



PROYECTOS DE INVESTIGACIÓN

Fallas en inteligencia

El ruido y los sesgos en el siglo XXI

Guillermo Tajan y Esteban Covelli

Un estudio que propone el análisis de la esencia de la inteligencia como una actividad histórica de carácter predictivo, donde el procesamiento de la información sirve para anticipar escenarios y enfrentar obstáculos no siempre libre de ruido. Y postula si la Inteligencia Artificial, aunque funcione mediante algoritmos lógicos, logra o no superar por completo los sesgos y problemas intrínsecos de la inteligencia tradicional, o si reproduce las limitaciones de la naturaleza humana. Exponiendo que la búsqueda por la objetividad pura sigue siendo un desafío pendiente.

FALLAS EN INTELIGENCIA

El ruido y los sesgos en el siglo XXI

Guillermo Tajan¹ y Esteban Covelli²

Resumen:

La Inteligencia es una actividad histórica, de carácter predictivo, que enfrenta constantes obstáculos. Su esencia intelectual choca con los sesgos del analista, lo que ha provocado fallos notorios. Los sesgos son atajos mentales que, junto al ruido, comprometen la objetividad del análisis.

La Inteligencia Artificial (IA) se postula como una solución imparcial, pero también es propensa a sesgos. Estos provienen de los datos de entrenamiento y de la influencia humana. En última instancia, la IA no logra superar los problemas de la Inteligencia, ya que su creación y operación siguen ligadas a las limitaciones de la naturaleza humana.

Palabras clave:

Inteligencia - Sesgos - Ruidos - Fallos - Inteligencia Artificial

¹ **Guillermo Martín Tajan.** Capitán de Navío (RE). Posee una licenciatura en Sistemas de Propulsión (Instituto Universitario Naval), una especialización en Análisis Estratégico (Instituto Universitario de la Fuerza Aérea Argentina) y estudios de Maestría en Metodología de la Investigación Científica (UB). Es director de la Especialización Gestión de la Información en los Espacios Marítimos y Fluviales y titular de la Cátedra de Fundamentos Metodológicos de la Inteligencia en el Instituto de Inteligencia de las Fuerzas Armadas (desde 2013), donde también es adjunto en la Cátedra de Inteligencia Militar. Sus principales publicaciones incluyen artículos sobre análisis estratégico, análisis mediante hipótesis competitivas (MAM), bioética, pensamiento complejo y formación militar, en revistas como la *EGN* y *Manual de Informaciones*. Es coautor del libro *Apuntes de Inteligencia Estratégica*. E-mail: guillermotajan@yahoo.com

² **Esteban Agustín Covelli.** Licenciado en Relaciones Internacionales por la Universidad Nacional de Rosario. Docente en la carrera de Relaciones Internacionales en la Universidad Austral. E-mail: esteban.covelli@gmail.com

“La raíz del problema –el eslabón más débil del proceso de inteligencia– es la naturaleza humana”.

Michael Handel

La inteligencia es una actividad que se realiza desde tiempos inmemoriales, pudiendo rastrearse antecedentes en el Antiguo Testamento, cuando Moisés envía a doce hombres a explorar la tierra de Canaán y que, en su regreso, advirtieron sobre los desafíos que el pueblo judío debía atravesar para llegar a su tierra prometida³. Así, brindando un panorama presente, la inteligencia tiene también su faceta predictiva o especulativa, buscando presentar imágenes no solo posibles, sino probables de un futuro de interés para quien debe tomar decisiones⁴.

Como se ha visto en el caso de los enviados de Moisés a Canaán, cuando estos brindaron detalles exagerados sobre los pueblos que observaron en sus viajes, la inteligencia es un ejercicio de fuerte contenido intelectual sujeto a un sinnúmero de condicionantes y factores que puedan alterarla, tanto propios o ajenos del sujeto. Por tal motivo, y para poder alcanzar un mejor producto, se requiere de una aplicación rigurosa y sin perturbaciones de la razón, que se encuentra atravesada.

Los diez espías hebreos, atravesados por el temor, usaron la presencia de guerreros de gran tamaño y las ciudades fortificadas para convencer al pueblo de que era imposible conquistar la tierra, contrariando la promesa y el poder de Dios. Esta relación entre la razón y las pasiones no es ninguna novedad. Platón ha representado esta tensión con el mito del carro alado entre la razón, representado por el caballo blanco; y las pasiones (que hoy llamaríamos

³ *Libro de los Números*, 13: 27-32.

⁴ La Armada Argentina ha adoptado para esa función el término especulativa como expresa en la publicación *Fundamentos Básicos de la Inteligencia de la Armada R.G-1-702*.

emociones), representado por el caballo negro⁵. Al practicar la disciplina de inteligencia, el analista se encuentra atravesado por esta problemática, en donde debe encontrarse vigilante de no caer en fallos propios de su razonamiento.

