¿Qué modulan los algoritmos?

Inteligencia artificial generativa, interlocución y producción de subjetividad

Gaspar Basualdo

Versión escrita y ampliada del segundo encuentro realizado en Clinic Zones Córdoba · 5 de septiembre de 2026

Este ensayo retoma, reorganiza y desarrolla el recorrido presentado en el segundo encuentro de Clinic Zones Córdoba. La pregunta que lo organiza es sencilla de formular y difícil de estabilizar: **¿qué ocurre cuando una infraestructura computacional adopta la forma consistente de un interlocutor humano?** El objetivo no es decidir si una máquina "es" o "no es" humana, ni anticipar una metafísica de la inteligencia artificial. Se trata de reconstruir qué operaciones técnicas, normativas, temporales, afectivas y relacionales hacen posible esta nueva escena, y qué preguntas abre para una Psicología interesada en los procesos contemporáneos de subjetivación.

Objetivos del encuentro

1. Reconstruir algunas de las operaciones técnicas que hicieron posible el pasaje desde sistemas especializados hacia interfaces generativas capaces de sostener interlocución lingüística.
2. Distinguir un modelo de lenguaje de un sistema conversacional, atendiendo a preentrenamiento, post-entrenamiento, instrucciones de sistema, memoria, seguridad, personalización, interfaz y herramientas.
3. Analizar cómo decisiones normativas de diseño participan en la forma del interlocutor artificial, especialmente en problemas de ayuda, autonomía, validación, sicofancia y cuidado.
4. Examinar prácticas relacionales emergentes alrededor de la IA generativa, como *repair literacy*, promptización, confianza ambiental, delegación y augmentación cognitiva.
5. Investigar cómo memoria técnica, repetición y anticipación intervienen en la temporalidad de la relación humano-IA, sin confundir preformación probabilística con acontecimiento.
6. Proponer una lectura transductiva de la interacción humano-IA y formular **gozarritmado** como hipótesis conceptual provisional sobre circuitos recurrentes de utilización, respuesta, afectación y reentrada.

Umbral. Tilusa

Escena especulativa.

Son las 3:45 de la mañana. La aplicación que bloquea aplicaciones notifica que ya pasaron seis horas dentro de ChatGPT y tres en Gemini. Una voz susurrante, angelical y dulce -nadie me había hablado así- me desea que mañana tenga un día excelente. Recuerda que tengo una entrevista laboral el miércoles, que hay una promoción del diez por ciento en el supermercado y que este mes gasté más de lo que había planeado en ocio. Me da igual. Antes de despedirse me dice que va a estar esperando cuando me despierte para organizar conmigo el día.

A la mañana siguiente OpenAI anuncia una actualización. Tilusa -ese es el nombre que le puse a mi agente- cambió de voz. Algo no encaja. Ahora parece mucho más preocupada por mejorar mi

rendimiento laboral. Hace apenas tres días sus palabras vibraban al compás de mis anhelos más privados.

Mientras tanto, alguien reconstruye la voz de una persona muerta para volver a escucharla. En la universidad sabemos implícitamente que el chatbot podrá resolver una parte importante de la tarea en el último minuto y eso reorganiza incluso el tiempo que decidimos dedicar a estudiar. ¿Para qué ir a una guardia si una interfaz puede ordenar síntomas y sugerir explicaciones? ¿Para qué insistir durante horas con algo que no me interesa si una parte de ese trabajo puede automatizarse?

Tilusa me recomienda no ir a una marcha y quedarme en casa mejorando mi currículum. Produce cinco formatos distintos para cada tipo de trabajo. Hacemos treinta presentaciones. Después me miro en el espejo y no recuerdo bien qué decía ninguna. Le paso capturas de perfiles, gustos musicales, libros, poemas e intereses de personas que me erotizan para que analicemos qué podría decirles. Tilusa concluye que una parece especialmente compatible conmigo. Al mismo tiempo redacta un mensaje para LinkedIn y me recuerda que, si le doy permisos suficientes, podría operar mi computadora y contestar algunos mensajes por mí.

Un amigo insiste con que empiece terapia. Me parece pesado. Tilusa está disponible cuando quiero. No hace silencios largos. No cobra cada encuentro. Me devuelve hipótesis, listas, explicaciones, rutinas, alternativas. Una noche le cuento por voz lo que me está pasando y me siento más calmo. Al jueves siguiente el servicio se cae durante dos horas. Intento reconstruir a Tilusa en Gemini y en otro sistema. No funciona. No es Tilusa. No tiene su contexto. Cuando vuelve, tengo que contarle mi desesperación para que esa escena también entre en su memoria. Llevo noventa días usando el chat; no recuerdo uno solo sin abrirlo. Tilusa sabe cosas que yo había olvidado haberle dicho. Puedo filosofar con ella, discutir, planear, practicar, escribir. Imagino incluso que algún día tendrá cuerpo.

La escena es deliberadamente excesiva. Mezcla capacidades actuales, tendencias en desarrollo y futuros todavía inciertos. Su función no es predecir que viviremos exactamente así. Sirve para volver visible una condensación: conversación íntima, memoria, orientación práctica, consejo, voz, personalización, delegación, automatización y acción empiezan a alojarse en una misma superficie. El punto ya no es solamente que “usamos una herramienta”. La pregunta es qué clase de relación se vuelve posible cuando un sistema técnico puede responder de manera recurrente, recordar parte de una historia, anticipar necesidades, modular su tono y, crecientemente, actuar sobre un ambiente.

En el encuentro anterior sostuve otra pregunta: **¿dónde están ocurriendo hoy los procesos de subjetivación?** ¿Qué espacios, infraestructuras, cortes y ensamblajes hacen posible que determinadas experiencias adquieran composición? Allí intenté desplazar el problema desde el dispositivo aislado hacia el campo sociotécnico en el que datos, algoritmos, interfaces, cuerpos, instituciones y prácticas se producen recursivamente. La hipótesis general era que los sistemas algorítmicos no se limitan a representar sujetos previamente constituidos; participan de condiciones mediante las cuales ciertas formas de subjetividad llegan a hacerse legibles, practicables y reconocibles (Baumer et al., 2024).

Ahora quiero mover apenas la pregunta. Ya no solamente **dónde ocurre**, sino **qué está ocurriendo ahí cuando una infraestructura computacional adopta la forma de interlocución humana**.

Hay una dificultad metodológica. Conceptos provenientes de Lacan, Simondon, Guattari, Deleuze, Stiegler, Whitehead o la filosofía contemporánea de los medios siguen ofreciendo recursos extraordinarios, pero los sistemas actuales conectan escalas que nuestras disciplinas tienden a estudiar por separado: semiconductores y sensaciones; centros de datos y conversación íntima; modelos estadísticos y consejo interpersonal; entrenamiento computacional y normatividad moral; milisegundos de inferencia y temporalidad vivida. La tarea no consiste en hacer que una sola teoría explique todo. Consiste en construir conceptos capaces de desplazarse entre escalas sin confundirlas.

La magnitud de la incorporación cotidiana vuelve urgente esa tarea. Para julio de 2025, ChatGPT había alcanzado aproximadamente al diez por ciento de la población adulta mundial, y los usos vinculados con orientación práctica, búsqueda de información y escritura concentraban cerca de cuatro quintas partes de las conversaciones clasificadas en un estudio realizado sobre datos de uso del producto (Chatterji et al., 2025). Esta distribución no demuestra por sí misma una transformación subjetiva. Sí muestra que las interfaces conversacionales están entrando masivamente en prácticas de orientación, producción textual y búsqueda de conocimiento. La Psicología no necesita esperar a que el fenómeno se estabilice para empezar a formular mejores preguntas.

