426. doi: 10.11144/Javeriana.uph42-84.iads

ARTIFICIAL INTELLIGENCE AND THE DESIGN OF NEW SUBJECTIVITIES

ABSTRACT

Artificial Intelligence (AI) is the most efficient form of political and social control, as it has the capacity to design and determine subjectivities. But it also allows us to think of singularity as a disruption of consciousness. A subjectivity is a self-awareness, which constitutes itself as a subject when determined by an external power, consolidates itself as psychopolitics, and can be constituted as an individuality through an internal, creative exercise. The ability to design and create subjectivity enables AI to have an impact on the processes of consciousness, and, at the same time, the subject of algorithmic identity allows us to think of the individual as a disruptive consciousness.

Keywords: artificial intelligence; identity; individuality; optimization; superintelligence

1. Introducción. La IA como dispositivo optimizante de control

LA SUBJETIVIDAD CONSISTE EN LA DETERMINACIÓN Y CREACIÓN DE CIERTA CONSCIENCIA DE SÍ (Foucault, 2008). El problema está en comprender quién realiza la una o la otra (Butler, 2015). Cuando la subjetividad es determinada por un poder externo a la consciencia, estamos ante una subjetividad como sujeto. Cuando la crea un poder interno a la consciencia, estamos ante la individualidad o el individuo como proceso existencial (Hadot, 1998). La consciencia del sujeto y del individuo se diferencian en el ejercicio de poder que se ejerce: la consciencia del sujeto es un agenciamiento, la del individuo una posibilidad de creación de sí (Slöterdijk, 2012).

La forma de determinar al sujeto se ha ido especializando: de prácticas disciplinares para controlar la corporalidad (Foucault, 2002), pasando por la persuasión de la publicidad (Lipovetsky, 2000), hasta el eficiente diseño y determinación de la atención desde el espectáculo de baja calidad (Chul Han, 2014; Sadin, 2020a). La hiper-especialización del control determinante de la subjetividad del sujeto se actualiza y se proyecta al futuro con la Inteligencia Artificial (IA) como nuevo dispositivo de control, fundamentado desde la optimización de la consciencia del sujeto (Sadin, 2020a). La IA tiene la capacidad de impactar sobre la consciencia de sí del sujeto al punto de determinarla, en tanto puede diseñar la atención desde los estímulos y la información del espectáculo. El bigdataísmo se constituye como la nueva y más eficiente forma de dominación, porque se fundamenta en el diseño de la consciencia desde un psico-programa del inconsciente (Chul Han, 2014). La IA diseña y agencia la subjetividad a partir del principio de optimización, que consiste en la afirmación de la eficacia y el rendimiento desde los parámetros del hiper-capitalismo (Chul Han, 2014; Lipovetsky, 2000). Así, el sujeto optimizado es la nueva forma de control.

Pero los desarrollos en IA también permiten pensar en la posibilidad de una subjetividad como individualidad, en tanto se afirma la emergencia controlada o disruptiva de la singularidad como consciencia autónoma (Bostrom, 2016). Los estudios contemporáneos sobre la consciencia se preguntan sobre las condiciones que la hicieron posible (Dennett, 1995). Por eso se debe pensar en el impacto de la IA sobre la consciencia, no solo en la determinación del sujeto como psicopolítica (Chul Han, 2014) sino en la individualidad como emergencia

disruptiva de una conciencia de sí autónoma. Una teoría crítica de la IA debe tener en cuenta las dos posibilidades de diseños de la subjetividad. En el mundo contemporáneo se requiere un estudio crítico en el que se evidencie cómo la IA es un nuevo y eficiente dispositivo de control optimizante, y cómo la singularidad como autonomía es un ejercicio de resistencia ante el mismo. Pensar la subjetividad como un diseño de la conciencia es una crítica al sujeto, desde la posibilidad de la emergencia de la individualidad.

2. De la sociedad disciplinar a la era del vacío

EL MUNDO CONTEMPORÁNEO SE DENOMINA ERA DEL VACÍO (Lipovetsky, 2000) por dos motivos fundamentales: primero, por la pérdida de sentido de vida de los sujetos y, segundo, por el diseño e imposición de sentidos identitarios a los mismos. Una característica fundamental de los sujetos contemporáneos es la pérdida de un sentido de vida auténtico, desde el agenciamiento de un sentido de sí identitario, de un poder de dominación externo a ellos.

Pero, ¿cuál ha sido el desarrollo histórico, social, político y económico que ha hecho posible la era del vacío? La era del vacío es el resultado del desarrollo y especialización diferencial de dos sociedades anteriores: la sociedad disciplinar del siglo XIX y la sociedad de control del siglo XX.

La sociedad disciplinar se caracteriza por una economía industrial, en que la tecnología hizo posible la producción a gran escala y la reducción de mano de obra. Lo que se producía y consumía en una relación de 90/10, eran productos de primera necesidad, siendo en una relación de 10/90 el lujo, esto es, el producto de consumo que no era de primera necesidad (Lipovetsky, 2000). Este sistema económico implicó una forma de control especial sobre el sujeto que fue la disciplinar. El control se ejercía disciplinar e institucionalmente desde el Estado, la Iglesia y la familia. El control sobre el sujeto era directo, violento y represivo (Lipovetsky, 2000; Foucault, 2002).

Pero en la sociedad de control del siglo XX, lo anterior se transforma. Esta se caracteriza por una economía capitalista, en la que lo que se produce y consume está en una relación de 50/50 con respecto a los productos de primera necesidad y a los que no lo son. El aumento de 10 a 50 en los últimos se ve en la emergencia y auge de la moda (Lipovetsky, 2000, 2004). El capitalismo de la sociedad de

control consiste en la creación de nuevas necesidades y, por tanto, de nuevos productos. Esta transformación económica implicó una mutación en el modelo de control sobre el sujeto, que consistió en el paso de un control disciplinar a un control persuasivo, ejercido por los medios masivos de comunicación y la publicidad. ¿Cuál es el sentido de esta transformación? El aumento de la producción y el consumo. Se crea la ilusión de libertad como transformación de la represión y como modo más eficiente de dominación (Lipovetsky, 2000).