Así es como, en múltiples ocasiones a lo largo de la historia, han resultado estados de cosas que pudieron ser catalogados como fallos de Inteligencia; como el ataque a Pearl Harbor, el fracaso de Joseph Stalin a la hora de anticiparse a la invasión alemana de 1941, el llamado fracaso de la inteligencia israelí en vísperas de la guerra del Yom Kippur, el atentado de Las Torres Gemelas, etc.

Por estos motivos resulta tentador y necesario trabajar sobre los asesoramientos propiamente dichos como fenómenos, durante la actividad y en su evaluación *ex post*, como así también buscar lo que está realmente detrás para explicar los éxitos como fallas y buscar su mejora⁶.

Como se ha mencionado, las fallas en inteligencia pueden tener origen tanto en el interior del analista como desde el exterior. Uri Bar-Joseph y Jack S. Levy (2009) han adoptado una variación de distintas tipologías de las fuentes de los fallos de inteligencia, utilizando un marco modificado de niveles de análisis. El entorno informativo externo incluye hipótesis basadas en la falta de información, el exceso de información en forma de una baja relación señal-ruido y el engaño estratégico por parte del adversario. Otras hipótesis se centran en factores internos, como la psicología individual, la dinámica de los grupos pequeños, el comportamiento organizativo y la politización de la inteligencia.

⁵ El mito del carro alado aparece en el diálogo de Platón, *El Fedro*. Para explicar su visión del alma, va a recurrir a una alegoría en la que un auriga, que representa la parte racional del alma humana, conduce un carro tirado por dos caballos, uno bueno y uno malo, que vendrían a ser la parte irascible y la parte concupiscible. El auriga, o la razón, ha de llevar al ser humano hacia la verdad, hacia la luz, lo que no será fácil, teniendo en cuenta el conflicto entre los dos caballos, que quieren ir en direcciones opuestas, lo mismo que sucede en el interior del alma humana.

⁶ Lo subyacente, según Kant es el noúmeno, la cosa en sí, no lo que se muestra. Esta diferencia conceptual se en la *Crítica de la razón pura*, en lo que hace al conocimiento como en la *Crítica de la Razón Práctica* referida a la conducta.

Sobre las fallas internas, se destaca el concepto de “sesgos congñtidos”, definidos como errores mentales causados por la simplificación de las estrategias de procesamiento de información; y las contrasta con los sesgos culturales, organizativos o de propios intereses. Así, un sesgo cognitivo es un error mental consistente y predecible surgido de la forma en la que se procesa la información a nivel subconsciente. Este tipo de falla puede ser comparada con ilusiones ópticas, que convencen de manera irresistible al sujeto y, por tanto, son muy difíciles de superar (Heuer, 1999).

Para facilitar la comprensión de los sesgos, es conveniente sumarse a la visión de Rayner (1985), dando por necesaria la existencia de la conciencia de forma independiente del sustrato material (cerebro). Esta aproximación parece darnos la mejor perspectiva acerca de cierto escenario en el que es posible que ocurran cosas que no se controlan completamente, tanto en el plano consciente como por debajo de él.

En este sentido, hay un estrecho paralelismo entre la Inteligencia y la toma de decisión ya que la primera conlleva la realización de innumerables decisiones (selección de evidencias, su juicio de importancia, la selección de hipótesis, etc.). De este modo es posible nutrirse de los trabajos investigativos de los procesos de decisión y tenerlos en cuenta y/o aplicarlos a la Inteligencia, teniendo en cuenta que, actualmente, en ambas disciplinas existe un consenso acerca de la participación de la intuición (Gigerenzer, 2008; Lehrer, 2011) y de la emoción (Damasio, 2011).

Nos hemos acercado al ámbito propio de los procesos de los individuos, más allá de lo que desde afuera les llegue. En este amplio espectro de acciones existen algunos hitos clave, a nivel individual, que contribuyen al fracaso de la inteligencia como por ejemplo los procesos heurísticos cognitivos y los factores afectivos que influyen en el procesamiento de la información. Al mismo tiempo interactúan con ellos sistemas de creencias y personalidades individuales y lo hacen el contexto de tendencias más generales. La hipótesis, razonablemente acreditada, es que los analistas de inteligencia, como todas las personas, intentan actuar racionalmente, pero no lo consiguen debido a la influencia inconsciente que ejercen sobre ellos los heurísticos cognitivos y los factores afectivos que influyen en el procesamiento de la información.

Por este motivo, algunos autores, como Gilovich y Griffin (2002), distinguen entre sesgos inmotivados y motivados. Los sesgos inmotivados o cognitivos se refieren a la influencia del sistema de creencias de un individuo y las estrategias simplificadoras que utiliza para dar sentido a un mundo complejo y ambiguo, independientemente de los intereses políticos o las influencias emocionales. Uno de los argumentos centrales es que la percepción es un proceso basado en teorías, y que las creencias previas se sobreponen a la nueva información en el proceso de juicio. En consecuencia, la gente tiende a ver lo que espera ver.

