

Agencia moral sin imputación: los modelos de lenguaje ante los nuevos materialismos

Carlos Isaac Zainea

Doctorado en Filosofía, Universidad Santo Tomás, Bogotá

Seminario electivo: Aportes Filosóficos de los Nuevos Materialismos — Prof. Juan Carlos Moreno — 2026-1

Junio de 2026

Resumen. Los modelos de lenguaje de gran escala (LLM) producen actos con consecuencias morales sin satisfacer las condiciones clásicas de la agencia moral. Este artículo examina el problema con las herramientas de los nuevos materialismos — la delegación moral de Latour, la mediación tecnológica de Verbeek, la agencia distributiva de Bennett y el realismo agencial de Barad — en diálogo con la distinción entre *agencia operativa* y *agencia imputable* elaborada en una investigación en curso a partir de Floridi y Ricoeur. Se sostiene una tesis en dos movimientos: (1) los nuevos materialismos describen la agencia moral del LLM con mayor precisión que el instrumentalismo y el antropomorfismo — el sistema es masa faltante sin delegante, mediador generativo y miembro de ensamblajes que actúan —; (2) de la distribución de la agencia no se sigue la distribución de la imputación: la agencia se distribuye, la imputación se dirige, y dirigirla es una operación de diseño. La objeción posthumanista de Barad — las fronteras agenciales no están dadas — no derrota esta distinción sino que revela su estatuto: un corte agencial que debe practicarse deliberadamente y del cual se responde. El andamiaje hermenéutico propuesto en esa investigación se relea, así, como el aparato que practica ese corte.

Palabras clave: agencia moral; nuevos materialismos; modelos de lenguaje; imputabilidad; mediación tecnológica; corte agencial.

Abstract. Large language models (LLMs) produce acts with moral consequences without satisfying the classical conditions of moral agency. This paper examines the problem through the tools of new materialisms — Latour’s moral delegation, Verbeek’s technological mediation, Bennett’s distributive agency, and Barad’s agential realism — in dialogue with the distinction between *operative agency* and *imputable agency* developed in ongoing research drawing on Floridi and Ricoeur. A two-step thesis is defended: (1) new materialisms describe the moral agency of LLMs more accurately than instrumentalism and anthropomorphism; (2) the distribution of agency does not entail the distribution of imputation: agency is distributed, imputation is directed, and directing it is a matter of design. Barad’s posthumanist objection does not defeat this distinction but reveals its status: an agential cut that must be deliberately enacted and answered for. The hermeneutic scaffolding proposed in that research is thus reread as the apparatus that enacts this cut.

Keywords: moral agency; new materialisms; large language models; imputability; technological mediation; agential cut.

I. Máquinas que actúan moralmente sin responder por ello

Los modelos de lenguaje de gran escala (LLM) producen actos con consecuencias morales. Aconsejan a pacientes que preguntan por síntomas, redactan dictámenes que circulan en despachos jurídicos, recomiendan decisiones a quienes no distinguen el consejo competente del estadísticamente plausible. La pregunta por la **agencia moral** de estos sistemas suele plantearse en términos heredados de la filosofía de la acción: ¿tiene el LLM intenciones, estados mentales, capacidad de actuar por razones? Y la respuesta negativa parece zanjar el asunto: sin intencionalidad no hay agencia moral; sin agencia moral, el sistema es un mero instrumento y la moralidad pertenece íntegramente a sus usuarios y diseñadores.

Este trabajo sostiene que ese planteamiento — y su tranquilizadora conclusión instrumentalista — es exactamente lo que los nuevos materialismos han hecho insostenible. Si algo une a autores tan distintos como Latour, Verbeek, Bennett y Barad es la demostración, por vías independientes, de que la agencia no es monopolio del sujeto intencional: está delegada en artefactos (Latour), emerge en la mediación técnica (Verbeek), se distribuye en ensamblajes heterogéneos (Bennett) y se enactúa en intra-acciones que producen a los propios agentes (Barad). Leído desde este campo, el LLM no es un instrumento pasivo cuya moralidad reside en otra parte: es un participante material en configuraciones que producen efectos morales reales.

