

Reflexiones éticas sobre la inteligencia artificial. Implicancias de la lógica algorítmica sobre la subjetividad.

Salomone, Gabriela Z.

Cita:

Salomone, Gabriela Z (2025). *Reflexiones éticas sobre la inteligencia artificial. Implicancias de la lógica algorítmica sobre la subjetividad. XVII Congreso Internacional de Investigación y Práctica Profesional en Psicología. XXXII Jornadas de Investigación XXI Encuentro de Investigadores en Psicología del MERCOSUR. VII Encuentro de Investigación de Terapia Ocupacional. VII Encuentro de Musicoterapia. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires.*

Dirección estable: <https://www.aacademica.org/000-004/13>

ARK: <https://n2t.net/ark:/13683/eNDN/D2r>

REFLEXIONES ÉTICAS SOBRE LA INTELIGENCIA ARTIFICIAL. IMPLICANCIAS DE LA LÓGICA ALGORÍTMICA SOBRE LA SUBJETIVIDAD

Salomone, Gabriela Z.

Universidad de Buenos Aires. Facultad de Psicología. Buenos Aires, Argentina.

RESUMEN

El texto reflexiona particularmente sobre la narrativa dominante en torno a la inteligencia artificial, que la presenta como solución a múltiples problemas, con promesas de eficiencia, precisión y automatización. Se retoman aportes de diversos pensadores, que analizan el efecto de la tecnología digital sobre las formas de pensar, sentir y actuar, para centrar el análisis en las transformaciones que impactan en las condiciones de producción de subjetividad. Se propone una interrogación ética sobre la inteligencia artificial, sus fundamentos conceptuales y su lógica de funcionamiento, particularmente la expansión de la lógica de los algoritmos computacionales de decisión en el mundo humano y sus efectos desubjetivantes. Se destaca la necesidad de no delegar totalmente la acción ni el pensamiento a los sistemas automatizados y la importancia de sostener una posición crítica que reintegre la dimensión subjetiva, interrogando las prácticas y sus fundamentos.

Palabras clave

Inteligencia artificial - Automatización ética - Desubjetivación - Algoritmos de decisión

ABSTRACT

ETHICAL REFLECTIONS ON ARTIFICIAL INTELLIGENCE:

THE IMPACT OF ALGORITHMIC LOGIC ON SUBJECTIVITY

The text offers a critical reflection on the dominant narrative surrounding artificial intelligence, which frames it as a solution to a wide range of problems, promising efficiency, precision, and automation. Drawing on insights from various thinkers who examine the impact of digital technologies on how we think, feel, and act, the analysis centers on the transformations that affect the conditions under which subjectivity is produced. It calls for an ethical inquiry into artificial intelligence —its conceptual foundations and its operational logic— focusing in particular on the growing influence of algorithmic decision-making in human affairs and its desubjectivizing effects. The text underscores the importance of not fully delegating action or thought to automated systems and stresses the need to maintain a critical stance that reintroduces the subjective dimension by questioning practices and their underlying assumptions.

Keywords

Artificial intelligence - Ethical automation - Desubjectification - Decision algorithms

Sigmund Freud plantea que, al igual que la relación con el semejante, la naturaleza y el propio cuerpo son fuentes de satisfacción y a la vez de permanente sufrimiento para el ser humano. “El desvalimiento y desconcierto del género humano son irremediables”, dijo en 1927, al referirse a esta condición constitutiva del ser humano, que lo acompaña a lo largo de toda su vida. Gracias a su capacidad simbólica, la humanidad ha generado en todos los tiempos modos de mitigar tal desvalimiento, mediatisando su relación con las fuentes de sufrimiento: en la forma de *mediaciones normativas* que regulan la amenaza que proviene del semejante, y *mediaciones instrumentales* para lidiar con la fragilidad del propio cuerpo y generar recursos para enfrentar los embates de la naturaleza y desasirse de las limitaciones que esta le impone.^[i]

El desarrollo científico-tecnológico ha jugado históricamente un rol central en la mitigación del desvalimiento humano en lo real de su relación con el entorno, pero también como recurso simbólico para mitigar el sentimiento de indefensión, sumado en este punto a sistemas de creencias, ideologías, religiones, concepciones compartidas, creación artística, etc. Al respecto, cabe señalar que las formas instrumentales —como sistemas normativos, instituciones sociales, avances científicos y tecnológicos, etc.— van generando nuevos sistemas de sentido, nuevas concepciones que interrogan, amplían o cuestionan los sistemas de representaciones vigentes.