Otros elementos reconocibles hallazgos clave son que las personas muestran más confianza en sus juicios de la que justifican los datos, permiten que sus evaluaciones de probabilidad se vean desproporcionadamente influidas por imágenes vívidas de sucesos pasados, actualizan sus creencias sólo lentamente y sobreponderan el impacto de probabilidades pequeñas (Kahneman, 2012).

A menudo existe una delgada línea entre el “sesgo motivado” inconsciente y la acción consciente y deliberada para distorsionar la inteligencia, así se agrega una dificultad adicional para reconocer y mitigar la acción de los sesgos inconscientes.

Asimismo, los sesgos, que son materia de trabajo en centenas de publicaciones, pueden ser categorizados según la clasificación ofrecida por Heuer (1999):

- Sesgos en la evaluación de la evidencia:
 - Impacto vivencial.
 - Ausencia de evidencia.
 - Exceso de sensibilidad a la coherencia.
 - Persistencia de impresiones basadas en evidencias descalificadas.
- Sesgos en la percepción de causa y efecto:
 - Explicaciones causales.
 - Percepción de una dirección centralizada.
 - Similitud entre causa y efecto.
 - Causas internas vs externas.

- Sobreestimar nuestra propia importancia.
- Correlación no significa causalidad.
- Sesgos en la estimación de probabilidades:
 - Regla de accesibilidad.
 - Anclaje.
 - Expresión de incertidumbre.
- Sesgos retrospectivos en la evaluación de informes de Inteligencia:
 - La perspectiva del analista (sobreestimación).
 - La perspectiva del usuario (subestimación) cognitivas.
 - La perspectiva del supervisor (autopsia).

Vale la pena enfatizar el hecho de que estas fallas son involuntarias e independientes de convicciones o deseos. Se dan en el marco de las estrategias para disminuir el costo de una permanente evaluación, utilizando para ello “atajos”. Según Bachrach (2012), todos los caminos dentro del cerebro están disponibles pero la energía fluye por aquellas llamadas “avenidas iluminadas”, que, siguiendo una metáfora del cerebro como una ciudad, representan los caminos mentales que recorremos con mayor frecuencia, aquellos que están bien conectados y activos gracias a la repetición, la experiencia y la cultura. Es decir, son los circuitos neuronales más utilizados y consolidados en nuestro cerebro.

Ruido, ¿un caso particular de los sesgos?

Por su parte, al momento hemos hecho referencia a los sesgos, como circunstancias que son inherentes al analista, que ocurren en su interior, más allá que pueden tener su inicio en el exterior. En cambio, al referirnos al ruido, este puede ocurrir tanto en el exterior del conjunto de analistas, manifestado como una perturbación en la cantidad y calidad de la información que obtenemos y se caracteriza por una mala relación señal/interferencia. En cambio, el “otro” ruido es producido por el propio analista o el conjunto de analistas y se reconoce en la dispersión de los resultados. Así las cosas, un error puede confundirse con “solo” un sesgo. Buscando más claridad se puede establecer una analogía.

Gráficamente se puede comprender porque algunos creen ver al ruido como un sesgo particular, aunque más adelante se harán notables algunas diferencias⁷.

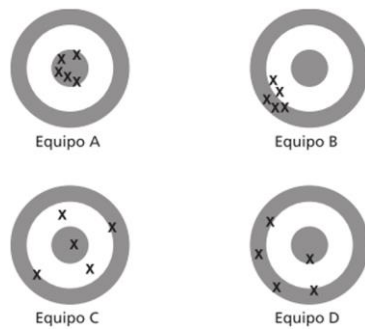


FIGURA 1. Cuatro equipos

Kahneman, Siboni y Susteín (2021) nos ayudan a ilustrar el concepto de ruido a partir de blancos. Los autores nos proponen imaginar cuatro equipos en un campo de tiro, cada uno formado por cinco personas que comparten un rifle y realiza cada integrante un disparo. En un mundo perfecto y sin errores, todos darían en el blanco. El primer equipo logra agrupar los disparos próximos al centro en un patrón perfecto. El Equipo B, en cambio, se lo observa “sesgado”, ya que mantiene los disparos agrupados, pero fuera del objetivo. Esto es debido a la consistencia del sesgo, pudiendo predecir el resultado de los disparos hacia la misma zona. Esta circunstancia ofrece también una relación de causalidad o explicación causal, como por ejemplo puede ser una falla en la mira del rifle desvía los disparos hacia ese sector.

El Equipo C, en cambio, ofrece un resultado “ruidoso” al encontrarse los disparos muy dispersos. No se puede evidenciar un sesgo en particular ya que forman una rosa irregular. Al no poder definir una tendencia predeterminada (sesgo) y tener semejante dispersión, resultaría muy difícil poder predecir dónde impactaría el siguiente tiro. No es fácil generar una hipótesis. Desconocemos por qué son tan ruidosos. El equipo D está sesgado y es ruidoso. Sus disparos se desvían sistemáticamente del objetivo y están muy dispersos.

⁷ Gráfico de Kahneman, Siboni y Susteín (2021).