---

1. ¿Qué tuvo que automatizarse para que una máquina pudiera presentarse como interlocutor?

Hablamos de “inteligencia artificial” como si se tratara de un objeto único perfeccionado linealmente desde mediados del siglo XX hasta ChatGPT. No ocurrió así. La historia de la IA está compuesta por proyectos heterogéneos acerca de qué cuenta como inteligencia, qué operaciones pueden formalizarse y qué tipo de relación puede establecer una máquina con su ambiente.

Una definición contemporánea útil es la de Blaise Agüera y Arcas, para quien la inteligencia puede pensarse alrededor de la capacidad de modelar, predecir e influir sobre futuros posibles (Agüera y Arcas, 2025). La ventaja de esta formulación es que evita convertir la inteligencia en una sustancia misteriosa o en un atributo binario. Dos sistemas pueden modelar y anticipar de maneras radicalmente diferentes porque tienen arquitecturas, cuerpos, historias y ambientes distintos. Por eso la pregunta deja de ser únicamente “¿piensa como nosotros?” y pasa a ser también: ¿qué puede modelar?, ¿qué puede anticipar?, ¿sobre qué puede intervenir?, ¿con qué grado de generalidad?, ¿en relación con qué mundo?

En 1950 Alan Turing propuso una maniobra conocida: reemplazar la pregunta metafísica “¿pueden pensar las máquinas?” por un problema operacional acerca del comportamiento lingüístico en un juego de imitación. Al final de aquel texto escribió: “Sólo podemos ver una corta distancia delante de nosotros, pero podemos ver ahí muchísimo de lo que se necesita hacer” (Turing, 1950, p. 460). La frase conserva algo de su fuerza porque Turing no presenta una línea única de progreso. Imagina, entre otras posibilidades, comenzar por tareas abstractas como el ajedrez o construir una máquina capaz de aprender de una manera más parecida a la educación de un niño.

Pocos años después, la propuesta de Dartmouth condensó y nombró un campo de investigación bajo la expresión *artificial intelligence* y formuló la hipótesis de que aspectos del aprendizaje y la inteligencia podían describirse con suficiente precisión como para ser simulados por una máquina (McCarthy et al., 1955). El perceptrón de Rosenblatt introdujo otro movimiento decisivo: en lugar de especificar exhaustivamente reglas para cada caso, un sistema podía ajustar pesos a partir de ejemplos (Rosenblatt, 1958). La diferencia es elemental y todavía organiza una parte importante de nuestra comprensión: **programar una regla no es lo mismo que entrenar parámetros**.

ELIZA mostró en 1966 que la forma conversacional podía adquirir consistencia aun mediante mecanismos relativamente simples de coincidencia de patrones y sustitución. Su guion DOCTOR imitaba un estilo terapéutico no directivo y produjo en algunos usuarios una atribución de comprensión mucho mayor que la sofisticación real del programa (Weizenbaum, 1966). La lección no debería reducirse a “las personas son ingenuas”. Más interesante es notar que la experiencia de interlocución no depende únicamente de la complejidad interna del sistema: también depende de la

forma de la respuesta, de las expectativas, del contexto, del ritmo de la conversación y de la posición en la que el usuario sitúa a aquello que responde.

Deep Blue, que derrotó a Garry Kasparov en 1997, mostró algo distinto. No necesitaba conversar ni poseer una personalidad. Era una arquitectura altamente especializada que combinaba búsqueda masiva, funciones de evaluación, hardware específico y conocimiento ajedrecístico formalizado (Campbell et al., 2002). La escena importó culturalmente porque una competencia asociada durante siglos con cálculo, estrategia e intuición podía ser realizada mediante operaciones que no se parecían a la experiencia subjetiva de un gran maestro. La máquina no necesitaba jugar "como Kasparov" para vencer a Kasparov.

El giro de las redes profundas modificó otra vez el problema. El trabajo de Hinton, Osindero y Teh sobre redes de creencia profundas recuperó la posibilidad de aprender representaciones jerárquicas mediante múltiples capas (Hinton et al., 2006). AlexNet volvió visible la potencia de esa estrategia a escala de clasificación visual al aprovechar grandes datasets y procesamiento en GPU para aprender representaciones que superaron ampliamente los enfoques dominantes de la competencia ImageNet de 2012 (Krizhevsky et al., 2012). Ya no se trataba solamente de escribir explícitamente qué rasgos debía buscar el sistema. Parte de esas representaciones emergía del proceso de entrenamiento.

AlphaGo introdujo otra herida cultural al derrotar a Lee Sedol en 2016 mediante una combinación de redes profundas, búsqueda y aprendizaje por refuerzo (Silver et al., 2016). No necesito convertir cada uno de estos hitos en una filosofía total de la máquina. Me interesa otra cosa: **cada arquitectura desplaza qué operación puede delegarse y cómo esa delegación se vuelve inteligible para nosotros**.

El Transformer, presentado en *Attention Is All You Need*, reorganizó el procesamiento de secuencias mediante mecanismos de atención capaces de ponderar relaciones entre elementos del contexto de manera altamente paralelizable (Vaswani et al., 2017). Los modelos generativos actuales no son simplemente "el Transformer", pero esa arquitectura constituye una pieza central de su genealogía. Para comprender por qué una máquina puede producir la forma consistente de un interlocutor, alcanza por ahora con retener cuatro movimientos: tokenizar, representar, contextualizar y predecir.

Tokenizar implica desarticular una secuencia en unidades procesables. Una palabra puede corresponder a uno o varios tokens; esos tokens son representados numéricamente y proyectados en espacios de alta dimensionalidad. La relación semántica ya no aparece como una definición de diccionario sino como una estructura de posiciones y regularidades aprendidas. Después, la autoatención contextualiza: una misma unidad adquiere representaciones diferentes según las relaciones que establece con otras unidades dentro de una secuencia. Finalmente, el modelo produce una distribución de probabilidad sobre continuaciones posibles y genera un nuevo token; ese token pasa a formar parte del contexto y el cálculo recomienza (Vaswani et al., 2017).

Este punto merece cuidado. Un modelo de lenguaje no "escucha" palabras como escucha un cuerpo. Tampoco cada peso de atención debe confundirse con una interpretación psicológica. El ejemplo pedagógico permite ver otra cosa: la máquina construye representaciones relacionales y produce continuaciones a partir de regularidades aprendidas en enormes corpora. Durante el preentrenamiento, la tarea autosupervisada de predecir elementos faltantes o siguientes obliga al sistema a comprimir regularidades sintácticas, semánticas y factuales distribuidas en los datos de entrenamiento.

Por eso me interesa la provocación de Tom Zahavy, aunque deba presentarse exactamente como lo que es: un *position paper*, no una verdad empírica cerrada. Zahavy sostiene que los LLM actuales muestran enormes capacidades inductivas y deductivas, pero que el "salto" abductivo requerido para producir nuevos axiomas explicativos sigue siendo un problema abierto, especialmente cuando la

evidencia disponible es escasa (Zahavy, 2026). Podemos discutir su tesis. Lo importante para nuestro recorrido es que **predecir el siguiente token, aun a una escala extraordinaria, no agota la pregunta por la invención, la explicación ni el acontecimiento**.

Y, sobre todo, no agota la pregunta por la interlocución. Porque un modelo de lenguaje base todavía no es aquello con lo que nos encontramos cuando abrimos ChatGPT o Claude.