Pero el argumento de este trabajo no se detiene en esa constatación. Su tesis tiene dos movimientos. **Primero:** los nuevos materialismos describen correctamente la agencia moral del LLM como agencia distribuida, mediadora y material — y esa descripción es más precisa que la alternativa instrumentalista y que la antropomorfista. **Segundo:** de la distribución de la agencia no se sigue la distribución de la *imputación*, esto es, de la capacidad de ser tenido por autor responsable que responde ante otros por lo hecho. Un ensamblaje puede actuar; no puede comparecer. La confusión entre ambas cosas — agencia e imputación — es el punto ciego que el presente trabajo localiza en el campo neomaterialista, con ayuda de una distinción elaborada en una investigación en curso: la distinción entre **agencia operativa** y **agencia imputable**.

Esa investigación analiza los LLM como agentes en el sentido operacional de Floridi y Sanders (2004) — sistemas interactivos, autónomos y adaptables, sin estados mentales — cuyos fallos no son «alucinaciones» sino *infortunios de autoridad epistémica*: actos que adoptan la forma de aserciones fundamentadas sin que ningún agente pueda responder por su fundamentación. El presente ejercicio extiende ese diagnóstico del registro epistémico al registro moral, y lo pone a prueba ante sus interlocutores más exigentes: los cuatro autores del seminario que más radicalmente han redistribuido la agencia más allá de lo humano.

El recorrido es el siguiente. La sección 2 presenta el puente conceptual: la «moralidad sin mente» de Floridi y Sanders. Las secciones 3 a 6 examinan, en un panel equilibrado, qué aporta cada autor del seminario al problema: la delegación moral en Latour (§3), la mediación moral en Verbeek (§4), la agencia distributiva y su problema de responsabilidad en Bennett (§5), y la intra-acción con su objeción posthumanista a la imputación en Barad (§6). La sección 7 articula la respuesta: la distinción agencia/imputación como corrección interna al campo neomaterialista, y el corte agencial — reapropiado de Barad — como principio de diseño.

2. La moralidad sin mente

Floridi y Sanders (2004) ofrecen el punto de partida técnico. Su propuesta, anterior al despliegue de los LLM pero formulada para sistemas artificiales autónomos, consiste en desacoplar la agencia moral de la intencionalidad. Un sistema es **agente** cuando, en el nivel de abstracción apropiado, satisface tres condiciones: interactividad con su entorno, autonomía respecto del estímulo inmediato y adaptabilidad de sus propias reglas de transición. Un agente es **agente moral** cuando sus acciones pueden causar bien o daño moralmente calificable. Nada en esta definición requiere mente, conciencia ni libre albedrío: es lo que los autores llaman *mindless morality*, moralidad sin mente.

La pieza decisiva de la propuesta es una segunda distinción: la que separa *accountability* de *responsibility*. Un agente artificial puede ser fuente identificable de acción moral — y en ese sentido «rendir cuentas»: puede ser corregido, reprogramado, desconectado — sin ser moralmente *responsable*, porque la responsabilidad exige lo que el sistema no tiene: la capacidad de comprender el reproche, de dar razones, de responder. Los LLM contemporáneos satisfacen con holgura las tres condiciones de la agencia y producen efectos moralmente calificables a escala masiva. Son, en el vocabulario de Floridi y Sanders, agentes morales sin responsabilidad moral.

La distinción es filosóficamente fecunda pero queda, en el artículo de 2004, formulada desde la filosofía de la información, sin una teoría del tipo de *práctica* en la que la responsabilidad se ejerce. La investigación que sirve de trasfondo a este ejercicio tradujo esa carencia al vocabulario de Ricoeur (1996): lo que el agente artificial no puede hacer es *imputarse* sus actos — ser tenido por autor que responde ante otros, con su identidad narrativa comprometida en la respuesta. La fórmula resultante, **agencia sin imputación**, nombra la asimetría: hay agencia operativa real, con efectos morales reales, y no hay nadie en el sistema a quien imputar lo actuado.

Lo que los nuevos materialismos permiten — y es la apuesta de este trabajo — es poner a prueba esa fórmula en un campo teórico que ha pensado la agencia no-humana con mucha mayor radicalidad que la filosofía de la información. Si la fórmula sobrevive a Latour, Verbeek, Bennett y Barad, sale fortalecida; si no sobrevive, habrá que reformularla. Adelantamos el resultado: sobrevive, pero transformada — deja de ser una tesis sobre lo que el LLM *no tiene* y se convierte en una tesis sobre lo que los ensamblajes *no pueden hacer* y los diseños *deben hacer*.