El sesgo y el ruido –desviación sistemática y dispersión aleatoria– son componentes diferentes del error. Los blancos ilustran la diferencia. El campo de tiro es una analogía de lo que puede fallar en el juicio humano, en particular en las diversas decisiones que se toman en nombre de organizaciones o grupos pluripersonales. En estas situaciones, nos encontraremos con las dos clases de error ya referidas. Unos juicios están sesgados; se desvían sistemáticamente del objetivo. Otros son ruidosos, ya que las personas que se espera que estén de acuerdo terminan en puntos muy diferentes, aunque en los alrededores del objetivo. Por desgracia, muchas organizaciones se ven afectadas tanto por el sesgo como por el ruido, aunque una distinción fundamental es que el ruido se puede reconocer y medir sin saber nada sobre el objetivo o el sesgo, es decir solo por los valores o posiciones relativas de los resultados.

Los sesgos y el ruido son problemáticas recurrentes en distintos sectores y ámbitos. Estos pueden ocurrir en toda instancia de toma de decisiones como también en proyectos investigativos, como el presente. En este sentido, se pueden encontrar aportes académicos que abordan a los sesgos y al ruido desde distintas aproximaciones, ya sea desde lo teórico o mediante revisiones de prácticas; desde análisis de un sistema de inteligencia propiamente dicho a la decisión que lleva un cuerpo judicial o médico con respecto a un caso o varios casos semejantes.

Así, Beck (2011) aporta una base conceptual que integra diversas desviaciones cognitivas con el objetivo de demostrar cómo mecanismos basados en el ruido pueden llevar al conservadurismo, sesgos bayesianos, correlaciones ilusorias, sesgo de autocomplacencia, expectativa exagerada, entre otros. Este marco teórico resulta útil para entender los fundamentos cognitivos que influyen en el análisis de inteligencia.

Por otro lado, Blumenthal-Barby y Krieger (2014), si bien trabajan sobre el ámbito médico, ofrecen un enfoque metodológico riguroso para identificar y clasificar heurísticas y sesgos. La adaptación de estas estrategias al campo de la inteligencia subraya la posibilidad de transferir metodologías interdisciplinarias para enriquecer el análisis en contextos de alta incertidumbre.

Desde una perspectiva práctica y específica en los sistemas de inteligencia, existen estudios sobre cómo el error inherente puede afectar la evaluación y la agregación de juicios de distintos analistas, destacando la importancia de incorporar mecanismos que permitan mitigar el ruido y compensar la influencia de sesgos individuales, sugiriendo que la integración de sistemas tecnológicos puede mejorar la robustez de las evaluaciones estratégicas.

En el ámbito militar, Williams (2011) examina cómo las heurísticas, si bien ofrecen respuestas rápidas ante escenarios complejos y críticos, frecuentemente conducen a errores sistemáticos en la toma de decisiones. Su análisis es relevante puesto que resalta la doble vertiente de estas estrategias cognitivas: funcionales en situaciones de alta presión, pero potencialmente perjudiciales cuando introducen sesgos que distorsionan el juicio estratégico.

Payá Santos y Delgado Morán (2017) discuten cómo funciona la mente humana en los análisis de inteligencia, estando en situaciones de incertidumbre, necesidad de valorizar las pruebas realizando juicios bajo presión y tiempo, factores que inciden en la calidad de la información procesada. Este estudio recalca la necesidad de reconocer y manejar tanto el ruido como los sesgos en ambientes donde la precisión es esencial.

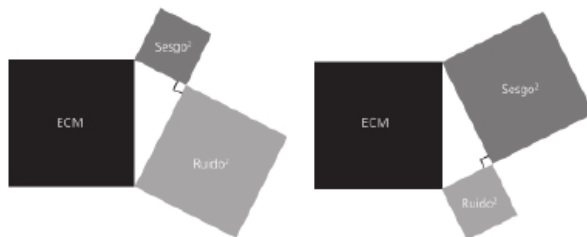
Por su parte, Welton Chang *et al.* (2018) proponen replantear las técnicas analíticas convencionales, enfatizando que un enfoque unidimensional para la mitigación de sesgos podría, inadvertidamente, intensificar la variabilidad (ruido) en ciertos contextos. Su propuesta de reestructuración aboga por la incorporación de métricas cuantitativas que permitan evaluar simultáneamente la presencia de sesgos y la variabilidad en los procesos analíticos, abriendo el camino a un diseño más equilibrado que combine rigor analítico y adaptabilidad al entorno.

Este estado del arte presentado evidencia la complejidad inherente a la toma de decisiones en entornos inciertos y dinámicos. La literatura revisada abarca desde fundamentos teóricos que explican la base cognitiva de estos errores hasta propuestas innovadoras en técnicas analíticas orientadas a su mitigación. Este panorama multidimensional sugiere la necesidad de continuar explorando estrategias que integren soluciones tecnológicas y

metodológicas, encaminadas a mejorar la precisión y confiabilidad del análisis en inteligencia.