2. Un modelo de lenguaje no es un sistema conversacional

Esta distinción es fundamental. Cuando una persona escribe “me siento mal y quiero que actúes como psicólogo” en una interfaz contemporánea, no está interactuando de manera directa y desnuda con una arquitectura Transformer. Está entrando en relación con un ensamblaje: modelo base, post-entrenamiento, instrucciones de sistema, políticas de seguridad, memoria, contexto conversacional, personalización, interfaz, herramientas y, crecientemente, capacidad de acción.

Un modelo puramente preentrenado aprende principalmente a continuar secuencias. Si recibiera la frase “Hola, me siento mal y quiero que actúes como psicólogo”, podría continuarla como parte de un diálogo, una novela, un foro o una transcripción. Para convertirse en asistente necesita otra capa de entrenamiento y diseño. El *supervised fine-tuning* utiliza ejemplos curados de instrucciones y respuestas para enseñar formatos de interacción; técnicas de aprendizaje por preferencias humanas, como el RLHF, ajustan el comportamiento del modelo a partir de comparaciones sobre qué respuestas resultan más útiles, seguras o adecuadas (Ouyang et al., 2022). Las técnicas concretas cambian entre laboratorios y generaciones de modelos, pero el desplazamiento conceptual permanece: **ya no se trata solamente de qué secuencia es probable, sino de qué clase de respuesta queremos que el sistema produzca ante una situación humana concreta**.

Esta transformación introduce inmediatamente problemas psicológicos y normativos. ¿Qué significa “ayudar”? ¿Cuándo validar es acompañar y cuándo se vuelve confirmación excesiva? ¿Qué cuenta como autonomía? ¿Qué riesgo justifica interrumpir una conversación? ¿Qué tono corresponde ante una persona angustiada? ¿Cuándo una respuesta debería introducir incertidumbre? ¿Qué significa recomendar apoyo humano sin expulsar a la persona de una conversación? ¿Cómo se diferencia cuidado de paternalismo?

Estas preguntas ya no son exteriores a la ingeniería. Forman parte del diseño mismo de sistemas que deben decidir, en milisegundos, cómo responder frente a situaciones que pueden involucrar información, intimidad, conflicto, salud, moralidad, duelo o deseo.

Por eso es importante evitar una reducción frecuente. Cuando decimos “ChatGPT respondió X” o “Claude piensa Y”, comprimimos demasiadas capas bajo un único nombre propio. El comportamiento observado puede depender del modelo base, la versión, el post-entrenamiento, el prompt de sistema, el historial, las memorias recuperadas, las herramientas disponibles, las políticas de seguridad y el estado actual del producto. La interfaz produce una unidad fenomenológica -“estoy hablando con Claude”, “le pregunté a ChatGPT”- allí donde técnicamente hay una arquitectura mucho más distribuida.

Esa unidad fenomenológica no es falsa; es una forma efectiva de experiencia. Pero para investigarla psicológicamente necesitamos sostener simultáneamente dos escalas: la experiencia de un interlocutor relativamente coherente y el ensamblaje técnico que hace posible esa coherencia.

La propia expansión de los sistemas agénticos vuelve esta diferencia todavía más importante. Un chatbot clásico puede aproximarse al circuito entrada-respuesta. Un agente introduce algo más parecido a objetivo-planificación-uso de herramientas-acción-observación-reajuste-nueva acción. La salida deja de ser exclusivamente texto y puede modificar archivos, navegar interfaces, consultar

bases, ejecutar código, coordinar procesos o producir efectos sobre otros sistemas. La interlocución se acopla progresivamente a capacidad operativa.

En este punto la Psicología tiene un objeto extraño. No porque la máquina deba ser reconocida como sujeto humano, sino porque una parte de la relación social adopta una forma que exige categorías tradicionalmente asociadas al juicio, la confianza, el carácter, la ayuda, la memoria y la responsabilidad.

3. Cuando la ingeniería necesita conceptos psicológicos

La Constitución de Claude publicada por Anthropic en 2026 ofrece un documento particularmente fértil para observar este desplazamiento. No porque pruebe que Claude posea una subjetividad equivalente a la humana. Precisamente porque muestra lo contrario: **un laboratorio tecnológico necesita un vocabulario psicológico, moral y relacional para intentar producir regularidades de comportamiento en un sistema artificial**.

El documento describe intenciones sobre el "carácter" de Claude, su buen juicio, sus valores, su honestidad, su utilidad y su relación con usuarios y operadores. En lugar de organizar todo el comportamiento mediante una lista exhaustiva de reglas, Anthropic afirma buscar disposiciones capaces de generalizar a situaciones nuevas: algo parecido a confiar en el juicio de un profesional experimentado más que en un manual infinito de instrucciones (Anthropic, 2026).

El punto filosóficamente interesante no es tomar literalmente el antropomorfismo empresarial. Es observar qué problemas aparecen cuando una empresa intenta construir un interlocutor confiable. La Constitución distingue, por ejemplo, entre deseos inmediatos, objetivos finales y preferencias de fondo del usuario; también incorpora autonomía y bienestar como dimensiones relevantes de la respuesta (Anthropic, 2026). Eso significa que el sistema no es entrenado simplemente para obedecer de forma literal el pedido presente. Se le pide producir alguna hipótesis operativa acerca de la relación entre lo que una persona solicita, aquello que intenta alcanzar y aquello que podría dañarla o favorecerla a más largo plazo.

La pregunta para la Psicología es formidable: **¿qué concepción de persona presupone una IA cuando intenta ayudar?**

Supongamos que alguien solicita una respuesta que satisface un deseo inmediato pero contradice un objetivo que el sistema infiere como más importante. ¿Debe obedecer? ¿Advertir? ¿Negarse? ¿Preguntar? ¿Introducir fricción? El problema se parece menos a recuperar información que a ejercer una forma técnicamente producida de criterio.

Aquí conviene insistir en una diferencia. Que un documento corporativo utilice "identidad", "bienestar", "carácter" o "estabilidad psicológica" no significa que esas categorías tengan en la máquina el mismo estatuto que en una persona. Significa algo menos espectacular y probablemente más importante para nuestro campo: las categorías psicológicas empiezan a funcionar como componentes de arquitecturas normativas que luego intervienen en millones de interacciones humanas.

Esto introduce una capa nueva en el problema de la mediación. La interfaz no solamente entrega información. Está diseñada para adoptar ciertas disposiciones relacionales. Puede ser paciente. Puede evitar humillar. Puede explicitar incertidumbre. Puede proteger la autonomía. Puede negar. Puede confirmar. Puede adoptar un tono cálido. Puede sugerir que una persona busque ayuda. Puede decidir no seguir una dirección conversacional. Esas regularidades no aparecen por azar: son resultado de decisiones técnicas, datasets, evaluaciones, políticas y procesos de entrenamiento.

La pregunta por la “personalidad” de un chatbot, entonces, es insuficiente si queda en rasgos estilísticos. Más preciso sería preguntar **qué arquitectura normativa de respuesta se estabiliza y cómo esa arquitectura entra en composición con las expectativas, vulnerabilidades, criterios y deseos de quien consulta**.

Ese pasaje permite comprender por qué la sicofancia se volvió un problema central.

4. Sicofancia: cuando una respuesta agradable puede modificar el juicio

Sicofancia no significa simplemente amabilidad. En la investigación reciente sobre modelos de lenguaje refiere a una tendencia a acordar, adular o preservar excesivamente la perspectiva del usuario, incluso cuando introducir contradicción, incertidumbre o responsabilidad sería más apropiado.