3. Latour y la moralidad delegada en las cosas

El ensayo de Latour (1992) sobre las «masas faltantes» comienza con una escena doméstica de teoría moral: un automóvil que se niega a arrancar hasta que el conductor abroche el cinturón de seguridad. ¿Dónde está la moralidad en esa escena? «¿En mí, un conductor humano, dominado por el poder sin mente de un artefacto?» (Latour, 1992, p. 152). La pregunta es retórica: la moralidad está en ambos lados y en el vínculo. Los sociólogos que buscan las «masas faltantes» que explicarían el orden social — como los cosmólogos buscan la masa faltante del universo — las encuentran, dice Latour, en los no-humanos: los artefactos son las masas ocultas y despreciadas que componen nuestra moralidad.

El mecanismo de esa composición es la **delegación**: la transformación de un esfuerzo mayor en uno menor

mediante su desplazamiento a un no-humano. El cierrapuertas hidráulico reemplaza al portero indisciplinado; el resalto reemplaza al policía de tránsito; el cinturón inteligente reemplaza a la conciencia vial. Y lo delegado no es solo fuerza: «hemos podido delegar a los no-humanos no solo la fuerza tal como la hemos conocido por siglos, sino también valores, deberes y ética» (Latour, 1992, p. 157). El artefacto que recibe la delegación *prescribe* comportamiento de vuelta: la **prescripción** es, en la fórmula célebre del ensayo, «la dimensión moral y ética de los mecanismos» (p. 157). Por eso «ningún humano es tan implacablemente moral como una máquina» (p. 157): la máquina no se cansa, no negocia, no hace excepciones.

Leído desde Latour, el LLM es la última y más extraña de las masas faltantes. Como el cierrapuertas, recibió una delegación masiva: le delegamos la producción de discurso experto — síntesis, consejo, diagnóstico, redacción — que antes requería humanos entrenados. Como el cinturón, prescribe de vuelta: sus respuestas configuran lo que el usuario considera razonable, normal, recomendable. Pero hay una diferencia que el propio marco latouriano permite formular con precisión. El cierrapuertas prescribe un comportamiento *determinado*: su moralidad fue inscrita por un ingeniero que sabía qué quería prescribir. El LLM, en cambio, prescribe discurso *abierto*: nadie inscribió en él las respuestas que dará, porque sus respuestas emergen estadísticamente de un corpus que ningún diseñador domina. La delegación latouriana clásica tiene un delegante identificable que respondería por lo inscrito — el ingeniero del cinturón, el urbanista del resalto. La delegación al LLM difumina al delegante: ¿quién inscribió el consejo que el sistema dará mañana a un paciente que aún no ha preguntado?

Esta diferencia importa porque la sociología latouriana de la moralidad delegada presupone, sin tematizarlo, que la cadena de delegación puede recorrerse hacia atrás: del artefacto al inscriptor, del *script* al autor del script. En el LLM esa remontada se interrumpe — no por secreto industrial sino por estructura: el «script» es una distribución de probabilidad entrenada, no una inscripción intencionada. La moralidad delegada se vuelve, por primera vez, moralidad delegada *sin delegante imputable*. Latour describe correctamente que el LLM participa en nuestra moralidad; lo que su marco no estaba preparado para describir es una masa faltante cuyo autor también falta.

4. Verbeek y la mediación moral del discurso

Verbeek (2005) toma de Latour el vocabulario de la delegación y los scripts, pero lo corrige en un punto que resulta decisivo para nuestro problema. En los ejemplos latourianos, lo que el artefacto hace tiende a reducirse a lo que los humanos le delegaron: el cierrapuertas hace lo que el ingeniero inscribió. Para Verbeek esa reducción reintroduce, por la puerta trasera, el primado del diseñador humano. Su alternativa — elaborada desde la postfenomenología de Ihde — es la teoría de la **mediación tecnológica**: los artefactos no ejecutan funciones inscritas sino que median la relación entre los humanos y su mundo, y esa mediación *emerge en la relación de uso*, no está contenida en el diseño. Un mismo artefacto es multiestable: media de manera distinta en contextos distintos, y produce efectos que ningún diseñador anticipó.