En lo referente a Ruido es muy valioso el trabajo realizado por Kahneman, Siboni y Sustain (2021) sobre el concepto del ruido en lo que hace a la toma de decisiones. Los autores definen al ruido como la variabilidad no deseada en juicios que deberían ser idénticos. Particularmente esto es en los juicios, entendiéndose que, allí el desacuerdo admisible tiene un límite y diferenciándose de las opiniones o gustos. El ruido, entonces, puede entenderse brevemente como la falta de constancia en los juicios realizados tanto como una persona como por parte de una organización. Se muestra entonces en este trabajo cómo el ruido, siendo menos conocido, ejerce una influencia significativa. Para mostrar esto recupera “método de los mínimos cuadrados” de Gauss, con su “error cuadrático medio” (ECM) como una herramienta propuesta para ponderar los errores, relacionando el ruido y con el sesgo, ambos factores componentes del ECM.

Ecuación del error (Kahneman, 2021).



$$\text{Error global (ECM)} = \text{sesgo}^2 + \text{ruido}^2$$

El ruido se puede dividir en categorías. Una de estas es el *ruido (error) de nivel*, la cual se entiende como una variación entre diferentes agentes al momento de dictar un juicio, teniendo estas variaciones en otras características, como los antecedentes, experiencias, opiniones, sesgos, etc., funcionando como un rasgo de personalidad. Por otro lado, el *ruido de patrón* indica la existencia de otras diferencias que reflejan un patrón complejo en las actitudes de quienes juzgan o deciden, en casos particulares. Se refiere a las desviaciones sistemáticas en los juicios de una persona o grupo, en situaciones similares. Este nace también

de diferencias sistemáticas en la capacidad de emitir juicios válidos sobre diferentes dimensiones de un caso.

Por su parte, el ruido se encuentra compuesto por el *ruido de ocasión*, siendo una variabilidad intrapersonal, a partir de un contexto favorable o desfavorable del actor y que surge de las posibilidades no vistas. Respecto a esta última, hay distintas maneras de medir su existencia al comprobar si las decisiones se vieron influidas por factores irrelevantes para la ocasión. Este ruido tiene como fuentes el estado de ánimo, el estrés y cansancio, el clima, el orden de los casos (Kahneman, Sibony y Sunstein, 2021).

La IA, ¿una esperanza que nace?

Ahora bien, siendo los sesgos (y el ruido) fenómenos complicados de analizar por su naturaleza inconsciente y subjetiva, es natural buscar soluciones objetivas e imparciales para contrarrestar este tipo de problemas. Aquí surge la posibilidad de contar a la Inteligencia Artificial (IA) como una herramienta posible.

Los sesgos, en cualquier actividad o disciplina humana, son preconceptos, prejuicios, es decir formas de percibir la realidad que nos llevan a organizar nuestras cadenas de razonamientos y por lo tanto nuestras decisiones y su influencia es meramente inconsciente. Podemos interpretarlos como un mecanismo evolutivo/adaptativo que resulta de gran utilidad para la supervivencia al acortar los tiempos de respuesta en un sinnúmero de situaciones transformándolo en un mecanismo natural de pensamiento. Es decir que forma parte de los mecanismos de defensa de nuestra mente primitiva, cuyo sentido es protegernos de posibles peligros en base al conocimiento previo.

Ahora bien, si bien la IA no deja de ser una máquina que funciona a partir de algoritmos que pueden ser considerados objetivos y carentes de todo tipo de pasión, independizándose del problema del carruaje de dos caballos; esta no se encuentra exenta de caer en sesgos. Esto es debido a que la IA cuenta con sesgos que no son generados por la máquina en sí, sino que provienen de los humanos: desarrolladores, usuarios y los datos que

alimentan los algoritmos. La IA hereda y amplifica prejuicios sociales ya existentes (Nadal Saifán, 2025).

Autores como Saifán (2025) o González (2024) entienden que existen una cierta cantidad de estos sesgos cuyas fuentes pueden ser identificadas y tipificadas:

1. **Sesgos contextuales:** resultan de la falta de comprensión del contexto en el que se utilizará la IA. Un algoritmo puede funcionar bien en un entorno específico, pero no en otros, lo que lleva a resultados parciales o erróneos.
2. **Sesgos algorítmicos:** tienen su fuente en el diseño y la implementación de los algoritmos. Pueden surgir debido a la elección de características o variables que pueden influir en la toma de decisiones y crear parcialidades no intencionadas en los resultados.
3. **Sesgos de confirmación:** se refieren a la tendencia de los algoritmos a favorecer o dar más peso a la información que confirma las expectativas preexistentes, lo que puede llevar a un ciclo de retroalimentación y reforzar los sesgos existentes.
4. **Sesgos de datos:** es el resultante del universo utilizado en el entrenamiento de un algoritmo. Presentan desigualdades la IA puede aprender y perpetuar esas desigualdades en sus resultados.
5. **Sesgos de interpretación:** surgen cuando los humanos interpretan los resultados de los modelos de IA de manera sesgada, favoreciendo información que confirma sus creencias preexistentes o malinterpretando la salida de la IA.
6. **Sesgos de omisión:** pueden ocurrir cuando los datos utilizados para entrenar la IA no incluyen ciertos grupos o situaciones, lo que resulta en la omisión de información relevante y, por lo tanto, en decisiones parciales o inexactas.
7. **Sesgos de retroalimentación:** se producen cuando las decisiones basadas en los resultados de la IA afectan el conjunto de datos utilizado para

entrenarla, creando un ciclo de retroalimentación que amplifica los sesgos originales.

