

---

HIC SVNT DRACONES

# FALACIAS ARGUMENTALES EN LA ERA DE LA IA

---

Mariano Boiero

JUN · MMXXVI

---

HICSUNTDRACONES.AR

---

El repertorio canónico es conocido. Las falacias que más aparecen en el debate público — las clásicas — son: ad verecundiam (autoridad), ad hominem (ataque a la persona), petición de principio (conclusión como premisa), generalización apresurada, hombre de paja (atacar una versión distorsionada del argumento), non sequitur (conclusión que no se sigue), falso dilema, pendiente resbaladiza, anécdota universalizante. Veamos desde dónde vienen.

---

## I. LA TRAMPA COMO ESTRATEGIA

Las primeras clasificaciones de falacias no nacen para describir errores involuntarios. Nacen para identificar maniobras. Aristóteles ya distinguía falacias de ambigüedad, de suposición y de irrelevancia — no como accidentes del pensamiento, sino como dispositivos para desplazar el debate, para ganar una discusión sin tener razón. La falacia no era un tropiezo: era una táctica.

En las disputas escolásticas medievales eso se profundizó. Confundir al interlocutor, manipular premisas, introducir supuestos sin evidencia, acumular autoridad sin argumentos: todo formaba parte del arsenal dialéctico. La falacia era deliberada. Y era efectiva.

Con la Ilustración y el surgimiento de la esfera pública moderna, las falacias dejaron de ser un asunto de lógica y se convirtieron en un instrumento político. Servían para legitimar poder, justificar desigualdades, construir narrativas que simplificaban fenómenos complejos hasta volverlos manejables. La falacia se volvió persuasión de masas.

El siglo XX la industrializó. La radio, la televisión y la publicidad transformaron las falacias en slogans, generalizaciones, causalidades mágicas, equivalencias falsas y apelaciones emocionales diseñadas para escalar. Goebbels lo sabía con precisión clínica: la mentira repetida suficientes veces se convierte en verdad aceptada. La falacia dejó de ser un error lógico y se convirtió en una industria comunicacional.

---

## II. LO QUE HICIERON LAS REDES SOCIALES

Las redes sociales no inventaron falacias nuevas. Hicieron algo más interesante: las industrializaron a escala individual. Cualquier persona con conexión y cuenta puede producir y viralizar una falacia con la misma eficacia que antes requería un aparato de propaganda.

---

¿Por qué funciona? Porque la falacia es breve, emocional, compartible, fácil de producir y fácil de consumir. Se convierte en la unidad mínima del contenido viral. El algoritmo no premia la coherencia: premia el impacto. Y la falacia impacta mejor que el argumento bien construido, que es más lento, más largo y más difícil de procesar en un scroll.

Tres transformaciones concretas emergieron de ese ecosistema.

La primera: la velocidad destruyó el escrutinio. El formato de Twitter/X premia la frase que llega primero. Una afirmación sin fuente que circula antes que su refutación ya ganó, independientemente de su valor de verdad. El *ad ignorantiam* y la generalización apresurada se volvieron estructurales, no excepcionales. "Todos los estudios muestran que el trabajo remoto destruye la productividad." Sin fuente. Catorce mil likes.

La segunda: el algoritmo seleccionó las falacias más emocionalmente eficaces. Las expresiones sin sustento se repiten hasta volverse verdades aceptadas, no por rigor sino por frecuencia. Una frase falsa sobre salarios docentes compartida cuarenta mil veces en cuarenta y ocho horas sobrevive mucho más que la corrección que llega dos días después con cien retuits.

La tercera: el *ad hominem* mutó. Apareció una variante que no ataca a la persona sino a su circunstancia: el interlocutor queda situado en un punto ideológico, de clase, de origen o de experiencia vital desde el cual se declara ilegítima su opinión. "Vos no podés opinar sobre pobreza porque sos de clase media." El argumento desaparece. Queda solo la descalificación del hablante.