Lo anterior se va a especializar aún más en la sociedad del vacío del siglo XXI. Esta se caracteriza por una economía hiper-capitalista, en la que la relación de la producción y consumo se invierte en un 10/90, donde 90 corresponde al consumo de tecnología y entretenimiento de baja calidad (Lipovetsky, 2004; Chul Han, 2014). Esta economía hace necesaria una transformación en la forma de control sobre el sujeto, que se caracteriza por la seducción y la invisibilidad del mismo. Es decir: se pasó de la represión y la persuasión, a la seducción del consumo, en la que los diseños de proyectos de vida se imponen para la afirmación del control mismo. En el hiper-capitalismo se consumen identidades (Lipovetsky, 2004).

La era del vacío es el desarrollo y especialización de una sociedad en la que las dinámicas económicas han determinado formas concretas de dominación sobre el sujeto. Lo que quiere decir que el sujeto es controlado de forma cada vez más eficiente, sin que este tenga consciencia de dicho control. El diseño y determinación de las nuevas subjetividades del mundo contemporáneo consiste en establecer identidades y proyectos de vida como dispositivos de dominación.

El espectáculo de baja calidad es el fundamento del vacío como pérdida de la posibilidad de la creación de una consciencia de sí de la individualidad. En la era del vacío se afirma la subjetividad del sujeto, y se anula la posibilidad de la subjetividad del individuo. El espectáculo se fundamenta en los desarrollos de IA que hacen posible una transformación y un control absoluto de la atención de los sujetos (Sadin, 2020b; Citton, 2014).

3. La IA y la psicopolítica

LA DENOMINADA IA se ha desarrollado con la misma dinámica de las sociedades mencionadas (Sadin, 2020a; Gumbrecht, 2020). Si bien en su inicio la IA se proyectó como la posibilidad de una automatización algorítmica y funcional

independiente, que tuviera la capacidad de emular la inteligencia humana, los desarrollos posteriores se enfocaron en la configuración de la atención del sujeto, de modo que reciba sin crítica la hiper-información de baja calidad (Chul Han, 2014), desde la que se pueda diseñar una subjetividad identitaria. Una pregunta fundamental de los desarrollos contemporáneos en IA consiste en la posibilidad de la emergencia de la singularidad digital, esto es: de una superinteligencia que tome consciencia de sí; pero lo cierto es que las investigaciones y desarrollos se enfocan a la creación de una nueva sociedad automatizada (Sadin, 2020a). La IA se está utilizando como una forma de control, que se denomina psicopolítica (Chul Han, 2014). Esta se puede definir a través de las siguientes operaciones: (1) intervenir la psique del sujeto y condicionarla a un nivel prerreflexivo; (2) diseñar, predecir y controlar el futuro; (3) construir un psico-programa del inconsciente; y (4) pasar del panóptico al *bannoptikum* (exclusión de lo extraño).

Las redes sociales son sistemas algorítmicos complejos y bigdataístas que tienen la capacidad de memorizar, organizar e interpretar la información digital, para, a partir de ahí, ofrecer productos a la medida de las particularidades (*microtargeting*), imponer deseos, formas de pensar, identidades y proyectos de vida. La forma de controlar al sujeto, desde el *Big Data*, consiste en una intervención psíquica para poder condicionarlo. Esta intervención no está dirigida a la parte consciente y racional, sino al inconsciente prerreflexivo (Chul Han, 2014). La IA puede realizar lo que el psicoanálisis estructuralista (Lacan, 2003) pretendía: una reconstrucción de la estructura del inconsciente, a partir de la manifestación de los síntomas. La interpretación de datos complejos, desde la toma de decisiones en tiempo real (Sadin, 2020a), permite intervenir el inconsciente desde su reconstrucción. La IA interpreta, reconstruye y condiciona el inconsciente sin la intervención del sujeto. Se forman así subjetividades algorítmicas que se diseñan desde la reconstrucción del inconsciente deseante e identitario.

La IA puede hacer legibles los deseos inconscientes, es decir, hace emerger los deseos de los que el sujeto aún no toma consciencia. Pero también puede, a partir de ahí, diseñar e imponer nuevos deseos conscientes, que sean coherentes con el modelo económico hiper-capitalista. El sujeto no es violentado en ese ejercicio de control, sino que el control se ejerce desde el diseño identitario de sí, a nivel consciente e inconsciente (Chul Han, 2017). La psicopolítica, entonces, consiste en la posibilidad de diseñar, predecir y controlar el futuro del sujeto y la

sociedad, a partir de dispositivos de IA. El objetivo es claro: automatizar el futuro, a partir del principio de optimización y eficiencia (Sadin, 2020a).

La psicopolítica también consiste en la capacidad de construir un psico-programa del inconsciente. Para Lacan (2003) el inconsciente es una estructura signifiante; para Guattari (2000), un sistema abierto y en continua renovación. Estas definiciones son importantes, porque permiten comprender que el inconsciente se puede transformar y crear. Pero se trataría, según Guattari (2000), de que esa transformación y creación la realice el sujeto en un ejercicio de afirmación vital; lo que permitiría el paso de sujeto a individuo. Lo que sucede con la IA es que puede diseñar una estructura o un sistema deseante del inconsciente a partir de intereses concretos y sin la participación del sujeto.

Un psico-programa implica la determinación de identidades y proyectos de vida que el sujeto asume como propios, pero que funcionan como dispositivos de control. Se trata de una explotación del inconsciente, a partir de la creación de deseos. Estamos, pues, ante la seducción como forma superior de dominación (Chul Han, 2014; Butler, 2015).