8. **Sesgos de retroalimentación:** se producen cuando las decisiones basadas en los resultados de la IA afectan el conjunto de datos utilizado para entrenarla, creando un ciclo de retroalimentación que amplifica los sesgos originales.
9. **Sesgo de medición:** esto ocurre cuando la forma en que se recopilan o miden los datos introduce sesgos. Esto puede conducir a resultados sesgados. Por ejemplo, si un sistema de reconocimiento de voz está entrenado principalmente en habla clara y lenta, puede tener un rendimiento deficiente al transcribir voz rápida o con acento.
10. **Sesgo lingüístico:** el sesgo lingüístico se refiere a los sesgos inherentes al lenguaje o su uso, que reflejan y refuerzan sesgos culturales, sociales o cognitivos. Los modelos lingüísticos entrenados con datos sesgados probablemente replicarán estos sesgos. Además, el idioma en el que se diseña el modelo juega un papel relevante en su eficiencia. . La dependencia de la traducción automática exacerba aún más estos desequilibrios, ya que alrededor del 50% de las fuentes están en inglés, mientras que el siguiente idioma más común solo comprende alrededor del 5%. Esta brecha significativa también se puede utilizar para orientar la IA de manera poderosa a favor de países/poblaciones o suministros tecnológicos específicos.

Si bien no es materia del presente artículo un abordaje profundo de los inconvenientes en los que se puede incurrir al utilizar la IA en un análisis de inteligencia, lo reseñado basta para alertar sobre nuestras percepciones e ideas acerca de la sustitución de la actividad humana por la aplicación de este tipo de herramientas.

Aunque hay pocas dudas acerca de algunos de los aspectos beneficiosos de su uso, parece igualmente indubitable el hecho de que no podremos encontrar en esta

herramienta la ecuanimidad y objetividad ascética que parece no estar disponible en el juicio humano. Aquellas herramientas como la IA, que no se sienten como propias, al igual que ocurre con artilugios cotidianos o las técnicas; producen en nosotros la tendencia a considerarlas “ajenas” sin tener en cuenta que en ellas hay mucha humanidad, en tanto son creaciones de y para el ser humano⁸. Fueron concebidas y desarrolladas para ayudar al hombre y están impregnadas de muchas de las limitaciones de su creador.

En otras palabras, a pesar que la aparición de la IA en el siglo XXI se presenta como una instancia computarizada y carente de emociones u factores externos que puedan afectar la objetividad de su juicio, esta no escapa al hecho de que su programación de origen como su utilización es realizada por sujetos que sí se mantienen afectados por los sesgos y ruidos.

En este sentido, los sesgos cognitivos y su relación con la IA, las opiniones de Jessica Nordell (2022) y Kate Crawford (2023) representan las tendencias acerca de cómo evaluar la IA en general y sus productos en particular.

Para Nordell, los sesgos cognitivos funcionan como un circuito: comienzan cuando adquirimos conocimientos culturales del mundo que nos rodea mediante toda clase de experiencias (familia, medios de comunicación, profesores, redes sociales, etc.).

Con el tiempo, estos conocimientos (si tiene cierta constancia y sesgo) se transforman en estereotipos arraigados. Se constituyen en un modo de pensar irracional e involuntario. Cuesta mucho esfuerzo reconocer los propios sesgos. Los sesgos, producen distancia entre la palabra y los hechos, es decir, producen inexactitudes, falsedades o desinformación.

Por su parte, Crawford expresa que los sesgos son un síntoma de aflicción más profunda: una lógica normativa centralizadora y de largo alcance que se usa para determinar en qué se tendría que mirar y en cómo evaluar el mundo.

⁸ Acerca de ello se extiende Gilbert Simondon en *El modo de existencia de los objetos inanimados*, donde se centra en las consecuencias de esa fuerte externalización de los productos creados por el hombre.

Para Crawford, la IA no es ni I ni A. Es una forma material, hecha de recursos naturales, combustible, mano de obra, infraestructuras, logísticas, historias y clasificaciones. Los sistemas de IA no son autónomos, racionales ni capaces de discernir algo sin un entrenamiento intenso, con errores, conjunto de datos o reglas y recompensas predefinidas. También, insiste en que la IA tiene implicaciones políticas. Hay preguntas válidas: 1) qué se está optimizando, 2) para quién y 3) quién toma las decisiones.