---

### III. LO QUE INTRODUCE LA IA

La IA no destruyó el debate público. Lo equipó con una muleta más sofisticada. Los grandes modelos de lenguaje operan por predicción estadística, no por comprensión. Ensamblan texto basándose en qué secuencias de palabras aparecieron juntas en sus datos de entrenamiento, no en ninguna comprensión de lo que esas palabras significan.

Eso crea un problema epistémico genuinamente nuevo: la plausibilidad sin verdad. Un output puede sonar riguroso, técnico, bien fundamentado — y no tener ningún anclaje empírico real.

Lo que hacen las personas en debates públicos es usar esa plausibilidad como si fuera evidencia. De ahí emergen variantes de falacias clásicas que tienen su propia lógica y su propio mecanismo.

---

El Oráculo de Silicio (*ad verecundiam*): La autoridad clásica tiene creencias, intenciones y responsabilidad epistémica. La IA no tiene ninguna de las tres. Apelar a ella como autoridad es invocar una fuente sin sujeto. "Lo pregunté a los tres motores de IA y los tres coinciden." Los tres reprodujeron el patrón estadístico más frecuente en su corpus. No condujeron ningún estudio. Sonaron seguros porque así está diseñado su output.

La Inferencia Huérfana (*non sequitur*): La IA puede producir una conclusión que suena técnica pero no se sigue de ningún dato real. La conclusión existe; las premisas que la justificarían, no. "De los datos de PISA se puede inferir que un joven de 18 en LATAM equivale a uno de 13 en Singapur." PISA no produce ese tipo de equivalencia. La IA construyó una conclusión verosímil sobre datos que no la contienen.

El Espejo que Asiente (*petición de principio*): La IA fue entrenada con texto humano. Cuando se le pregunta algo, reproduce los patrones más frecuentes en ese corpus. Si una idea ya circulaba ampliamente, la IA la confirma — no porque sea verdad, sino porque es estadísticamente dominante. "Le pregunté a ChatGPT si el liderazgo transformacional es el más efectivo y me dijo que sí." ChatGPT repitió lo que miles de artículos de management anglosajón dicen desde hace treinta años. No evaluó evidencia. Devolvió la premisa con mejor prosa.

La Omnisciencia Delegada: "No necesito contexto, la IA ya sabe." Pero la IA generaliza sobre lo frecuente, no sobre lo específico. Sin contexto, produce output verosímil sobre el caso promedio. El caso particular — esta empresa, esta cultura organizacional, esta historia específica — no existe en ningún corpus.

El Sicario de Fuente: Ataca el instrumento presunto para evadir el argumento. Si el argumento es sólido, el origen es irrelevante. Pero en el debate de redes, la sospecha de autoría algorítmica basta para descartarlo sin refutarlo. El contenido sobrevive intacto; el debate, no.

El Doble Escudo: Combina dos movimientos contradictorios. Cuando la evidencia falla: "es solo mi opinión." Cuando quiere peso: "la IA lo confirma." Los dos escudos se activan en momentos distintos, se contradicen entre sí y ninguno asume la carga de la prueba.

El Paracaídas de IA: La afirmación original no tenía fuente. Queda expuesta. Entonces se reencuadra retroactivamente: "La metodología es inferencia de IA." La IA no era la fuente desde el principio. Apareció como metodología de rescate cuando el edificio ya se caía.

El Testigo Universal: Anécdota personal más "la IA coincide" igual a estadística. La IA reproduce lo frecuente en su corpus. Si un estereotipo circuló suficientemente en millones de textos, la IA lo confirmará. No valida la anécdota: la amplifica con eco digital.

---

#### IV. LA TESIS

Lo que todas estas variantes comparten no es un error lógico nuevo. Es un mecanismo de evasión epistémica: la IA se usa no para conocer mejor, sino para no tener que sostener lo que se dice. Funciona porque produce texto que suena riguroso sin serlo, lo que permite apropiarse del tono de la evidencia sin asumir la carga de la prueba.

Desde Aristóteles hasta LinkedIn, la estructura es la misma: ganar la discusión sin tener razón. Lo que cambió es la velocidad, la escala y la sofisticación del disfraz.

La IA no inventó la falacia. Le dio un traje de datos.

---

*Entro donde el mapa dice 'dragones'.*