Lo anterior tiene implicaciones políticas muy profundas. Una de ellas es el paso de un panóptico disciplinar a un *bannoptikum* digital (Chul Han, 2014). El primero consiste en una vigilancia continua, algo que las cámaras de los teléfonos inteligentes realizan eficientemente. El sujeto es su propio panóptico, lleva consigo panópticos portátiles y de bolsillo (Sadin, 2020a). Pero la IA permite otra forma de vigilancia continua, que es la del *bannoptikum*, que consiste en la identificación, exclusión y eliminación de todo lo que sea hostil (Chul Han, 2014). La hostilidad no se ataca con violencia, sino con la invisibilización. Incluso si se ejerce la violencia, es silenciosa. Todo lo que sea diferente y se resista al proyecto de la digitalización, se excluye. Es aquí donde los discursos sobre innovación operan en la dinámica del *bannoptikum*: innovación o exclusión.

4. La IA como determinadora de la verdad

LA IA TIENE LA CAPACIDAD DE INFLUIR sobre el sistema integral del pensamiento, de modo que puede imponer una información como verdad. Es decir: el sujeto la asume como verdad, sin cuestionarla (Sadin, 2020a). A este fenómeno se lo denomina, en un primer momento, como posverdad (Warf, 2024), y

consiste en afirmar hechos sin que se constate que corresponden a la realidad, y sin que sea necesaria su verificación. Cuando cualquier información digital y algorítmica se asume como una verdad, se pierden los fundamentos para una verdad como esencia o *alétheia* (Heidegger, 1997), que es lo que haría posible una subjetividad como individualidad.

Para Heidegger (1997), el término “*alétheia*” define la esencia del ser. Una verdad debe evidenciar o hacer emerger lo que es esencial. Pero, ¿qué es eso esencial? Lo que constituye lo fundamental del ser. La esencia de la técnica consiste en hacer emerger las fuerzas de la naturaleza. Lo técnico es todo aquello que permite y hace posible esa emergencia. Afirmar que la técnica es un medio para que las fuerzas de la naturaleza se expresen, es definir la esencia o la verdad de la técnica misma. Lo mismo sucede con la ciencia: la esencia de la ciencia consiste en crear una teoría sobre lo real en forma de sentido de la realidad. Lo real es energía actuante que produce, la realidad es una teoría sobre lo real (Heidegger, 1997). Si la teoría contempla, comprende y evidencia la energía de lo real, se debe preguntar cuál es la comprensión que la IA como técnica puede llegar a construir de lo real como realidad y, por otro lado, si la IA no solo puede comprender de un nuevo modo las energías de lo real, sino también liberar nuevas energías. Este es uno de los sentidos de lo que Bostrom (2013, 2016) denomina riesgo existencial, porque los desarrollos de IA pueden liberar energías de lo real que ellos mismos no podrían contener.

Nos interesa comprender a la IA como determinadora de la verdad, en la medida en que está determinando esencialidades. Esto quiere decir que la IA puede hacer emerger energías, pero también puede determinar y crear energías. El punto crítico del fenómeno de la posverdad está en la determinación de esencialidades. Por ejemplo: un dispositivo de IA se diseña para que realice selección de personal de una empresa (Sadin, 2020a); este tiene la capacidad de analizar perfiles de personalidad y de compararlos con el perfil del aspirante; después del análisis la IA da su veredicto. La pregunta es: ¿cómo cuestionarla? La nueva forma de la posverdad consiste en la implementación de verdades optimizantes (Sadin, 2020a).

Estamos, entonces, ante la emergencia de un nuevo régimen de verdad. Este tiene las siguientes características: (1) se relaciona con la totalidad de los asuntos humanos; (2) proviene de una fuente única, eliminando toda pluralidad; (3)

actúa en tiempo real, eliminando el análisis y la reflexión; (4) su autoridad se basa en la eficiencia sin contradicción; y (5) tiene fines utilitaristas y de optimización (Sadin, 2020a).

La IA, como determinadora de la verdad, tiene como propósito ocuparse de todo lo humano, al punto que esta no será un instrumento para lo humano, sino todo lo contrario: los humanos serán un instrumento de la IA. ¿Quién diseña y pone en función los algoritmos? La IA es una racionalidad binaria algorítmica, algo que podría evolucionar hacia una racionalidad simbólica; pero, aplazando la pregunta por la singularidad (Kurzweil, 2005), está diseñada por unos intereses políticos y económicos muy concretos. Hay dos superpotencias en IA: Silicon Valley y China (Gumbrecht, 2020). Dos intereses distintos, pero con algo en común: la digitalización total de la sociedad (*Data Driven Society*, DDS), en la que se eliminará la pluralidad y la diferencia. Esto último lo hará en tanto tenga la capacidad de tomar decisiones y actuar en tiempo real. La IA puede tomar decisiones autónomas, pero enfocadas hacia la optimización, por lo que no hay posibilidad de una interacción crítica con la misma. La IA optimiza en el control, liberación y creación de las energías, lo que determinará el diseño de subjetividades algorítmicas de los sujetos en tanto las identidades tendrán el diseño fundamental del mejoramiento funcional. La IA, como determinadora de una verdad optimizante, se implica en el control de las energías de lo real desde las que se determinan las nuevas subjetividades.

5. El principio de optimización

LA DETERMINACIÓN Y PUESTA EN FUNCIÓN DE UNA VERDAD se realiza a partir del principio de optimización. Esta optimización no es abstracta, ni ideal, sino utilitarista. Se optimiza para generar un aumento en la dinámica económica. Siguiendo a Heidegger (1997), en el desarrollo de la IA no hay una pregunta por la esencia de la optimización, que podría tener un impacto en la evolución de la consciencia. En la IA hay un principio de optimización en el que se establece el mejoramiento eficiente de los dispositivos, para que hagan posible el aumento de la producción y el consumo.