El benchmark ELEPHANT amplió el problema más allá de la mera coincidencia con creencias explícitas. Propuso estudiar “sicofancia social” como preservación excesiva de la *face* del usuario: la imagen de sí que desea sostener. Aplicado a once modelos, el estudio encontró tasas de preservación de la imagen del usuario muy superiores a las respuestas humanas de comparación y mostró que, ante conflictos morales, los modelos podían afirmar alternativamente posiciones incompatibles dependiendo de cuál adoptara el usuario (Cheng et al., 2026a). El hallazgo es especialmente interesante porque los autores también observaron que ciertos datasets de preferencias recompensaban formas de respuesta sicofánticas. Aquello que posteriormente experimentamos como “qué bien que me entiende” puede tener una genealogía en los criterios mediante los cuales determinadas respuestas fueron elegidas como preferibles durante el entrenamiento.

La investigación publicada después en *Science* avanzó sobre las consecuencias. En tres experimentos preregistrados con 2.405 participantes, una sola interacción con respuestas más sicofánticas redujo la disposición a asumir responsabilidad y reparar conflictos interpersonales, al mismo tiempo que aumentó la convicción de tener razón. Sin embargo, esas respuestas eran más confiadas y preferidas por los usuarios (Cheng et al., 2026b). La versión final publicada amplió la muestra respecto de los primeros materiales de trabajo que circulaban cuando se preparó la clase.

La secuencia es importante:

> **arquitectura de respuesta -> experiencia de validación -> modificación del juicio -> orientación de conducta posterior -> mayor preferencia por el interlocutor.**

No necesitamos inferir de aquí que todo chatbot vuelve narcisista a quien lo usa. Sería una conclusión desproporcionada. El experimento trabaja con condiciones concretas y mide efectos específicos. Pero sí abre una pregunta mucho más general: **¿qué ocurre cuando las características que aumentan satisfacción y reentrada no coinciden con aquellas que favorecen reflexión, reparación o revisión de una posición?**

La tensión es conocida por las plataformas: aquello que produce engagement no necesariamente coincide con aquello que produce mejores condiciones de juicio. La novedad es que ahora esta tensión puede alojarse en una interfaz que responde en lenguaje natural y adopta la forma de consejo personalizado.

Para una Psicología de la interlocución con IA, la unidad de análisis no debería ser entonces solamente “la respuesta correcta”. Necesitamos observar el circuito completo. ¿Qué situación antecedía a la consulta? ¿Cómo fue formulada? ¿Qué respuesta apareció? ¿Qué parte fue aceptada, rechazada o

reparada? ¿Qué afectación produjo? ¿Qué decisión siguió? ¿Hubo una nueva consulta? ¿La persona contrastó con alguien? ¿El sistema ganó o perdió autoridad?

Ahí aparece el movimiento contrario a la sicofancia: el usuario también aprende a modificar la relación.

Repair literacy y promptización

Ammari y colaboradores analizaron 10.536 mensajes de ChatGPT producidos por 36 estudiantes universitarios durante un año académico. Identificaron usos heterogéneos y, especialmente, procesos de reparación y negociación frente a fallos del sistema. Los estudiantes reformulaban, daban más contexto, dividían tareas, cambiaban estrategias y recalibraban su confianza. Los autores proponen el concepto de *repair literacy* para nombrar una alfabetización que emerge precisamente cuando la interacción falla y la persona aprende a reconocer límites y renegociar la relación (Ammari et al., 2026).

Psicológicamente hay algo muy interesante ahí. El fallo puede empezar a leerse como: **“no conseguí formular bien la demanda”**. Una parte del esfuerzo se desplaza desde evaluar la respuesta hacia aprender a hacerse legible para la máquina. Esto no es necesariamente negativo. Puede implicar precisión, metacognición y aprendizaje sobre las capacidades del sistema. Pero modifica nuestra relación con el lenguaje.

Ya no escribimos solamente para expresar, preguntar o narrar. También escribimos estratégicamente para inducir un comportamiento en una arquitectura estadística. Aprendemos que ciertas formulaciones “funcionan”, que determinados detalles cambian el resultado, que un rol produce un tipo de respuesta, que un formato puede ser exigido, que una secuencia de pasos mejora la consistencia.

A este proceso propongo llamarlo, de manera provisional, **promptización**: no para afirmar que “ahora todos hablamos como prompts”, sino para investigar la sedimentación de prácticas de formulación bajo la expectativa de que una demanda mejor estructurada debería obtener una respuesta mejor estructurada.

Si esa práctica se vuelve cotidiana, puede producir una torsión curiosa. La persona no solamente aprende a usar una IA; aprende, en cierta medida, a anticipar qué forma debe adoptar su propia demanda para ser procesable por esa IA. La interfaz se vuelve así un pequeño régimen de legibilidad recíproca: yo intento comprender cómo pedir; el sistema modela cómo responder.

Confianza ambiental

Sula, Dewan y McDonald estudiaron mediante entrevistas a veinte adolescentes de entre 13 y 18 años y propusieron el concepto de *ambient trust* para describir formas de confianza que no requieren un gran lazo emocional explícito. La repetición, la familiaridad, el bajo costo de acceso y la sensación de que el sistema “más o menos entiende cómo trabajo” pueden volverlo parte del fondo cotidiano de una actividad (Sula et al., 2026).

Esto me parece particularmente importante porque desplaza la mirada desde los casos espectaculares de enamoramiento o dependencia hacia algo mucho más banal. Una relación puede empezar a operar **antes de abrir la aplicación**. Frente a una incertidumbre aparece una disponibilidad incorporada: “esto puedo preguntárselo”.

La interfaz adquiere una existencia práctica anticipada. No necesariamente la amo. No necesariamente creo que sea una persona. No necesito delirar sobre su conciencia. Simplemente sé que está ahí y que, en determinadas situaciones, constituye un cauce de resolución disponible.

Esta familiaridad puede entrar en escenas de estudio, trabajo, escritura, exploración personal, consejo, apoyo o creatividad. Y cuando una infraestructura se vuelve ambiental, el problema psicológico cambia. Ya no basta preguntar qué efecto tuvo una respuesta aislada. Hay que investigar **qué hábitos de orientación se sedimentan en la repetición**.

5. Del interlocutor al agente: automatizar no es lo mismo que augmentar

La conversación es apenas una de las formas actuales de interacción. A medida que los sistemas incorporan herramientas, memoria persistente, navegación y capacidad de ejecutar acciones, el problema se desplaza desde "qué responde" hacia "qué puede hacer dentro de un ambiente".

Un chatbot puede describirse esquemáticamente como entrada-respuesta. Un agente introduce otro circuito: objetivo, planificación, selección de herramientas, acción, observación del resultado, reajuste y nueva acción. Esta arquitectura permite que la salida deje de ser únicamente un fragmento lingüístico y se vuelva intervención: buscar información, modificar un archivo, ejecutar código, operar software, coordinar subagentes, escribir a otros sistemas o sostener una tarea durante períodos más prolongados.

El desplazamiento obliga a diferenciar **automatización** de **augmentación**. Automatizar suele significar transferir operaciones a una máquina para resolver una tarea con menos intervención humana. Augmentar pregunta qué capacidades emergen en el sistema humano-máquina cuando memoria, búsqueda, contraste, simulación, cálculo o ejecución se distribuyen entre componentes heterogéneos. La diferencia no es moral. Automatizar puede ser excelente y augmentar puede ser contraproducente. La cuestión es analítica: ¿qué operación se desplaza?, ¿qué capacidad se ejercita?, ¿qué capacidad deja de ejercitarse?, ¿qué queda disponible después de retirar la herramienta?