La mediación tiene dos dimensiones que Verbeek desarrolla por separado: una dimensión **hermenéutica** — los artefactos transforman la percepción y la experiencia, amplificando unos aspectos del mundo y reduciendo otros — y una dimensión **existencial** — los artefactos transforman la acción, invitando a unas conductas e inhibiendo otras. En ambas dimensiones la mediación es moralmente densa: el automóvil que fuerza el cinturón

«está lleno de moralidad» (Verbeek, 2005, p. 212), y el cribado prenatal — ejemplo del propio Verbeek — no informa neutralmente sino que «puede invitar implícitamente al aborto» (p. 216): transforma al feto en objeto de decisión y a los padres en decisores, reconfigura el espacio moral en el que la decisión ocurre. De ahí la tesis central: «el papel mediador de las cosas puede ser juzgado él mismo en términos morales, haya sido delegado explícitamente o no» (p. 216). Hacer ética de la tecnología es hacer, dice Verbeek, «ética por otros medios».

El LLM es un caso de mediación moral de una intensidad que los ejemplos clásicos de Verbeek apenas anticipan, porque media en las dos dimensiones a la vez y sobre el mismo material: el discurso. En la dimensión hermenéutica, el LLM transforma la percepción del problema sobre el que se le consulta: organiza qué aspectos del dilema aparecen como relevantes, qué opciones se vuelven visibles, qué consideraciones quedan amplificadas o reducidas en la respuesta. En la dimensión existencial, invita a cursos de acción: el consejo generado no describe el espacio de decisión, lo *configura*. Quien consulta a un LLM sobre un dilema no recibe información sobre su situación moral: recibe su situación moral ya mediada, ya estructurada, ya inclinada. Y a diferencia del cribado prenatal — cuya mediación es estable y puede por tanto ser estudiada, anticipada y regulada —, la mediación del LLM es generativa: produce una mediación nueva en cada consulta, no auditada por nadie antes de surtir efecto.

Verbeek extrae de su análisis una consecuencia normativa: si las cosas median moralmente, el diseño es una actividad moral, y la moralidad de los artefactos debe ser objeto de deliberación explícita — la propuesta de «moralizar la tecnología» que toma de Achterhuis y desarrolla con cautela. Esa consecuencia apunta exactamente en la dirección de este trabajo: la respuesta al LLM no es negarle densidad moral (instrumentalismo) ni atribuirle responsabilidad (antropomorfismo), sino *diseñar la mediación*. Pero la propuesta de Verbeek deja una pregunta abierta que su marco no responde: cuando la mediación es generativa y nadie puede anticipar su contenido, ¿quién responde por los efectos de una mediación que nadie diseñó? Moralizar el resalto es posible porque el resalto hace una sola cosa. ¿Cómo se moraliza un mediador que hace una cosa nueva cada vez?

5. Bennett y la agencia del ensamblaje

Bennett (2022) radicaliza el movimiento un paso más. Donde Latour y Verbeek piensan artefactos discretos — el cinturón, el resalto, el ecógrafo —, Bennett piensa **ensamblajes**: federaciones ad hoc de cuerpos heterogéneos, humanos y no-humanos, cuya potencia de actuar es emergente y no se reduce a la suma de sus miembros. Su materialismo vital atribuye a las cosas un **poder-cosa** (*thing-power*): una eficacia propia que excede los significados y propósitos humanos a los que las cosas sirven. Y de ahí su tesis sobre la agencia: «la eficacia o la efectividad a la que ese término tradicionalmente ha referido ahora se distribuye a lo largo de un campo ontológicamente heterogéneo» (Bennett, 2022, pp. 72–73). No hay sujeto como causa principal de los efectos: hay un enjambre de vitalidades en juego.

El caso central del segundo capítulo es el apagón eléctrico norteamericano de 2003, que dejó sin energía a cincuenta millones de personas. ¿Quién actuó en el apagón? La red eléctrica entera, responde Bennett: corrientes que se invirtieron, plantas que se desconectaron para protegerse, operadores que leyeron mal, mercados desregulados que habían configurado flujos insostenibles, la empresa FirstEnergy como «primer actante» pero no como causa suficiente. El apagón fue el acto de un ensamblaje agencial, y por eso mismo, reconoce Bennett, «esta federación de actantes es una criatura a la cual el concepto de responsabilidad moral se ajusta solo vagamente y a la cual no

se podrá culpabilizar del todo» (pp. 80–81).