Recordemos que la IA fue alimentada por una gran mayoría de varones, blancos y, fundamentalmente, ingenieros (la ONU informa que solo el 22% de quienes trabajan en IA son mujeres). El *software* siempre muestra la esencia del grupo que lo creó, sostiene Nordell. Entonces, este nace con los sesgos propios de etnia, género, edad, profesión, que constituyen los más frecuentes cuando usamos cualquier tipo de IA y que también son los más frecuentes en la vida cotidiana, privada y pública.

Algunas reflexiones

El juicio, tal como ha sido expuesto aquí, no es lo mismo que pensar. Es algo más concreto: una forma de medir, con una mente humana como instrumento. Igual que cualquier otra medición, el juicio asigna un valor –no necesariamente un número– a lo que observa.

Decir “el tumor de Mary Johnson es probablemente benigno” es un juicio, igual que afirmar que la economía está inestable o que cierto candidato sería el mejor para un puesto.

No son cálculos ni fórmulas cerradas: son evaluaciones globales que integran fragmentos de información. Un profesor los usa al corregir una redacción, pero no para calificar un examen múltiple.

La vida cotidiana está llena de juicios profesionales: médicos, entrenadores, empresarios, ingenieros y analistas de inteligencia. Todos deciden, y esas decisiones impactan, de manera distinta, pero con regularidad e intensidad en sus correspondientes ámbitos.

Algunos juicios son predictivos –y en ciertos casos comprobables–, aunque muchas veces, como en inteligencia (cuando los planteos son hipotéticos) y/o en predicciones de largo plazo, no hay forma de verificar su corrección. En esos casos, solo podemos evaluar la calidad del razonamiento detrás. Otros juicios son evaluativos: un jurado de clavados, un crítico de arte. No existe un “valor verdadero” con el que compararlos, aunque solemos comportarnos como si lo hubiera.

Ahora bien, los juicios fallan. Cuando los errores van siempre en la misma dirección hablamos de sesgo; cuando los errores son desordenados y se dispersan hablamos de ruido. Ambos distorsionan.

El sesgo es más fácil de identificar; el ruido, en cambio, se esconde en la variabilidad de juicios que deberían ser iguales. Por eso se pueden y quizás deben realizar “auditorías de ruido”: se pide a distintos profesionales que evalúen los mismos casos y se mide la dispersión, como ver los disparos en un blanco. A veces esas auditorías revelan carencias de formación; otras muestran incoherencias estructurales.

La pregunta clave es: ¿qué daña más, el sesgo o el ruido? Depende. Lo cierto es que ambos suman al error total, y que reducir cualquiera de las dos mejoras la calidad de las decisiones.

Con respecto a las herramientas que nos ayudarían a superar a los sesgos y ruido, particularmente la IA, se debe tener presente que las tecnologías no son neutrales ni puramente objetivas, en tanto que los sesgos humanos se incrustan en los sistemas. Esto puede ser por una transferencia directa de la raíz cultural y objetiva del creador humano; o por la mera lógica de poder que ordena al mundo, con su distribución estructural y material. Entonces, se pueden llegar a ver en el uso de cualquier tipo de IA los sesgos propios del hombre.

Para concluir, se rescata la reflexión de Guadalupe Nogués (2024), que nos explica que un sesgo cognitivo provoca defectos en el razonamiento. Una falacia es un argumento no válido que aparenta serlo. Estos dos conceptos van muchas veces de la mano, ya que muchas falacias ‘funcionan’ porque aprovechan sesgos que tenemos.

Presentaciones falaces, *deep fakes*, sesgos cognitivos conscientes o no llevan, hoy más que nunca, a la necesidad de debates, en el caso de la IA acerca de sus capacidades y debilidades, para evitar la desinformación.

El objetivo es siempre el mismo: lograr juicios más consistentes, con menos sesgo y menos ruido.

REFERENCIA BIBLIOGRÁFICA

- Abbagnano, N. (1982). *Historia de la Filosofía*. Hora.
- Bachrach, E. (2013). *Ágilmente*. Sudamericana.
- Baidot, N. (2010). *Neuromanagment*, Granica.
- Beck, J. S. (2011). *Cognitive behavior therapy: Basics and beyond* (2.^a ed.). Guilford Press.
- Bennett, D. J. (2000). *Aleatoriedad*. Alianza Editorial.
- Blumenthal-Barby J. y Krieguer H. (2014). Cognitive Biases and Heuristics in Medical Decision Making: A Critical Review Using a Systematic Search Strategy. *Medical Decision Making*. <https://www.researchgate.net/publication/264990410>
- Cal, C. et al. (2019). *La Cuestión Estratégica*. Escuela de Guerra Naval.
- Carpio, A. (1995). *Principios de Filosofía*. Glauco.
- Clauser, J. (2008). *An introduction to intelligence research and analysis*. Scarecrow professional intelligence education series, N.º 3. The Scarecrow Press.
- Cohen Agrest, D. (2008). *¿Qué piensan los que no piensan como yo?* Sudamericana.
- Damasio, A. (2011). *El error de Descartes*. Paidós.
- Crawford, K. (2023). *Atlas de la inteligencia artificial. Poder, política y costos planetarios*. Fondo de Cultura Económica.
- Evers, K. (1980). *Neuroética*. Katz.
- Freud, S. (1980). *Proyecto de una Psicología para Neurólogos y otros escritos*, Alianza.
- Gigerenzer, G. y Brighton, H. (2009), *Homo Heuristicus*, “*Why Biased Minds Make Better Inferences*, Max Planck Institute for Human Development”, Austin, Topics in Cognitive Science, 1, 1.
- Gigerenzer, G. (2008). *Decisiones Instintivas: La inteligencia del Inconsciente*. Ariel.