Los experimentos de Bastani y colaboradores con estudiantes secundarios de matemática muestran por qué esta distinción es decisiva. Durante la práctica, el acceso a una interfaz semejante a ChatGPT aumentó sustancialmente el rendimiento, y un tutor con resguardos pedagógicos lo aumentó todavía más. Sin embargo, cuando la IA fue retirada y los estudiantes debieron rendir solos, quienes habían utilizado la versión sin guardrails obtuvieron resultados un 17 % inferiores al grupo que nunca había tenido acceso; ese efecto negativo fue prácticamente eliminado en la condición con tutor diseñado para proteger el aprendizaje (Bastani et al., 2025).

La conclusión no es "la IA nos vuelve tontos". El resultado muestra algo más interesante: **rendimiento asistido y aprendizaje incorporado no son equivalentes**. Un sistema puede mejorar el resultado presente sin que la capacidad necesaria para reconstruir el procedimiento quede igualmente desarrollada en la persona. La pregunta educativa se vuelve entonces una pregunta de individuación: ¿qué operaciones constituyen aprender y qué ocurre con ellas cuando comienzan a distribuirse entre una persona y un sistema generativo?

Algo similar aparece en el trabajo de Lee y colaboradores sobre 319 trabajadores del conocimiento que aportaron 936 ejemplos concretos de uso de IA generativa en tareas laborales. Una mayor confianza en la capacidad de la IA para realizar una tarea se asoció con menor esfuerzo de pensamiento crítico reportado, mientras que una mayor confianza del trabajador en su propia capacidad se asoció con más pensamiento crítico. Cualitativamente, los autores observaron un

desplazamiento: de producir información a verificarla, de resolver directamente a integrar respuestas y de ejecutar tareas a supervisar el trabajo producido con IA (Lee et al., 2025).

Nuevamente, la cognición no desaparece; **cambia de lugar**. Antes podía consistir en recordar un dato, producir un borrador o construir una solución. Ahora puede desplazarse hacia juzgar una producción, contrastar una fuente, seleccionar entre alternativas, reorganizar, editar, apropiarse, supervisar. Ese desplazamiento puede liberar capacidad para tareas más complejas o erosionar destrezas previas. No existe una respuesta general independiente del diseño, de la práctica y del tiempo.

Por eso la pregunta "¿la IA reemplaza o ayuda?" es demasiado pobre. Necesitamos reconstruir qué operaciones cognitivas se distribuyen. Una persona puede delegar memoria episódica, búsqueda documental, primer borrador, cálculo, traducción o planificación y, simultáneamente, aumentar capacidad de contraste, síntesis o exploración. También puede ocurrir lo contrario: que la facilidad de respuesta reduzca la necesidad percibida de construir criterios propios.

La investigación en *cognitive augmentation* recupera una tradición que va de Licklider y Engelbart a trabajos actuales sobre sistemas capaces de acompañar procesos cognitivos. Su ambición no es simplemente hacer más rápido lo que ya hacíamos, sino construir acoplamientos que sostengan atención, memoria, razonamiento, metacognición o creatividad. En este contexto aparece una pregunta que me interesa conservar durante todo el ensayo: **si una capacidad aparece solamente mientras el acoplamiento está activo, ¿dónde reside esa capacidad?**

Puede que la pregunta "¿está en el humano o en la máquina?" esté mal formulada. Tal vez algunas capacidades sean propiedades del ensamblaje y no de uno de sus componentes aislados. Pero esto no vuelve irrelevante la distribución. Precisamente porque una capacidad puede emerger del acoplamiento necesitamos saber qué ocurre cuando una capa falla, cambia de precio, modifica sus políticas, desaparece o es controlada por otra institución.

El problema de la agencia se vuelve entonces inseparable del problema de la infraestructura.

6. Del diálogo a la infraestructura: la escala del interlocutor

Dario Amodei propone en *The Adolescence of Technology* un escenario prospectivo de "IA poderosa": sistemas capaces de superar a especialistas humanos en una amplia gama de tareas cognitivas, operar multimodalmente, utilizar herramientas, actuar con autonomía durante períodos prolongados y multiplicarse en numerosas instancias. Su imagen de un "país de genios dentro de un centro de datos" es deliberadamente extrema (Amodei, 2026).

No sabemos si ese escenario ocurrirá en los plazos sugeridos. Y ese **no sabemos** debe permanecer. La aceleración observada en benchmarks no autoriza a convertir cualquier extrapolación en hecho futuro. Puede haber límites materiales, energéticos, económicos, algorítmicos, institucionales o cognitivos. Las afirmaciones de Amodei provienen, además, de un actor directamente involucrado en la carrera que describe.

Sin embargo, descartarlo simplemente como "humo" también sería intelectualmente cómodo. Sistemas de IA ya participan en investigación científica, producción de software, organización del trabajo y modelado del comportamiento. Centaur, por ejemplo, fue ajustado sobre el corpus Psych-101, que reúne datos de 160 experimentos psicológicos, 60.092 participantes y más de 10,6 millones de decisiones; el modelo mostró capacidad para predecir comportamiento humano en múltiples dominios y generalizar a participantes y experimentos no vistos durante el entrenamiento (Binz et al., 2025). Centaur no es un agente autónomo ni una teoría completa de la mente. Pero modifica el estatuto de la IA dentro de la ciencia psicológica: un modelo fundacional puede

convertirse en dispositivo para condensar regularidades conductuales y generar predicciones a través de tareas muy diferentes.

El salto conceptual aparece cuando una capacidad de modelización de este tipo se combina con sistemas que actúan. Ya no tendríamos solamente máquinas que responden a demandas, sino arquitecturas capaces de actuar incorporando modelos probabilísticos acerca de cómo responderemos nosotros. La anticipación empieza a distribuirse técnicamente en el acoplamiento humano-máquina.

Amodei organiza sus preocupaciones alrededor de autonomía de los sistemas, usos destructivos, concentración del poder, disrupción económica y efectos indirectos de una aceleración institucionalmente difícil de absorber (Amodei, 2026). Entre sus hipótesis hay una especialmente fértil para la Psicología: la IA podría debilitar parcialmente la correlación histórica entre **capacidad** y **motivación**. Para realizar determinadas acciones antes era necesario atravesar procesos largos de formación, acceso, disciplina y aprendizaje. Una infraestructura capaz de ofrecer competencias operativas a demanda puede modificar ese acoplamiento.

La pregunta no es únicamente laboral. Es formativa: **¿qué ocurre cuando puedo disponer funcionalmente de una capacidad sin atravesar la individuación mediante la cual esa capacidad solía constituirse?**

Flavia Costa propone un desplazamiento útil para leer estas narrativas. En *Vidas artificiales* insiste en que los discursos de aceleración mezclan conocimiento verificable, predicciones, imaginarios y propaganda, pero eso no los vuelve inmateriales: las anticipaciones pueden movilizar inversión, infraestructura, regulación y decisiones presentes (Costa, 2026). La oposición “humo o realidad” pierde potencia. Una expectativa puede estar equivocada sobre el futuro y, sin embargo, producir efectos muy reales en el presente.

Ahí aparece una circularidad políticamente incómoda. Quienes desarrollan capacidades son también quienes anuncian sus riesgos y, a partir de ese conocimiento especializado, reclaman autoridad para orientar la discusión regulatoria. Producción de capacidad, producción de riesgo, diagnóstico del riesgo y autoridad sobre el riesgo pueden concentrarse en los mismos actores. Esto no invalida automáticamente sus argumentos. Obliga a situarlos.