Entendiendo que corresponde a esta serie, nos proponemos reflexionar sobre la Inteligencia Artificial (IA), cuya narrativa la presenta como una tecnología con el potencial de asistir al ser humano en prácticamente todas las áreas de su existencia, mediante sistemas informáticos capaces de imitar ciertas funcionalidades de la inteligencia humana.

La IA ya se aplica en diversos ámbitos, como el diagnóstico médico, herramientas predictivas del comportamiento, estimación de probabilidad delictiva, redes sociales, automatización laboral, selección de personal, gestión de recursos humanos, asignación de créditos y planes sociales. Además, participa en la generación de soluciones innovadoras en educación,

creatividad, justicia, gestión urbana, ciberseguridad, economía y epidemiología.

Para referirse a la actual era de la información, Diana Fernández Zalazar toma el término del investigador catalán Manuel Castells, *nueva revolución sociotecnológica* (Castells, 2006), para referirse a los cambios cualitativos que van aconteciendo. La autora plantea lo siguiente:

A diferencia de las revoluciones anteriores, en esta última etapa donde surgen tecnologías emergentes (nanotecnología, biotecnología, Inteligencia Artificial (IA), robótica, computación cuántica, etc.) (...) Esta posibilidad de coordinar lo físico, lo orgánico y lo informacional a partir de la transversalidad de las tecnologías digitales, hace de este mundo una nueva forma de extensionalidad, velocidad e impacto en los sistemas complejos que organizan nuestro devenir histórico y cultural. (Fernández Zalazar, 2024)

En este sentido, el despliegue de la inteligencia artificial no es solo técnico: el desarrollo tecnológico transforma concepciones, valores y modos de relación, imprimiendo cambios en la “subjetividad de la época”, entendida como un producto de diversos dispositivos sociohistóricos.

Estas transformaciones nos llevan a interrogarnos éticamente sobre la Inteligencia Artificial, sus fundamentos conceptuales y su lógica de funcionamiento, atentos a su impacto sobre las condiciones de producción de subjetividad (Bleichmar, 1999). Seguimos en este punto la indicación de Jacques Lacan (1953) sobre no desentenderse de los cambios culturales, socio-políticos, económicos, tecnológicos, pues son el contexto del sujeto de nuestras prácticas y el propio. En sus formulaciones se lee con claridad que no se trata de algún tipo de complicidad con las innovaciones de la época, sino de sostener una posición de reflexión, interrogación y lectura crítica.

Diversos pensadores contemporáneos analizan el desarrollo y uso de la inteligencia artificial y su impacto en el orden social y subjetivo. Éric Sadin (2015) habla de la “algoritmización de la vida”, señalando cómo la lógica algorítmica de la IA atraviesa todos los aspectos culturales lo cual, según su percepción, va produciendo una concentración del poder en las grandes corporaciones multinacionales vinculadas al campo de la tecnología digital (Meta, Amazon, Google, Nvidia). Al respecto utiliza el término “siliconización del mundo”, para referirse a un proceso global de instauración de los modelos, valores y lógicas de Silicon Valleyⁱⁱⁱ, epicentro de la industria tecnológica en Estados Unidos.

Por su parte, el filósofo surcoreano-alemán Byung-Chul Han describe la “dataización de la vida” como el proceso en el que todos los aspectos de la existencia humana son traducidos a datos cuantificables, para ser gestionados algorítmicamente (mediante IA), facilitando así el control y la predicción del comportamiento. En este *régimen psicopolítico*, según Han: “Hoy no es necesario oprimir a las personas para dominarlas. Se entregan voluntariamente, incluso con entusiasmo, a la lógica de la vigilancia, a

través de la transparencia y el rendimiento” (Han, 2014).

En relación con estas transformaciones que los autores mencionados señalan, Susana La Rocca, especialista en bioética, plantea que: “En esta cosmovisión se inaugura una racionalidad a-normativa, a-política, sin sujeto reflexivo y que se alimenta de datos infra-individuales otorgados voluntariamente por los mismos usuarios del sistema digital”. (La Rocca, 2025)

Miguel Benasayag (2024), psicoanalista argentino y actual investigador en neurofisiología en Francia, advierte sobre la “colonización algorítmica” de la vida, entendida como una dominación creciente de la lógica de tecnologías de IA, en detrimento de la experiencia subjetiva y del error. Señala que esta lógica promueve un ideal de “buen funcionamiento” que también alcanza al ser humano, con un efecto de debilitamiento de la subjetividad y la singularidad, favoreciendo la deshumanización de las prácticas sociales.