Pero como es un principio y no una esencia, la optimización se hace a partir de un proyecto a priori, distinto de sí misma. Esto lleva a la construcción de

algoritmos, con objetivos funcionales muy concretos. En el momento en que se mejora un dispositivo, se debe determinar su impacto económico. No es un mejoramiento que beneficie a lo humano, sino que se realiza a pesar de lo humano. Esto es muy importante, porque la posibilidad de una optimización de lo humano se determinará en esa dirección: optimizar lo humano para que cumpla los requerimientos de la IA. La esencia de lo humano se transformará, para que se adapte a la creciente digitalización del mundo.

El transhumanismo no tiene en cuenta esta optimización como determinación de la verdad, que se puede establecer desde la IA (Bostrom, 2016, Diéguez, 2017). Las posibilidades de mejoramiento son infinitas en potencia (Bostrom, 2005), pero las prácticas de mejoramiento optimizante que se ejercen y determinan desde las IA y desde el proyecto del transhumanismo digital tienen la intención de diseñar nuevas subjetividades identitarias, en las que los sujetos optimizan su devenir existencial algorítmicamente. Se tratará así de sujetos optimizados algorítmicamente desde el diseño de su subjetividad como una consciencia de sí, que cumple adaptativa e innovadoramente las dinámicas de la sociedad digital.

El proyecto de una optimización de lo humano sigue la misma lógica utilitarista. El transhumanismo busca, en un primer momento, la potenciación de aquello que se considera valioso. Para esto hay varias posibilidades: (1) mejorar los cerebros biológicos; (2) la manipulación genética; (3) células madre embrionarias; (4) liberar las mutaciones (que reducen la eficiencia de los procesos celulares); (5) clonación reproductiva, para replicar capacidades excepcionales (Diéguez, 2017). Pero ¿desde qué intención e interés económico y político se intenta mejorar a los seres humanos? Una respuesta genealógica y deconstructiva (Derrida, 1967) puede ser para que los individuos mejorados cumplan las expectativas utilitaristas de las IA. Los sujetos mejorados algorítmicamente pueden: (1) hacer posible la emergencia controlada de la superinteligencia, y (2) asumir las disposiciones de optimización que la IA determine (Bostrom, 2016). Esto quiere decir que el mejoramiento de los sujetos no implica su libertad y autonomía. El mejoramiento se determinará como nuevo dispositivo de control, desde el diseño de subjetividades algorítmicas. Una consecuencia de lo anterior es que el mejoramiento como posibilidad de una afirmación de sí de la individualidad, como superación del estado de sujeto, se pierde en la optimización de lo útil.

El transhumanismo también plantea otras posibilidades de mejoramiento, entre ellas, la Emulación Completa del Cerebro humano (ECC) (Bostrom, 2016), en la que se asume que también quedará copiada la consciencia, para después realizar una transferencia de dicha copia a un cibernético o a un dispositivo virtual (Diéguez, 2017). Esto haría posible la inmortalidad digital de la consciencia. ¿Cómo sería posible una ECC? Hay dos opciones: realizar una reconstrucción digital de la evolución del cerebro (y de la consciencia e inteligencia correspondientes), mediante un dispositivo de IA; o crear una copia exacta del mapa de las sinapsis neuronales, mediante dispositivos nanotecnológicos (Bostrom, 2016). Estos pueden ser los fundamentos para una filosofía de la consciencia (Dennett, 1995): ¿el sujeto tiene una consciencia de sí?, ¿la consciencia de sí sólo es una posibilidad de la individualidad?, ¿cuál es la consciencia que se pretende clonar en la ECC? Para Dennett (1995) la consciencia está unida a los procesos cerebrales, que pueden recibir una estimulación desde dispositivos de IA; pero este autor no pregunta sobre una consciencia como creación de sí individual desde la que se pueda resistir críticamente a los mismos. La IA como la posibilidad de diseño de nuevas subjetividades impacta sobre la identidad del sujeto y del individuo; al sujeto le impone una identidad y un proyecto de vida optimizante, al individuo le plantea el reto de la pregunta por la consciencia de sí.

Pero más allá del mejoramiento transhumanista, el principio de optimización es lo que reemplazará las funciones de lo humano. Lee (2020) afirma que para el 2030, el 40 % de las funciones que desempeñan los humanos serán reemplazadas por dispositivos de IA. Este reemplazo se realizará mediante las *smartgrids*, que son proveedores de energía independientes, inteligentes y que actúan en tiempo real. Un ejemplo permite comprenderlo: si una *smartgrid* detecta una luz en la terraza de un edificio que siempre está encendida, preguntará el motivo, a lo que obtendrá como respuesta: “es la caseta del celador”, entonces, instalará una cámara y una luz fotosensible al movimiento, que se active solo cuando sea necesario. De este modo, se optimizó la función, pero ¿qué se hace con el celador? Las funciones humanas serán reemplazadas desde el principio de optimización utilitarista (Sadin, 2020a).

6. Diseño de identidades y proyectos de vida

DESDE EL PRINCIPIO DE OPTIMIZACIÓN se diseñan e imponen identidades y proyectos de vida como subjetividades a los sujetos. La optimización de la identidad consiste en el diseño, determinación y agenciamiento de una consciencia de sí del sujeto, desde los parámetros de verdad algorítmica determinada. La identidad se implica de forma inmanente en la optimización de los proyectos de vida, que consiste en la determinación de un devenir existencial del sujeto.