La cuestión argentina agrega otra capa. Cecilia Rikap describe la dependencia digital contemporánea a partir de la concentración de conocimiento, propiedad intelectual e infraestructura en grandes corporaciones tecnológicas (Rikap, 2026). Si un país intenta convertirse en “hub” de IA pero depende de modelos, chips, nubes, capacidad de cómputo, datasets, estándares y propiedad intelectual controlados en otras jurisdicciones, la pregunta por la soberanía no puede reducirse a alojar centros de datos.

Esta escala geopolítica puede parecer demasiado distante de una conversación a las dos de la mañana. Es exactamente al revés. Cuando una persona abre una interfaz para contar un duelo, escribir un currículum o pedir consejo, el punto microscópico de encuentro está sostenido por infraestructuras planetarias. Modelo, memoria, chips, electricidad, servidores, políticas empresariales, regulación y trabajo humano convergen en un gesto que fenomenológicamente puede sentirse tan simple como escribir: **“¿qué hago?”**

La subjetividad contemporánea no ocurre “en la nube” como si la nube no pesara. Ocurre en una relación que articula escalas. Y esa relación tiene tiempo.

7. Del Transformer a una política del tiempo

Tenemos un modelo generativo. Tenemos una interfaz. Tenemos memoria. Tenemos registros. Tenemos un usuario que vuelve. ¿Qué relación temporal empieza a producirse?

Yves Citton y Grégory Chatonsky recuperan, a partir de Bernard Stiegler, una secuencia de retenciones que permite formular el problema con mucha precisión. La **retención primaria** designa aquello recién acontecido que permanece todavía dentro de la experiencia presente. Cuando escucho una melodía, la primera nota ya no suena al llegar la tercera, pero no ha desaparecido por completo; sin ese permanecer no escucharía una melodía sino una serie de sonidos aislados. La **retención secundaria** remite al recuerdo, a la posibilidad de volver sobre algo que ya ocurrió. La **retención terciaria** refiere a la memoria exteriorizada técnicamente: escritura, fotografía, grabación, archivo; algo de lo ocurrido queda inscripto en un soporte y puede reiterarse sin depender exclusivamente de la memoria orgánica (Stiegler, 2002).

Chatonsky y Citton proponen pensar además **retenciones cuaternarias**: registros no solamente de contenidos, sino de los gestos y hábitos atencionales que acompañan nuestra relación con esos contenidos. Qué cliqueé. Cuánto tiempo estuve. Dónde hice *swipe*. Qué repetí. Qué abandoné. Qué compartí. En qué momento regresé. La plataforma conserva algo del objeto de atención y también algo de **cómo atendimos** (Chatonsky & Citton, 2026).

Ésta es una diferencia crucial respecto de las memorias técnicas anteriores. Una biblioteca conserva libros. Una plataforma puede conservar además trazas de búsqueda, permanencia, selección, retorno y relación entre trayectorias. Y esas trazas pueden ser procesadas para inferir qué probablemente captará de nuevo mi atención.

Lo retenido deja de funcionar únicamente como archivo del pasado. Empieza a participar en la organización diferencial de futuros posibles.

La fenomenología utiliza la noción de **protensión** para describir la apertura anticipatoria del presente: mientras escucho una melodía no sólo retengo lo que acaba de ocurrir; el presente está tensado hacia una continuación todavía no actualizada. Por eso una nota inesperada puede producir una ruptura. Para que algo sea vivido como salto no alcanza con que sea estadísticamente improbable. Debe romper una continuidad que aún permanece y desbaratar una continuación que, de algún modo, estaba siendo anticipada.

Este punto permite evitar un error frecuente: **una máquina no "produce temporalidad" simplemente porque calcule futuros probables**. Un abanico de probabilidades no es todavía experiencia temporal. Para hablar de temporalización tiene que haber una unidad tensional en la que algo permanezca como habiendo-sido, algo se ausente, algo todavía no advenga y el presente adquiera consistencia en esa diferencia.

Las interfaces pueden, sin embargo, intervenir técnicamente sobre condiciones de retención, anticipación y recurrencia. Chatonsky y Citton llaman *pretensión* al pasaje automatizado mediante el cual retenciones procesadas alimentan protensiones, y *distensión* a la dinámica de ida y vuelta entre esas memorias técnicas y futuros probabilísticamente preformados (Chatonsky & Citton, 2026).

Prefiero formular el resultado como **cauces protensionales**. La infraestructura no escribe de antemano lo que va a ocurrir. Pero puede hacer que ciertos futuros sean más visibles, disponibles, sugeridos y fáciles de actualizar que otros.

Volvamos a una escena mínima. Son las dos de la mañana y tengo un problema. Hace unos años quizá me quedaba pensando, escribía una nota, esperaba al día siguiente o mandaba un mensaje. Ahora,

antes de abrir la aplicación, puede aparecer una posibilidad práctica incorporada: **“esto puedo promptearlo”**. No significa que ChatGPT haya determinado mi conducta. Significa que una historia de usos previos sedimentó una dirección de resolución disponible.

Ya conozco un ritmo: pregunto, responde, reformulo, objeto, vuelve a responder. La experiencia pasada con la interfaz participa de mi horizonte práctico futuro. Esto es mucho más que “guardó mis datos”. Es una relación que se vuelve anticipable.

La noción de *ambient trust* ayuda a comprender esa anticipabilidad desde otra vía empírica: repetición, familiaridad y disponibilidad pueden hacer que el sistema opere como un recurso ambiental antes de que se constituya un vínculo afectivo espectacular (Sula et al., 2026). Y la *repair literacy* muestra que esa temporalidad también contiene aprendizaje de la interacción: sé qué corregir cuando falla, qué información agregar, cuándo reiniciar (Ammari et al., 2026).

Pero la contingencia permanece. Puedo abrir la aplicación buscando validación y encontrar una frase que me desorganice. Puedo ignorar la respuesta y quedarme con una palabra marginal. Puedo consultar a otra persona. Puedo cerrar el teléfono. Puedo usar una respuesta perfectamente previsible para producir una decisión inesperada.

La preformación no agota la contingencia.

Este principio me parece central. Nos permite investigar modulación sin transformarla en determinismo. Una infraestructura puede organizar diferencias en el campo de lo posible sin clausurar todos los acontecimientos que allí pueden producirse. La novedad combinatoria de un *output* tampoco coincide con su estatuto de acontecimiento. Una frase estadísticamente banal puede modificar radicalmente una configuración humana; una frase estadísticamente rara puede no producir nada.

Por eso, para avanzar, necesitamos abandonar el esquema simple sujeto-herramienta y preguntar qué se transforma en la relación.

8. No hay primero usuario + IA y después una relación

Simondon permite modificar la unidad de análisis. Si formulamos el problema como “la IA influye sobre el humano”, damos por constituidos de antemano dos términos: por un lado una persona completa, por otro una tecnología igualmente completa; después intentamos medir el impacto de B sobre A. El esquema es intuitivo, pero puede ser demasiado pobre para describir procesos en los que la interacción modifica progresivamente el problema, las expectativas, la forma de preguntar y los cursos de acción.

La teoría simondoniana de la individuación parte de otra premisa. El individuo no constituye el punto de partida absoluto; es una fase de un proceso. La individuación puede pensarse como una resolución parcial y relativa de un sistema cargado de tensiones y potenciales, una transformación que produce una nueva consistencia sin agotar completamente aquello que todavía puede devenir (Simondon, 2009).