Las configuraciones tecnológicas, así como las políticas y sociales, producen transformaciones en la subjetividad que cada época instaura; transformaciones que moldean a los sujetos de un modo funcional para ese orden social. En este sentido, es necesario sostener una responsabilidad ética frente a los modos actuales de subjetivación, que configuran asimismo modos particulares del malestar en la cultura.

La Inteligencia Artificial encarna una nueva racionalidad tecnológica que reorganiza el conocimiento y los modos de conocer, la experiencia, la acción, la relación con el otro. En el campo de la tecnología digital, los algoritmos de decisión se presentan como procesos computacionales capaces de analizar datos, evaluar opciones y arribar a conclusiones en función de criterios y objetivos predefinidos. Este proceso es denominado “toma de decisión”, por lo que se dice que el algoritmo “decide”, emulando procesos humanos. Tal vez haya que advertir que no es inocente la elección del término “decidir/decisión”ⁱⁱⁱ que, al equiparar procesos humanos y no humanos termina en una dañina simplificación de lo humano, reduciendo a datos numéricos la experiencia –incluida la del razonamiento–.

Los algoritmos computacionales de decisión están formalizados en secuencias lógicas bajo una racionalidad estadística y de ordenamiento de datos. Mientras que la decisión humana no puede reducirse a esa lógica, pues implica principalmente un proceso de deliberación, que incluye una operación subjetiva (no solo razonamiento). Incluso, un acto de decisión supone un salto cualitativo respecto de los elementos de los que parte, produciendo un plus fuera de todo cálculo.

La narrativa dominante sobre la IA enfatiza su capacidad para automatizar tareas y optimizar procesos, lo cual resulta de gran beneficio para el ser humano en diversas áreas de su existencia. No obstante, esta narrativa también promete reducir errores humanos y sortear el *sesgo humano* en la toma de decisiones, entendido como un problema a eliminar. Según los objetivos expresados, esto llevaría a una mayor eficiencia y precisión en la toma de decisiones.

Sin embargo, esta pretensión de objetividad absoluta es problemática. Por una parte, es razonable esperar que un programa computacional funcione sin errores, sobre todo si de él dependen resultados de gran importancia para la humanidad como una intervención quirúrgica realizada por un robot, el envío de un cohete al espacio, el funcionamiento de un microchip implantado en un cerebro humano, solo por nombrar algunos. No obstante, señalemos que la narrativa sobre la IA en general enfatiza la necesidad de reducir errores *humanos* para lograr la deseada precisión. Retomaremos este punto.

Por otra parte, al considerar la cuestión de la objetividad, se debe tener en cuenta que la automatización de decisiones por parte de la IA se basa en decisiones tomadas por seres humanos, en relación con la elección, carga y tratamiento de datos, con sus propios sesgos, concepciones y valores. Estas decisiones encontrarán fundamento en perspectivas teóricas, ideológicas, políticas, morales, tal vez de los programadores, pero sobre todo de las instancias superiores de decisión, relativas a las corporaciones dedicadas a su desarrollo. De este modo, las respuestas que ofrezcan las IA pueden fácilmente reproducir discriminaciones y estereotipos existentes –como sesgos culturales o de género– o llegar incluso a resultados erróneos.

Este proceso en el que interviene el ser humano es inevitable, por lo tanto, la reflexión sobre el sesgo humano se torna necesaria. No se trata solo del sesgo de la discriminación o de los estereotipos sociales, como suele subrayarse. Al respecto, tal vez la clave sea elegir cuál es el sesgo aceptable, los límites de la desviación y, sobre todo, estar advertidos de su existencia. Esto llevaría a un uso y a un diseño responsable de los sistemas de inteligencia artificial.

No obstante, la cuestión de las operaciones subyacentes no suele explicitarse en la narrativa predominante sobre la IA. Por el contrario, se enfatiza una pretendida objetividad –en los términos de una asepsia libre de contaminación– que genera un efecto muy negativo: la idea de que los procedimientos automatizados arriban a la mejor solución posible. Esto condiciona la disposición a la reflexión sobre los sesgos, así como al cuestionamiento crítico de sus conclusiones. Peligrosa consecuencia, si se tienen en cuenta, por ejemplo, los usos de la IA en salud y salud mental, en el ámbito de la justicia, en la gestión de poblaciones, entre otros.