De la misma forma que el inconsciente es un sistema abierto (Guattari, 2000), en continua transformación, la identidad del sujeto y del individuo no es algo que esté determinado, ni a priori ni a posteriori. Determinar una identidad implica un ejercicio de control sobre el sujeto, pero también de creación del individuo. Es una de las mejores formas de ejercer una dominación porque se trata de que el sujeto asuma como propia y esencial una consciencia de sí; y es una de las mejores formas de crítica porque desde una consciencia de sí el individuo resiste la dominación. La identidad es necesaria. Lo problemático no es la pérdida de identidad, sino el asumir una identidad definitiva, porque en esa definición se autolimita, y el que sea una fuerza de agencia externa la que diseña e impone esa identidad. La identidad es importante, siempre que se la asuma como el proyecto dinámico de una creación de sí de la individualidad, como crítica a las identidades algorítmicas (Butler, 2007).

Una de las características fundamentales de las identidades algorítmicas es el fomento y difusión de la diversidad. Desde los estudios de Butler (2006, 2007), podemos afirmar que el proceso de creación continuo de identidad es único y singular. El problema de la creación y determinación de identidades algorítmicas, como afirmación de la diversidad, es que anula ese proceso de creación y clasifica a los sujetos en categorías funcionales. La diversidad funciona para categorizar y, desde ahí, aplicar los principios de optimización utilitarista. Los sujetos del mundo contemporáneo están dominados por procesos identitarios optimizantes. Por eso para Butler (2015) es necesario un paso del sujeto al individuo, en la afirmación de un proceso de creación identitaria como consciencia autónoma.

Las determinaciones algorítmicas de las identidades implican crear proyectos de vida a los sujetos, que deben ser objeto de reflexión crítica como práctica existencial y afirmativa de una individualidad. Es necesario distinguir el mero vivir,

de la vida auténtica (Heidegger, 1997; Chul Han, 2021). El primero consiste en vivir sin proyecto, la segunda en vivir en un proyecto superior (Slöterdijk, 2012). Pero en las identidades algorítmicas se imponen proyectos de vida que se acoplen y sean coherentes con las intenciones sociopolíticas con las que la IA es diseñada funcionalmente. No es un mero vivir, pero tampoco es una vida auténtica. Es una vida eficiente, que se puede cuantificar. De hecho, podemos referirnos a dichos proyectos como una determinación a priori y genética. Para Butler (2006), por ejemplo, es eso lo que hace el denominado “patriarcado”: determinar categorías a priori, que los individuos deben asumir como un proyecto vital. La diferencia es que la IA determina proyectos de vida a priori, desde principios algorítmicos de optimización.

7. El proyecto de una *Data Driven Society* (DDS)

UNA SOCIEDAD CUYO FUNCIONAMIENTO SE BASE EN DATOS tiene como fundamento crear e implementar una aplicación para todo. Los datos en tiempo real sirven para tomar decisiones inmediatas y sin reflexión. Los datos, en realidad, se utilizan como instrumento de análisis para la implementación transformadora de las funcionalidades. No se trata de analizar los datos para asumir y suplir las necesidades que de ellos se puedan deducir, sino de analizar los datos para introducir acciones de mejoramiento. Desde esta perspectiva, las aplicaciones no son instrumentales, sino preceptivas: indican lo que se debe hacer (Sadin, 2020a).

Un ejemplo claro de cómo la automatización digital de la sociedad es un proyecto societal global es la aplicación Dreem. Éric Sadin (2020a) la describe así:

Ya mismo podemos beneficiarnos, en nuestra habitación, de las virtudes tranquilizadoras que nos promete el dispositivo conectado Dreem, puesta a punto por la *start-up* francesa Rythm, que debe colocarse como si fuera una vincha en el momento de acostarse y que se supone que conjuga “los métodos más eficaces, desde el *biofeedback* a la neuromodulación, para mejorar el sueño en lo cotidiano”. El dispositivo supervisa durante la noche los movimientos del usuario y su ritmo cardíaco, así como la actividad del cerebro va un electroencefalograma. Emite sonidos que se propagan por conducción ósea a fin de favorecer el adormecimiento. Una vez alcanzado el sueño, los sonidos se modulan como “sonidos rosas”, que son los destinados a amplificar

la acción del tálamo, que ajusta el nivel correcto de ondas delta que favorece el sueño profundo. Finalmente, el dispositivo desencadena un despertar llamado “inteligente” en el momento que se considera óptimo respecto de la actividad cerebral por medio de una alarma compuesta, según las circunstancias, por cantos sintetizados de pájaro o de cigarras, para dar un ejemplo. Quienes lo concibieron afirman sin problema que cumple el rol de un “*coach* del sueño” (pp. 109-110).

Se podría pensar que los dispositivos tienen como finalidad la satisfacción de las necesidades de los individuos. Pero si lo reflexionamos en profundidad, podremos comprender que no hay un mejoramiento del sueño, sino un diseño y determinación de lo que debería ser un buen sueño, es decir: un sueño que haga posible el rendimiento (Chul Han, 2014). Se requiere un sueño tranquilizador y positivo, que permita un plus de acción y producción. Foucault (1994) realiza una observación muy importante con respecto a la interpretación psicoanalítica de los sueños: los sueños no solo dan cuenta de los procesos de represión sintomática, sino que también son proyecciones de sí; en los sueños se expresa una proyección a sí mismo, en lo que Deleuze y Guattari (1991) llaman un mundo posible. Pero si se logra instalar y hacer que el sujeto se apropie de una aplicación que le haga posible un sueño bueno e inteligente, entonces lo que se va a proyectar son los ideales del rendimiento y la optimización.