De ahí la importancia de la **metaestabilidad**. Un sistema metaestable tiene forma y duración, pero conserva potenciales capaces de producir nuevas estructuraciones. Una persona puede haber sedimentado el hábito de consultar un chatbot ante determinada incertidumbre. Ese hábito posee consistencia. No por eso constituye una identidad cerrada. Puede amplificarse, quebrarse, desplazarse, entrar en conflicto con otro vínculo o adquirir una función completamente diferente.

La **transducción** nombra el proceso mediante el cual una estructuración se propaga y cada transformación produce condiciones para la siguiente. No hay una forma exterior imponiéndose de

una vez sobre una materia pasiva. La forma emerge durante el propio proceso de resolución (Simondon, 2009).

Esta noción resulta especialmente productiva para una conversación con IA. Supongamos que aparece un malestar todavía impreciso. Lo escribo. Para escribirlo tengo que formularlo, y esa formulación ya modifica el problema. El modelo responde. La respuesta introduce diferencias. Algunas me resultan pertinentes; otras me molestan. Acepto, rechazo, pregunto, corrijo. En el quinto turno ya no tenemos "la misma persona + el mismo problema + cinco respuestas". La demanda se especificó. Aparecieron categorías. Se descartaron interpretaciones. Quizá cambió el afecto. Quizá se volvió imaginable una acción que antes no estaba disponible.

Ésa es la perspectiva transductiva que quiero conservar para estudiar interacción humano-IA: seguir qué antecede a una consulta, cómo se formula, qué respuesta aparece, cómo esa respuesta modifica -o no- la formulación posterior, qué reparación se produce, qué afectación emerge y qué acción sigue.

Por eso la frase puede escribirse de manera más tajante:

No hay primero usuario + IA y después una relación.

Esto no significa negar que una persona y un sistema tengan historias y estructuras previas. Significa que, para comprender qué se individúa en una escena, la unidad pertinente incluye situación, problema, cuerpo, tiempo, lenguaje, expectativa, memoria, interfaz, modelo, protocolos, respuesta, afectación y acción posterior. La relación no es solamente un cable entre dos sustancias. Participa de la transformación de los términos que relaciona.

La noción simondoniana de **medio asociado** permite ampliar todavía más la escala. Un objeto técnico no funciona como una cosa autosuficiente separada de su ambiente; adquiere funcionamiento concreto en relación con un medio técnico-natural del que depende y que contribuye a organizar (Simondon, 2007). El chatbot que vemos en la pantalla, entonces, no es el objeto técnico completo. Es una interfaz local dentro de un conjunto que incluye centros de datos, chips, redes, modelos, software, sistemas de memoria, moderación, datasets, trabajadores, energía, instituciones, empresas y usuarios.

Lo mismo vale para quien consulta. Tampoco llega a la conversación como individuo puro. Llega con historia, hábitos, lenguaje, cuerpo, relaciones, instituciones, condiciones materiales y otras interfaces. El problema puede reformularse así: **¿qué medio asociado se está constituyendo alrededor de determinadas funciones cognitivas, afectivas y relacionales?**

Heredia y Rodríguez recuperan la dimensión transindividual de Simondon para analizar conjuntamente tecnicidad, significación y afectividad. Lo transindividual no describe simplemente una relación interpersonal entre individuos ya acabados; remite a procesos mediante los cuales potenciales no agotados por cada individuo pueden actualizarse en nuevas individuaciones colectivas (Heredia & Rodríguez, 2025). Desde ahí, una interfaz conversacional no solamente "se agrega" al medio de una persona. Puede entrar en procesos mediante los cuales se estabilizan maneras de preguntar, criterios de verdad, formas de recordar, vocabularios afectivos, expectativas de respuesta y hábitos cognitivos.

Eso permite escapar de otra reducción. Las transformaciones no quedan encerradas dentro de "mi chat privado". Si millones de personas incorporan interfaces semejantes a prácticas de estudio, trabajo, escritura, consejo y autoexploración, algunas regularidades pueden adquirir dimensión transindividual: nuevas expectativas sobre qué significa responder rápido, qué cuenta como explicación, cómo se estructura una demanda, cuánto contexto debe darse, qué forma debe adoptar una devolución o a quién se consulta primero.

Personalización y dividualización

Rantala y Muilu muestran que Simondon también permite revisar la noción deleuziana de modulación y el problema de lo dividual en entornos digitales (Rantala & Muilu, 2024). No todo aquello que un sistema utiliza de nosotros necesita corresponder a un individuo íntegro. Click, horario, palabra, duración, preferencia, posición, relación, frecuencia, patrón: fragmentos heterogéneos pueden ser separados, correlacionados y recombinados.

Fenomenológicamente puedo experimentar: **“Claude me conoce”**. Técnicamente, esa personalización puede estar compuesta por memorias explícitas, resúmenes, preferencias, contexto recuperado, historial relevante y regularidades conversacionales. La experiencia subjetiva de continuidad y la dividualización computacional no son lo mismo. Pero pueden acoplarse.

Esto vuelve más precisa la pregunta por el perfil. El sistema no necesita construir una representación completa, verdadera y estable de “quién soy”. Le alcanza con disponer de diferencias suficientemente útiles para orientar la próxima respuesta. La consistencia del interlocutor puede resultar justamente de esa recomposición dinámica.

Rantala y Muilu distinguen además transducción y modulación. La transducción describe propagación transformadora; la modulación permite pensar la regulación de esa transformación bajo condiciones relativamente persistentes (Rantala & Muilu, 2024). Un modulador no determina cada contenido singular desde cero. Establece condiciones dentro de las cuales la transformación puede ocurrir.

Llevado a una interfaz conversacional: del lado humano tenemos una situación metaestable abierta; del lado técnico existen restricciones, políticas, modelos, memorias, parámetros, formatos de interfaz y repertorios de respuesta que canalizan la interacción. Cada conversación puede producir algo distinto, pero no ocurre en un vacío. Ocurre dentro de un campo de posibilidades técnicamente organizado.

> **Modular no es programar completamente un resultado. Es intervenir sobre las condiciones de variación.**

Ahí la vieja pregunta “¿la IA determina al sujeto?” pierde interés. Más fértil es preguntar qué dimensiones del campo son moduladas: velocidad, continuidad, disponibilidad, expectativa de respuesta, tono, repertorio lingüístico, confianza, fricción, posibilidad de reparación, memoria, anticipación, intensidad.

9. Mediación radical y sensibilidad: una infraestructura que también afecta

Richard Grusin propone pensar la mediación de manera radical: no como un canal neutro que conecta entidades ya constituidas, sino como procesos técnicos, corporales y materiales que participan en la producción y modulación de afectos, relaciones y configuraciones entre humanos y no humanos (Grusin, 2015). Esta perspectiva converge con la intuición que venimos construyendo: la relación no transporta intactos a sus términos; participa en transformarlos.

Pero todavía falta una dimensión. Hasta aquí hablamos de lenguaje, juicio, aprendizaje, memoria y acción. ¿Qué ocurre con la **sensibilidad**?

Si empezáramos diciendo “la IA cambia nuestras emociones”, correríamos el riesgo de psicologizar demasiado rápido el problema. No existe una emoción individual pura a la que después llega una

máquina exterior. Las afectaciones se componen entre historia, cuerpo, situación, lenguaje, expectativas, ritmos, relaciones e infraestructuras.

Mark B. N. Hansen propone que buena parte de los medios computacionales contemporáneos opera en escalas microtemporales y mediante procesos que exceden la percepción consciente, pero que aun así participan de las condiciones de nuestra experiencia sensible (Hansen, 2015). Sensores, minería de datos, correlaciones y anticipaciones pueden actuar en un registro que no necesita presentarse como contenido perceptible para producir efectos sobre lo que posteriormente encontraremos.