- Gilovich, T. y Griffin, D. (2002). Heuristics and Biases: Then and Now. En T. Gilovich, Griffin, D. Kahneman (Eds.), *Heuristics and Biases. The Psychology of Intuitive Judgment*. Cambridge University Press.
- Goleman, D. (1996). *La Inteligencia Emocional*. Javier Vergara Editor.
- González, M. H. (05/09/2024). *Sesgos en la Inteligencia Artificial Generativa: un análisis exhaustivo*. CISA, CRISC. <https://ciberprisma.org/2024/09/05/sesgos-en-la-inteligencia-artificial-generativa-un-analisis-exhaustivo/>
- Handel, M. I. (1977). The Yom Kippur War and the inevitability of surprise. *International Studies Quarterly*, 21(3), 461-502.
- Han, Byung Chuln (2013). *La sociedad de la transparencia*. Herder Editorial.
- Heuer, R. J. (1999). *Psicología del Análisis de Inteligencia, Centro para el Estudio de la Inteligencia*. Agencia Central de Inteligencia, Estados Unidos.
- Heuer, R. y Pherson, R. (2011) *Structured analytic techniques for intelligence analysis*. CQ Press.
- Kahneman D. (2016). *Pensar Rápido, Pensar Despacio*. Debate.
- Kahneman, D., Sibony O. y Sustain, C. (2021). *Ruido*. Penguin Random House.
- Kant, I. (1928). *Crítica de la Razón Pura*, V. Suárez.
- Lehrer, J. (2011). *Cómo decidimos y cómo tomar mejores decisiones*. Paidós.
- Morin, E. (2002). *La mente bien ordenada* (2.^a ed.). Seix Barral.
- Morris, D. (1968). *El mono desnudo*. Plaza y Janés.
- Morris, D. (2009). *El hombre desnudo*. Planeta.
- Nadal Saifán, I. (17/01/2025). *Sesgos de la IA: Más humanos que artificiales*. Comisión de Estudios Institucionales (CEEII). <https://ceeii.ar/scripta/sesgos-de-la-ia-mas-humanos-que-artificiales>
- Navarro Bonilla, D. (s.f.). *Cuadernos Constitucionales de la Cátedra Fadrique Furio Ceriol*, (48). Universidad Carlos III Madrid.

- Nietzsche, F. (1994). *Opiniones y sentencias diversas*. Editores Mexicanos Unidos.
- Nogués, G. (2024). *Pensar con otros: Una guía de supervivencia en tiempos de posverdad*. El Gato y la Caja.
- Nordell, J. (2022). *El fin del sesgo*. Tendencias.
- Payá Santos, C. A. y Delgado Morán, J. J. (2017). Incertidumbres del análisis dimensional de la inteligencia/ Uncertainty of dimensional analysis of intelligence. URVIO. *Revista Latinoamericana de estudios de seguridad*, (21), 225-239. <https://doi.org/10.17141/urvio.21.2017.2962>
- Popper, K. R. (1999). *El mundo de Parménides: Ensayos sobre la ilustración presocrática*. Paidós.
- Popper, K. R. (2008). *La lógica de la investigación científica*. Tecnos.
- Rayner, C. (1985). *El cuerpo y la mente*. Folio.
- Searle, J. (2006). *La mente*. Norma.
- Smith, J. E. (s.f.). Irracionalidad. Fondo de Cultura Económica.
- Somiedo, P. (s.f.). *El papel de la epistemología en el análisis de inteligencia*. Intel Times. <https://www.researchgate.net/publication/277565787>
- Taleb, N. N. (2009). *¿Existe la suerte?* Paidós.
- Taleb, N. N. (2013). *Antifrágil*. Paidós.
- Thaler, R. H. (2017). *Portarse mal*. Paidós.
- Uri Bar, J. y Levy, J. S. (2009). Conscious Action and Intelligence Failure. *Political Science Quarterly*, 124(3).
- Welton Chang, E., Berdini, D., Mandel, R. y Tetlock, P. E. (2017). *Restructuring structured analytic techniques in intelligence, Intelligence and National Security*. https://www.researchgate.net/publication/321006524_Restructuring_structured_analytic_techniques_in_intelligence
- Whaley, B. (1973). *Codeword Barbarossa*. MIT Press.

Williams, B. (2011, enero-febrero). Heurísticas y sesgos en la toma de decisiones militares.
Military Review.

Wohlstetter, R. (1962). *Pearl Harbor: warning and decision*. Stanford University Press.