La circulación por las redes sociales y otras plataformas digitales nos muestra una permanente –y hasta insistente– invitación a la toma algorítmica de decisiones; invitación a delegar en programas de IA acciones y funciones cotidianas. Publicaciones que presentan un texto como el que sigue: “¿Por qué nadie está hablando de cómo la IA puede hacer tu trabajo en tres minutos?”. Esta invitación tiene algunas consecuencias. En primer lugar, claramente, la propuesta no es el uso de la inteligencia artificial como ayuda para la realización de tareas, sino que se trata de la sustitución. En los términos planteados más arriba, la automatización de saberes técnicos, procedimientos y normas lleva

a la *des-humanización*, pero ya no solo en el sentido de privar de caracteres humanos a los seres humanos, sino en un sentido más radical como *des-infectar*, suprimir lo que infecta, lo que corrompe.

En segundo lugar, señalamos el fenómeno de la *delegación masiva de funciones* hacia la IA, tanto individuales como sociales, al que se refiere Benasayag. El autor advierte que, tal como lo muestra la historia evolutiva, dejar de usar ciertas áreas del cerebro produce cambios neurofisiológicos que involucran modificaciones en las redes neuronales. Pero señala que, en el caso actual relativo a la delegación de tareas hacia la IA, al producirse de un modo tan masivo y tan rápido ocasiona un empobrecimiento del funcionamiento cerebral, porque no hay tiempo para la regulación, es decir, para el desarrollo de otras áreas cerebrales. Respecto de los efectos de la delegación de funciones, una reciente investigación del Instituto Tecnológico de Massachusetts (MIT) arrojó también datos interesantes (Kosmyna et al., 2025). El objetivo fue evaluar el impacto de herramientas digitales como ChatGPT en la escritura académica. El estudio se realizó sobre 54 estudiantes universitarios, organizados en tres grupos, a quienes se les solicitó escribir un breve ensayo. El primer grupo utilizó ChatGPT, el segundo accedió a un motor de búsqueda tradicional sin funciones de inteligencia artificial, y el tercero no usó ningún soporte tecnológico (grupo “brain only” [solo cerebro]). La actividad cerebral se registró durante la escritura, mediante electroencefalograma (EEG). Entre otros resultados que no explicitaremos en esta ocasión, el experimento detectó disminuida en un 55% la conectividad cerebral en el grupo que utilizó el ChatGPT, y el 83% de ese grupo presentó dificultades para recordar sus propios textos y citar algún pasaje. Nos interesa particularmente este último dato, puesto que permite sumar otra perspectiva de lectura, más allá del aspecto meramente cerebral relativo a la memoria como función cognitiva. Conjuntamente con el empobrecimiento del funcionamiento cerebral que mencionamos anteriormente debemos considerar aquí un *empobrecimiento subjetivo*, producto de la no participación activa en el desarrollo de la tarea. Al ser sustituido en la escritura, se pierde la operación subjetiva que un proceso deliberativo supone, necesario para decidir los diferentes momentos de producción de un escrito. La ajenidad respecto del producto realizado produce efectos desubjetivantes.

En tercer lugar, interesa señalar esta idea de un *saber total con garantías de eficacia*, que se va generando a partir de estas narrativas. Un discurso sin fisuras, que no incluye el equívoco, ni la equivocación, ni el error. Quienes trabajamos en el campo del sujeto y la subjetividad sabemos que un discurso tal no es propicio para el despliegue subjetivo. Los discursos totalitarios producen efectos desubjetivizantes, el borramiento mismo del sujeto.

En este sentido, la perspectiva ética nos lleva a interrogar el *uso automatizado* de la inteligencia artificial. No nos referimos ahora a su diseño e implementación, sino a las distintas modalidades de su uso, los diferentes modos de relacionarse con ella.

Nos centramos entonces en el sujeto y en el campo de la responsabilidad. No es lo mismo utilizar la IA como herramienta y recurso, que tomarla como relevo o sustituto del ser humano. El sujeto se borra a sí mismo cuando hace consistir un saber total, soslayando incluso los puntos de inconsistencia de esa supuesta totalidad. En cambio, una posición de cuestionamiento y reflexión reintroduce el campo subjetivo en la pretendida objetividad y asepsia del discurso digital.