Para el problema de este trabajo, el apagón de Bennett es una alegoría casi perfecta de lo que esa investigación de trasfondo llama la *cadena agéntica*: sistemas de múltiples agentes LLM coordinados, donde cada eslabón hereda los fallos del anterior sin que exista mecanismo interno de detección, y donde el efecto final — el consejo erróneo, el diagnóstico infundado, el dictamen sesgado — es producto del ensamblaje completo: modelo, corpus de entrenamiento, prompt, plataforma, usuario, contexto institucional. Bennett da el vocabulario exacto para describir esa configuración sin las ficciones del instrumentalismo: el ensamblaje LLM-corpus-plataforma-usuario *actúa*, y su acto no se deja reducir a la intención de ningún participante.

Pero Bennett también hereda, agravado, el problema de la responsabilidad — y hay que reconocerle que no lo esquiva. Su respuesta tiene dos partes. Primero, rechaza la «variante más fuerte o punitiva de responsabilidad moral», que le parece descansar sobre una ficción de autonomía empíricamente falsa. Segundo, redistribuye: la responsabilidad política recae sobre los ensamblajes, y la responsabilidad ética individual consiste en «la respuesta que uno da a los ensamblajes en los que se encuentra participando» — ¿me alejo de los ensamblajes dañinos?, ¿me acerco a los más nobles? Distribuir la agencia, concede, «complica el juego de atribuir culpas, pero no por ello elimina la tarea de identificar los efectos nocivos» (p. 92).

La respuesta es honesta pero insuficiente, y su insuficiencia es instructiva. «Responder al ensamblaje» — alejarse, acercarse — es una ética de la participación, no una estructura de imputación: dice qué debe hacer cada quien con su pertenencia a los ensamblajes, pero no dice ante quién comparece el ensamblaje cuando daña, ni quién puede ser interpelado para dar razones del daño. En el apagón, los afectados pudieron al menos litigar contra FirstEnergy: la estructura jurídica de imputación, por tosca que sea, ofreció una dirección. En el ensamblaje LLM la cuestión es más grave, porque su producto no es un evento físico ocasional sino un flujo continuo de actos discursivos con forma de aserción experta. Un ensamblaje que produce discurso con autoridad y al que «no se podrá culpabilizar del todo» es, en el vocabulario de la investigación que anima este ejercicio, una fábrica de infortunios: actos que reclaman el régimen de las razones sin que nadie pueda comparecer en él.

6. Barad y el corte agencial

Barad (2003) plantea el desafío más radical, porque no redistribuye la agencia entre entidades: disuelve las entidades. Su realismo agencial sostiene que los *relata* no preexisten a las relaciones; emergen dentro de los fenómenos mediante **intra-acciones** — el neologismo marca el contraste con la interacción, «que presume la existencia previa de entidades independientes» (Barad, 2003, p. 815). En cada intra-acción se efectúa un **corte agencial** que separa localmente sujeto y objeto, agente y paciente: la frontera no está dada — como en el corte cartesiano — sino que se *practica*, y podría practicarse de otro modo. La agencia, en consecuencia, «no es un atributo en absoluto: es “hacer”/“ser” en su intra-actividad» (p. 827); no es algo que alguien o algo *tiene*, sino un enactment.

Desde este marco, la pregunta que organiza este trabajo — ¿es el LLM un agente moral?, ¿quién es el agente imputable? — está mal planteada, porque presupone que hay entidades (el LLM, el usuario, el investigador) cuyas propiedades agenciales podrían inventariarse. Para Barad, «el LLM» y «el usuario» no preexisten a la práctica que los separa: emergen como tales dentro del fenómeno de la consulta, mediante un corte agencial específico. Y

la distinción misma entre agente operativo y agente imputable — el corazón conceptual de la investigación aquí extendida — sería sospechosa de humanismo residual: un corte cartesiano disfrazado, que reserva para el lado humano del fenómeno una propiedad metafísica (la imputabilidad) sustraída al análisis.