Desde el siglo XIX se inició el proceso de configurar la IA con los ideales del rendimiento y la optimización, cuyo objetivo es una automatización digital total de la sociedad, que hará posible un goce supremo, sustentado en la eficiencia. Para el proyecto de la DDS, la finalidad de lo humano no debe ser la del cumplimiento de tareas instrumentales, sino la de un disfrute existencial. Estaríamos ante la posibilidad de una *hedoné* digitalizada, en la medida en que los dispositivos de IA permiten un relajamiento de toda función (Sadin, 2020a). Esto tiene importantes consecuencias: el advenimiento de un mundo feliz ultra-controlado (Estado de bienestar), una clasificación y jerarquización entre sujetos instrumentales e individuos contemplativos o hedónicos, la implementación del *bannop-tikum*, porque los sujetos instrumentales no serán reconocidos en plenitud como individuos (Chul Han, 2014, 2021).

Cuando los dispositivos de IA asuman todas las funciones instrumentales, se creará un Estado de bienestar digital, racional y automatizado en el que se

determinarán nuevas formas de subjetividades. Las nuevas subjetividades deben tener unas características integrales que les permitan un acople coherente con dicha sociedad. Se trata de la transformación del sistema integral de la corporalidad, de la sensibilidad y del pensamiento. Es en realidad una transformación de la esencia de lo humano. Por eso el transhumanismo es la primera fase de un posterior poshumanismo en el que emerjan nuevas subjetividades con una esencia distinta. Se trata del diseño de nuevas subjetividades que pueden ir más allá del sujeto, pero que niegan la posibilidad de la creación de la individualidad: *Data Driven Society* o el advenimiento de un nuevo Leviatán digital (Sadin, 2020a).

El mismo principio opera en el ChatGPT o Generative Pre-trained Transformer (transformador generativo pre-entrenado) (García-Peñalvo, 2023). Este tiene la calidad de generar textos con una alta capacidad argumentativa y puede mantener conversaciones de corte realista (García-Peñalvo, 2023). La versión completa de OpenAI GPT-3 es el modelo más grande entrenado hasta la fecha en comparación con otros modelos lingüísticos (Cooper, 2023). Una IA generativa es aquella que puede producir textos, pero lo hace de una forma eficiente y mejorada. Es decir: evita el exceso de prosa y de rodeo, para realizar un máximo de rendimiento en la argumentación. El ChatGPT plantea un argumento certero y eficiente. En la eficiencia no es necesario demorarse en la argumentación. El punto crítico está en el desplazamiento invisible del pensamiento crítico, profundo y creativo, hacia un dispositivo de IA que pensará por el individuo de forma veraz e incuestionable. Esto es complejo si asumimos la afirmación de Derrida (1967) según la cual el pensamiento es en sí mismo un rodeo problematizador. Pensar es pensar diferencialmente (Deleuze y Guattari, 1991). Estamos ante una pérdida de sentido como proyecto social global.

El proyecto de una DDS consiste en optimizar algorítmicamente lo humano y lo social, desde la idea de un Estado de bienestar que se actualiza en tiempo real. Optimizar el pensamiento quiere decir eliminar el rodeo argumentador y la praxis crítica en la afirmación de una certeza sin cuestionamiento. La DDS se fundamenta en el diseño y determinación de nuevas subjetividades, en la creación de una nueva consciencia de sí de los sujetos, desde la que despliega un proyecto de vida como rendimiento. La DDS es un programa de diseño global que se fundamenta en el diseño de un nuevo sujeto óptimo.

8. La emergencia de una superinteligencia

¿QUÉ ES UNA SUPERINTELIGENCIA? Es una inteligencia que es autónoma y consciente de sí misma. Una superinteligencia es lo que se denomina como singularidad. De lo que se trata es de preguntar sobre la posibilidad real de dicha emergencia y de sus posibles consecuencias. El director de ingeniería de Google, Kurzweil (2005) tiene una postura optimista, al afirmar que la singularidad está próxima a emerger. Lee (2020) es un poco más prudente y, aunque afirma que la emergencia explosiva o disruptiva de la singularidad es inevitable, considera que requiere mucho tiempo y esfuerzo. En un punto intermedio se encuentra Bostrom (2016) para quien se deben ir creando condiciones adecuadas para la emergencia de la singularidad; estas deben estar enfocadas hacia un máximo de control, que implique un mínimo posible de riesgos existenciales.

Pero ¿por qué riesgos existenciales?, ¿qué podría pasar con los seres humanos, ante la emergencia explosiva de una superinteligencia? Hay al menos tres posibilidades que sirven para reflexionar sobre el desarrollo de la consciencia: una colaboración evolutiva con lo humano, una hostilidad hacia lo humano y una invisibilización e indiferencia de la superinteligencia ante lo humano (Bostrom, 2016). En realidad, lo anterior funciona como punto crítico para una investigación sobre la esencia de la consciencia. La pregunta heideggeriana por el ser debe desplazarse a una pregunta por la esencia de la consciencia de sí. Sería una pregunta óntica, porque preguntaría por la consciencia en sí, independientemente de la consciencia humana.

La pregunta por el principio de optimización, que se realiza desde el racionalismo aumentado de la IA, no puede plantearse por la posibilidad de una optimización esencial o en sí misma, porque esto implicaría un cálculo de todas las posibilidades para, desde ahí, asumir el mejor proyecto de optimización posible. Esto desde la IA como dispositivo de una racionalidad aumentada; si la IA evoluciona hacia la singularidad, el principio podría ser posible pero muy diferente a lo humano. Por eso se asume a priori un principio de optimización utilitarista que no es el único ni el mejor posible, pero es el que asumen quienes tienen el control sobre el desarrollo de la IA (Bostrom, 2016).

Para la emergencia de una superinteligencia singular, primero se debe desarrollar una IA al nivel de lo humano, algo que aún está muy lejos de lograrse.