Desde esta perspectiva, la dimensión ética no se reduce a la aplicación de normativas y procedimientos, sino que supone una interrogación y una toma de posición que se juegan en campo de la responsabilidad (Salomone, 2021). Como señaló Ignacio Lewkowicz, la ética es una disciplina del pensamiento, no del conocimiento. Es decir, no se agota en la aplicación de saberes ya existentes, sino que involucra una operación subjetiva, un posicionamiento singular y una lectura crítica que le permitan al sujeto dar una respuesta.

En palabras de Jacques Lacan: "La ética consiste esencialmente en un juicio sobre nuestra acción" (Lacan, 1960). Todo el contexto del seminario muestra con claridad que no significa un juicio moral, sino que ese juicio al que se refiere indica una interrogación que vuelve sobre el sujeto, sobre su posición y sobre su deseo.

En función de esta idea, sostenemos una noción de ética reflexiva y conceptual (Salomone, 2021), que nos permita interrogar las prácticas, sus fundamentos, nuestra posición, desde herramientas conceptuales que transcinden la propia moral o la de los discursos que la enmarcan.

NOTAS

[i] Utilizamos aquí la terminología utilizada por Narciso Benbenaste (1993).

[ii] Región ubicada en el norte del estado de California, Estados Unidos, que incluye una decena de ciudades. El nombre "Silicon Valley" proviene de Silicon (silicio): el material base de los chips y microchips (circuitos integrados) en la industria informática.

[iii] En diferentes idiomas se verifica la elección de este mismo término: Algoritmos de decisión (portugués), algorithmes de décision (francés), algoritmi decisionali (italiano), decision-making algorithms (inglés).

BIBLIOGRAFÍA

- Benasayag, M. (2024). Asistimos a una delegación masiva de funciones del cerebro a las máquinas sin que haya tiempo para un reciclaje. Entrevista de Santiago Liaudat. *Revista Ciencia, Tecnología y Política*. vol. 7, núm. 12, 2024. Universidad Nacional de La Plata, Argentina. <https://revistas.unlp.edu.ar/CTyP/article/view/16846/16271>
- Benbenaste, N. (1993). *Sujeto = política x Tecnología / MERCADO*. Buenos Aires: Oficina de Publicaciones del C.B.C., Universidad de Buenos Aires.
- Bleichmar, S. (1999). Entre la producción de subjetividad y la constitución del psiquismo. En *Revista del Ateneo Psicoanalítico N° 2*, Buenos Aires.
- Castells, M. (2006). *La sociedad red: una visión global*. Madrid. Alianza Editorial.
- Fernández Zalazar, D. (2024). Transhumanismo: ¿la amplificación o el fin de lo humano? XVI Congreso Internacional de Investigación y Práctica Profesional en Psicología. *Memorias 2024*. ISSN 2618-2238. Universidad de Buenos Aires. Facultad de Psicología. pp. 47-50.
- Han, B.-C. (2014). *La sociedad de la transparencia*. Barcelona: Herder Editorial.
- Han, B.-C. (2016). *Psicopolítica: Neoliberalismo y nuevas técnicas de poder*. Barcelona: Herder Editorial.
- Jones, N. (2025) *Does using chatGPT affect brain activity? Study sparks debate*. Nature 643, 15-16 (2025) doi: <https://doi.org/10.1038/d41586-025-02005-y>
- Kosmyna, N. et al. (2025) *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task*. Preprint at rXiv <https://doi.org/10.48550/arXiv.2506.08872>
- La Rocca, S. (2025) *¿Qué puede hacer la ética con las subjetividades dataizadas por la IA?* Centro de Investigación sobre Sujeto, Institución y Cultura UNMdP @cisic.unmdp
- Lacan, J. (1953) Función y campo de la palabra y del lenguaje en psicoanálisis, *Escritos 1*. Buenos Aires: Amorrortu editores.
- Lacan, J. (1960) Seminario 7: La Ética Del Psicoanálisis, Buenos Aires: Paidós.
- Sadin, É. (2015) *La Vie algorithmique : Critique de la raison numérique*. Paris : Éditions L'Échappée.
- Sadin, É. (2018) *La silicolonización del mundo. La irresistible expansión del liberalismo digital*. Buenos Aires: Caja negra editora
- Salomone, G. Z. (2021) *Ética profesional en Salud Mental: campo normativo y dimensión clínica. Un aporte conceptual y metodológico a su estudio teórico, su aplicación práctica y su transmisión en la formación profesional*. Tesis doctoral. <https://repositoriouba.sisbi.uba.ar/>