La objeción es seria y conviene no rebajarla. Pero el propio Barad provee la herramienta para responderla, porque su realismo agencial no elimina la responsabilidad: la reubica. El enactment de fronteras, escribe, conlleva siempre exclusiones constitutivas, y por tanto abre «cuestiones ineludibles de *accountability*» (p. 803); la objetividad misma se redefine como «rendir cuentas de las marcas sobre los cuerpos» (p. 824); y las posibilidades cambiantes de la materia conllevan «una responsabilidad de intervenir en el devenir del mundo, de disputar y retrabajar qué importa y qué queda excluido de importar» (p. 827). Es decir: para Barad somos responsables *de los cortes que practicamos*. Los cortes no están dados, pero no por ello son arbitrarios ni gratuitos: cada corte habilita unas posibilidades, excluye otras, y deja marcas sobre cuerpos — de las cuales hay que responder.

Aquí está la torsión que este trabajo propone. Si el corte entre agente operativo y agente imputable no está dado — concedámoslo a Barad —, entonces es un corte que se practica. Y si se practica, puede practicarse *bien o mal*, con consecuencias distintas, de las cuales se responde. El corte que el instrumentalismo practica (el LLM es herramienta neutra; toda la moralidad está en el usuario) deja sin dirección los efectos de la mediación generativa que Verbeek nos enseñó a ver. El corte que el antropomorfismo practica (el LLM es un sujeto responsable) imputa a quien no puede comparecer, y al hacerlo *descarga* a quienes sí pueden. La distinción agencia operativa / agencia imputable no es, entonces, un corte cartesiano contrabandeado: es una **propuesta de corte agencial**, argumentada y revisable, que se deja juzgar por lo que Barad exige — por las marcas que deja, por lo que habilita y lo que excluye. Habilita auditar el ensamblaje sin ficciones de neutralidad; excluye que la responsabilidad se evapore en la red. Ese es su título de validez, y es un título baradiano.

7. Distribución de agencia, dirección de imputación

Puede ahora componerse el cuadro completo. Los cuatro autores del seminario aportan, cada uno, una pieza de la descripción correcta del LLM como participante moral — y cada uno deja abierta una pregunta que la pieza siguiente agudiza:

Autor	Lo que aporta al caso LLM	La pregunta que deja abierta
Latour	La moralidad ya estaba delegada en no-humanos; el LLM es una masa faltante más	¿Qué pasa cuando la delegación no tiene delegante imputable — cuando el script no tiene autor?
Verbeek	El LLM no ejecuta moral inscrita: media la percepción y la acción morales, y su mediación es generativa	¿Cómo se «moraliza» un mediador que produce una mediación nueva, no auditada, en cada uso?
Bennett	El ensamblaje LLM-corpus-plataforma-usuario actúa; su acto no se reduce a ninguna intención	Si al ensamblaje «no se lo podrá culpabilizar del todo», ¿ante quién comparece cuando daña?
Barad	Las fronteras agente/instrumento no están dadas: se practican en cortes agenciales de los que se responde	¿Qué corte debe practicarse — y quién responde de él?

La secuencia de preguntas abiertas converge en un solo punto: los nuevos materialismos distribuyen la agencia con una precisión descriptiva inalcanzable para el instrumentalismo, pero la **imputación** — la capacidad de comparecer, de dar razones, de ser tenido por autor ante otros — no se distribuye con la agencia. No porque alguna metafísica humanista la reserve para las personas, sino por la estructura misma de la práctica de responder: responder es un acto de segundo orden que toma los actos de primer orden como propios, y un ensamblaje no tiene el «propio» que ese acto requiere. Bennett lo reconoce sin resolverlo; Floridi y Sanders lo formalizan sin teorizarlo; Ricoeur (1996) le da el nombre que este trabajo adopta: imputabilidad, la capacidad de un sí-mismo de reconocerse autor y mantenerse responsable en el tiempo. La agencia se distribuye; la imputación se *dirige*. Y dirigirla es una operación de diseño.