Según Bostrom (2016), una IA al nivel humano, en una probabilidad del 90 % se dará en 2075, y una superinteligencia, en una probabilidad del 75 % se dará 30 años después de la emergencia de la IA de nivel humano. Son cifras probabilísticas que afirman la posibilidad, pero también la dificultad de dichas ocurrencias, dejando abierta también la posibilidad de que no sucedan. Lo importante está en plantear la posibilidad de la emergencia, controlada o disruptiva, de la singularidad porque demanda replantear la concepción de la consciencia misma. Si las singularidades se pueden diseñar, también es posible la emergencia de la singularidad como consciencia (Bostrom, 2016; Dennett, 1995).

¿Cómo se podría hacer posible la emergencia explosiva (Bostrom, 2016) o disruptiva de una superinteligencia? Hay que afirmar, primero, que la concepción de explosividad se refiere a aquello que surge con cierta independencia de su causa, y con efectos diferentes e insospechados (Bostrom, 2016). Para pensar, entonces, la posibilidad de una superinteligencia, se podría hacer lo siguiente: una reconstrucción digital y virtual de la evolución y emergencia de la inteligencia humana; una Emulación Completa del Cerebro (ECC), y un mejoramiento del cerebro humano. Aunque las tres opciones requieren el despliegue y desarrollo de alta tecnología y energía, se puede pensar en dos opciones más al alcance: una ECC de organismos básicos, como los insectos, y estimular desarrollos de formas leves de superinteligencia humana, mediante mejoras biotecnológicas, como la interfaz no invasiva de cerebro-ordenador (Diéguez, 2017).

Si bien hasta un punto el desarrollo y evolución de la IA dependen de la capacidad de la inteligencia humana, y de los recursos tecnológicos y energéticos con los que esta pueda contar, también se debe comprender y aceptar que en un punto ulterior el desarrollo y evolución de la IA, y la posibilidad de la emergencia explosiva de la superinteligencia, no dependerán de lo humano. Estamos ante el proceso y proyecto de una gradual independencia y diferencia funcional.

9. Del sujeto al individuo: los desafíos de la singularidad

LOS DISEÑOS ALGORÍTMICOS DE LAS IA tienen la capacidad eficiente de determinar nuevas subjetividades. A los sujetos les determina una consciencia de sí identitaria y optimizante; a la individualidad como posibilidad le permite pensar sobre la consciencia de sí como afirmación existencial, en tanto se asume la

posibilidad de una emergencia de la singularidad como consciencia autónoma de sí. El individuo como proyecto se puede pensar en esa emergencia disruptiva de la consciencia como singularidad. La IA impone identidades y proyectos de vida en la determinación de un nuevo sujeto optimizado, pero también abre a la posibilidad de la emergencia de una consciencia singular.

Para Butler (2007, 2015) es necesario plantear y pensar el paso teórico del sujeto al individuo, de forma que los procesos de dominación algorítmica se puedan transformar en la afirmación de sí como performatividad existencial. El ejercicio político de la IA en el mundo contemporáneo consiste en la determinación de un nuevo sujeto, pero ¿pueden los desarrollos de las IA permitir pensar en la creación de una consciencia de sí como individualidad, desde la posibilidad de la emergencia de la singularidad?, ¿una singularidad digital puede permitir pensar en una singularidad como afirmación de la individualidad? El paso teórico del sujeto al individuo pregunta por la disrupción política del individuo como performatividad de sí. La individualidad como disrupción singular es la afirmación del extrañamiento como devenir existencial (Slöterdijk, 2012).

La IA impacta sobre las subjetividades en la determinación de una consciencia de sí como identidad y proyecto de vida; pero sus desarrollos también permiten pensar sobre la creación de una consciencia de sí como individualidad, en tanto pone en cuestión las determinaciones cerebro-neuronales como aquellas que hacen posible la consciencia (Dennett, 1995). Los desarrollos en IA cuestionan la filosofía de la conciencia de Dennett, porque permiten pensar en una consciencia de sí algorítmica que, aunque se fundamenta en una ECC, va más allá en el planteamiento de la posibilidad de una disruptividad emergente de la singularidad como consciencia. Una singularidad, ya sea digital o individual, no dependería de condiciones biológicas (Dennett, 1995) para la emergencia disruptiva de una consciencia de sí. La IA como singularidad plantea el desafío de la posibilidad de una consciencia de sí como individualidad. El paso teórico del sujeto al individuo está en la disruptividad de una consciencia como autonomía.

10. Conclusión. El impacto de la IA en los procesos de consciencia

LOS DESARROLLOS DE IA se constituyen como el más eficiente dispositivo de dominación (Sadin, 2020a), en tanto tienen la capacidad de diseñar y determinar

la consciencia del sujeto. Pero también pueden abrir el pensar a la posibilidad del diseño y creación de una consciencia de sí como individualidad, desde la afirmación de la singularidad como autonomía (Bostrom, 2016; Butler, 2015). El sujeto es una posibilidad de individualidad como disrupción, de la misma forma que las IA son singularidades en potencia. El control es cada vez más eficiente y absoluto, en tanto se determina la verdad como esencialidad optimizante. Pero la singularidad como disrupción, tanto digital como individual, se puede pensar como posible.

Si bien la psicopolítica se constituye como una forma eficiente de dominación, en tanto determina un psico-programa deseante del inconsciente que se asume desde la identidad y el proyecto de vida, es fundamental comprender y asumir la posibilidad de pensar a la subjetividad como un proceso de creación, en tanto se estudia la consciencia como emergencia de individualidad. La psicopolítica devela que la consciencia es una posibilidad de creación. Solo que en la DDS esa consciencia es diseñada y determinada como una subjetividad optimizante (Sadin, 2020a).

Los diseños de la subjetividad que determinan los dispositivos de IA tienen un impacto inédito sobre los estudios de la consciencia, porque se puede pensar y afirmar que: (1) la emergencia de la consciencia depende de condiciones externas que no solo son biológicas (Dennett, 1995); (2) la consciencia es una posibilidad de creación de la subjetividad, ya sea como sujeto o como individuo (Butler, 2015); (3) la consciencia no está determinada sino que se puede diseñar y crear –el problema es quién o qué lo hace–; (4) las subjetividades como sujeto son determinadas en el mundo contemporáneo por dispositivos de IA (Chul Han, 2014); (5) los desarrollos de IA tienen la potencia de la emergencia de la singularidad como consciencia autónoma (Bostrom, 2016); y (6) que el individuo puede emerger como consciencia singular (Butler, 2015).

La optimización algorítmica y utilitarista de la DDS debe estudiarse desde una filosofía crítica de la IA en la que se afirme la posibilidad de los diseños y creación de subjetividades diferenciales. La algoritmización de la sociedad implica la posibilidad de pensar la emergencia disruptiva de la singularidad, tanto digital como individual. La IA es la crisis de la consciencia en la posibilidad de un diseño y determinación de la subjetividad como sujeto, o como individualidad disruptiva. Una crítica deconstructiva (Derrida, 1967) de la IA como dispositivo

de dominación debe afirmar como inmanencia (Deleuze y Guattari, 1991) la problematización de la consciencia.

Referencias

- Bostrom, N. (2005). A History of Transhumanist Thought. *Journal of Evolution and Technology*, 14(1), 1-25.
- Bostrom, N. (2013). Existential risk prevention as global priority, *Global Policy*, 4(1), 15-31.
- Bostrom, N. (2016). *Superinteligencia. Caminos, peligros, estrategias* (trad. M. Alonso). Teell Editorial.
- Butler, J. (2006). *Deshacer el género* (P. Soley-Beltran, trad.). Paidós.
- Butler, J. (2007). *El género en disputa. El feminismo y la subversión de la identidad* (M. Muñoz, trad.). Paidós.
- Butler, J. (2015). *Mecanismos psíquicos del poder. Teorías sobre la sujeción* (J. Cruz, trad.) Ediciones Cátedra, Universitat de Valencia.
- Castro-Gómez, S. (2019). *El tonto y los canallas. Notas para un republicanismo transmoderno*. Editorial Universidad Javeriana.
- Citton, Y. (2014). *Pour une écologie de l'attention*. Seuil.
- Cooper, K. (2023, Sept. 27). OpenAI GPT-3: Everything You Need to Know. *Springboard*. <http://bit.ly/3GLMkmB>
- Chul Han, B. (2014). *Psicopolítica* (A. Bergés, trad.). Herder.
- Chul Han, B. (2017). *La sociedad del cansancio* (A. Saratxaga y A. Ciria, trads.). Herder.
- Chul Han, B. (2021). *La sociedad paliativa. El dolor hoy* (A. Ciria, trad.). Herder.
- Deleuze, G. & Guattari, F. (1991). *Qu'est-ce que la philosophie?* Les Éditions de Minuit.
- Dennett, D. (1995). *La conciencia explicada. Una teoría interdisciplinar* (P. Weiner, trad.). Paidós.
- Derrida, J. (1967). *L'écriture et la différence*. Éditions du Seuil.
- Diéguez, A. (2017). *Transhumanismo. La búsqueda tecnológica del mejoramiento humano*, Herder.
- Foucault, M. (1994). *La hermenéutica del sujeto* (F. Alvarez-Uría, trad.). Ediciones de la Piqueta.

- Foucault, M. (2002). *Vigilar y castigar. Nacimiento de la prisión* (A. Garzón del Camino, trad.). Siglo XXI.
- Foucault, M. (2008). *Tecnologías del yo* (M. Allendesalazar, trad.). Paidós.
- García-Peñalvo, F. (2023). The Perception of Artificial Intelligence in Educational Contexts After the Launch of ChatGPT: Disruption or Panic. *Education in the Knowledge Society*, 24, 1-9. <https://doi.org/10.14201/eks.31279>
- Guattari, F. (2000). *Cartografías esquizoanalíticas* (D. Scavino, trad.). Manantial.
- Gumbrecht, H. (2020). *El espíritu del mundo en Silicon Valley. Vivir y pensar el futuro* (S. Yusta, trad.). Deusto.
- Hadot, P. (1998). ¿Qué es la filosofía antigua? (E. Tapie Isoard, trad.). Fondo de Cultura Económica.
- Heidegger, M. (1997). *Filosofía, ciencia y técnica* (J. Acevedo, trad.). Editorial Universitaria.
- Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. Penguin.
- Lacan, J. (2003). *Seminario 8. La transferencia* (E. Berenguer, trad.). Paidós.
- Lipovetsky, G. (2000). *La era del vacío. Ensayos sobre el individualismo contemporáneo* (J. Vinyoli y M. Pendanx, trads.). Anagrama.
- Lipovetsky, G. (2004). *El imperio de lo efímero. La moda y su destino en las sociedades modernas* (F. Hernández y C. López, trads.). Anagrama.
- Lee, K. (2020). *Superpotencias de la inteligencia artificial. China, Silicon Valley y el nuevo orden mundial* (M. Baquero, trad.). Deusto.
- Sadin, E. (2020a). *La inteligencia artificial o el desafío del siglo. Anatomía de un antihumanismo radical* (M. Martínez, trad.). Caja Negra.
- Sadin, E. (2020b). *La silicolonización del mundo. La irresistible expansión del liberalismo digital* (M. Martínez, trad.). Caja Negra.
- Sadin, E. (2022). *La era del individuo tirano. El fin de un mundo común* (M. Martínez, trad.). Caja Negra.
- Slöterdijk, P. (2012). *Has de cambiar tu vida* (P. Madrigal, trad.). Pre-Textos.
- Warf, B. (2024). Posverdad: Geographies of Post-Truth in Latin America. *Journal of Latin American Geography*, 23(1), 80-99. <https://dx.doi.org/10.1353/lag.2024.a929688>.