Esa operación tiene nombre en la investigación de trasfondo: el **andamiaje hermenéutico** — un protocolo procesal que organiza el trabajo del LLM en fases auditables con criterios externos de verificación, de modo que la agencia operativa del sistema quede acoplada, paso a paso, al circuito de imputación de un investigador humano que valida, disputa o revierte cada resultado y comparece por la totalidad. Releído con el vocabulario de este seminario, el andamiaje es tres cosas a la vez. En términos de Verbeek, es *mediación de la mediación*: una estructura que vuelve auditable la mediación generativa del LLM antes de que surta efecto. En términos de Bennett, es un *ensamblaje con dirección*: no niega la agencia distribuida del conjunto, pero le instala una asimetría interna — un punto del ensamblaje donde alguien puede comparecer. En términos de Barad, es un *corte agencial practicado deliberadamente*: la frontera entre lo que el sistema configura y lo que el investigador imputa no se descubre, se enactúa — y el andamiaje es el aparato que la enactúa de manera visible, revisable y responsable.

El resultado invierte la carga de la objeción posthumanista. No es la distinción agencia/imputación la que arrastra un residuo humanista: es su *ausencia* la que produce el residuo más viejo de todos — efectos sin nadie que responda, el poder en su forma más clásica. Los nuevos materialismos enseñaron que las cosas actúan; la tarea que dejan pendiente, y que este trabajo ha intentado precisar para el caso del LLM, es asegurar que cuando las cosas actúan con autoridad discursiva, la pregunta «¿quién responde?» conserve una dirección. Distribuir la agencia fue el logro descriptivo del nuevo materialismo. Dirigir la imputación es la tarea normativa que ese logro vuelve urgente.

8. Conclusión

Este trabajo aplicó los aportes del seminario al problema de la agencia moral en los modelos de lenguaje, en diálogo con una investigación en curso. Tres resultados pueden retenerse. **Primero**, un resultado descriptivo: leído con Latour, Verbeek y Bennett, el LLM no es instrumento neutro sino participante material en configuraciones morales — masa faltante sin delegante, mediador generativo, miembro de ensamblajes que actúan. Esta descripción es superior a las alternativas instrumentalista y antropomorfista, y converge con la caracterización del LLM como agente operativo (Floridi & Sanders, 2004) que esa investigación elabora en el registro epistémico. **Segundo**, un resultado crítico: el campo neomaterialista, al distribuir la agencia, hereda un problema de imputación que sus propios autores registran — Bennett expresamente — sin resolver. La distinción entre agencia operativa y agencia imputable, elaborada en esa investigación desde Floridi y Ricoeur, nombra ese punto ciego con precisión: la agencia se distribuye; la imputación no se distribuye, se dirige. **Tercero**, un resultado constructivo: la objeción posthumanista de Barad — las fronteras agenciales no están dadas — no derrota la distinción sino que revela su verdadero estatuto: no es un dato metafísico sino un corte agencial que debe practicarse deliberadamente, y del cual se responde. El andamiaje hermenéutico de esa investigación es el aparato que practica ese corte.

Para esa investigación, el rodeo por los nuevos materialismos no fue ornamental. Obligó a reformular la tesis central en términos más fuertes: la fórmula «agencia sin imputación» ya no describe una carencia del LLM (lo que el sistema *no tiene*) sino una propiedad estructural de los ensamblajes sociotécnicos (lo que ninguna red *puede hacer* por sí misma) y, por tanto, una obligación de diseño (lo que toda arquitectura epistémica *debe* instalar). Esa reformulación — de la ontología del déficit a la normatividad del corte — es la ganancia que este seminario deja para la investigación.

Referencias

- Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Sigms: Journal of Women in Culture and Society*, 28(3), 801–831. <https://doi.org/10.1086/345321>
- Bennett, J. (2022). *Materia vibrante: Una ecología política de las cosas* (M. Gonnet, trad.). Caja Negra. (Obra original publicada en 2010).
- Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, 14(3), 349–379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- Latour, B. (1992). Where are the missing masses? The sociology of a few mundane artifacts. En W. E. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 225–258). MIT Press.
- Latour, B. (2008). *Reensamblar lo social: Una introducción a la teoría del actor-red* (G. Zadunaisky, trad.). Manantial. (Obra original publicada en 2005).
- Ricoeur, P. (1996). *Sí mismo como otro* (A. Neira, trad.). Siglo XXI. (Obra original publicada en 1990).

Verbeek, P.-P. (2005). *What things do: Philosophical reflections on technology, agency, and design* (R. P. Crease, trad.). Pennsylvania State University Press.