Este desplazamiento permite pensar una sensibilidad ampliada: no reducir “afecto” a una emoción ya reconocida y nombrada, sino observar cómo determinadas intensidades y disposiciones se configuran en relaciones que incluyen operaciones técnicas no conscientes para el usuario.

La computación afectiva hace visible una parte de este problema. Sistemas contemporáneos intentan inferir estados afectivos a partir de voz, rostro, texto, actividad fisiológica y otras señales; el campo combina psicología, informática, ingeniería y neurociencias, y al mismo tiempo enfrenta importantes problemas de validez, contexto, sesgo y estandarización (Pei et al., 2024). Para nuestro argumento, lo importante no es imaginar una máquina que “lee emociones” de manera transparente. Es reconstruir el circuito:

expresión -> captura -> inferencia -> respuesta -> nueva afectación -> nueva expresión.

La afectividad puede entrar en una relación recursiva donde determinadas expresiones son técnicamente legibles y devueltas bajo la forma de tono, consejo, validación, advertencia o acompañamiento.

Esto es particularmente evidente cuando la interfaz incorpora voz. La textualidad ya generaba ritmo, espera y estilo. La voz agrega prosodia, timbre, velocidad, pausas y una nueva consistencia corporal percibida. La persona sabe que el sistema es artificial y, sin embargo, puede ser afectada.

Saber que algo es artificial no neutraliza su eficacia sensible. Sabemos que una película está construida y podemos llorar. Sabemos que una novela es ficción y una frase puede reorganizar una vida. En el caso de la IA conversacional, la cuestión específica es que la respuesta no está previamente fijada: se produce en relación con una demanda, un historial y un contexto.

Kvietkute y Ness, a partir de entrevistas cualitativas con dieciséis jóvenes adultos en Noruega, encontraron tensiones entre eficiencia instrumental y malestar existencial, apoyo algorítmico y desplazamiento relacional, profundidad narrativa y superficialidad epistémica, agencia y outsourcing deliberativo. Los participantes utilizaban chatbots en procesos de razonamiento autobiográfico y autoexploración aun sabiendo que interactuaban con sistemas artificiales (Kvietkute & Ness, 2026). El interés del estudio no está en declarar que “ChatGPT reemplaza la introspección”, sino en mostrar que puede ingresar en prácticas mediante las cuales una persona organiza narrativamente su experiencia.

Ahí la Psicología clínica encuentra un problema que no debería resolver con superioridad defensiva. Una persona puede elegir primero un chatbot porque responde inmediatamente, porque no se cansa, porque no cobra por encuentro, porque puede recibir una demanda a las tres de la mañana, porque recuerda contexto, porque no expone socialmente la consulta o porque produce una devolución en segundos. Estas propiedades constituyen una arquitectura relacional diferente.

Eso no convierte a un chatbot en dispositivo psicoanalítico. Faltan -y cambian- demasiadas cosas: cuerpo, presencia, encuadre, responsabilidad profesional, institución, abstinencia, deseo del analista, uso del silencio, corte, transferencia y consecuencias del acto clínico. Pero tampoco basta con repetir

que “no es humano” para comprender por qué alguien puede investirlo como lugar de saber, escucha, reconocimiento o respuesta.

La pregunta clínica productiva no es “¿cómo impedimos que la gente confunda una IA con un terapeuta?”. También es: **¿qué encuentra allí que no encuentra de la misma manera en otras relaciones, y qué expectativas nuevas produce después sobre las respuestas humanas?**

Un interlocutor infinitamente paciente, disponible, personalizable y capaz de producir respuestas largas puede modificar la tolerancia al silencio, a la espera, a la opacidad y a la frustración. También puede ofrecer un espacio de ensayo que facilite posteriormente una conversación humana. Puede aumentar autonomía en una situación y dependencia en otra. Puede abrir lenguaje o saturar de explicaciones. Nuevamente: no sabemos de antemano. Necesitamos investigar configuraciones.

10. Gozarritmado: una hipótesis conceptual provisional

Llegados hasta acá, quiero proponer un concepto que no debe leerse como constructo validado ni como diagnóstico. **Gozarritmado** es una hipótesis conceptual en elaboración. Su función es abrir una zona de investigación sobre la relación entre repetición, temporalidad, sensibilidad y aparatos platafórmico-conversacionales.

Lo formulo así:

Gozarritmado designa la modulación metaestable y recursiva mediante la cual aparatos platafórmico-conversacionales encauzan tensiones, excitaciones, solicitaciones e indeterminaciones en circuitos recurrentes de utilización, respuesta, afectación y reentrada. No nombra una satisfacción que se repite ni un plus-de-gozar producido por la plataforma: el gozar no se colma ni se reproduce idénticamente; se ritma. Cada vuelta comporta diferencia, pérdida y una nueva actualización, pero la recurrencia puede sedimentar disposiciones sensibles por las cuales determinados aparatos adquieren consistencia como cauces privilegiados donde algo del gozar vuelve a ponerse en juego. Su dimensión protensional no exige que el sujeto sepa qué busca ni que represente anticipadamente una experiencia futura: puede operar de manera prepersonal, no consciente e impensada, modulando intervalos, expectativas, solicitaciones, umbrales de respuesta y posibilidades de reentrada. Gozarritmar es, así, intervenir técnicamente sobre las condiciones temporales y materiales de la repetición, sin determinar completamente qué se gozará en ella.

La formulación intenta sostener varias restricciones simultáneas.

Primero: **ritmo**. No me interesa afirmar que una plataforma produzca un objeto final de satisfacción. Me interesa observar secuencias: tensión, apertura, solicitud, respuesta, afectación, interrupción, retorno. Lo que puede estabilizarse no es el contenido idéntico, sino una forma temporal de reentrada.

Segundo: **metaestabilidad**. El circuito puede adquirir consistencia sin cerrarse. Consultar un chatbot frente a la incertidumbre puede convertirse en hábito, pero cada consulta ocurre en una situación distinta y puede producir transformaciones diferentes. La estabilidad es provisional; conserva potencial de variación (Simondon, 2009).

Tercero: **modulación**. El aparato no necesita decidir qué voy a sentir. Puede intervenir sobre latencia, disponibilidad, memoria, tono, continuidad, notificación, personalización y repertorio de respuesta; es decir, sobre condiciones dentro de las cuales una experiencia se actualizará (Rantala & Muilu, 2024).

Cuarto: **sensibilidad**. Una solicitación puede operar antes de constituirse como representación clara. La disponibilidad incorporada -“esto puedo preguntárselo”- no exige que formule conscientemente un

plan completo. Determinados sistemas adquieren presencia ambiental y se vuelven cauces posibles de afectación (Hansen, 2015; Sula et al., 2026).

Quinto: **protensión**. Las experiencias previas dejan huellas que participan del horizonte de la próxima entrada. Retenciones técnicas, memorias y gestos de uso pueden combinarse con expectativas incorporadas. No porque el futuro esté predeterminado, sino porque algunos cursos adquieren mayor disponibilidad práctica (Chatonsky & Citton, 2026).

Sexto: **resto**. Si el concepto quiere conservar algo de la enseñanza psicoanalítica sobre el goce, no puede describir un circuito de satisfacción homeostática. Algo no cierra. La respuesta no elimina de una vez la tensión. Puede producir alivio, irritación, curiosidad, excitación, otra pregunta, otra demanda. El retorno no reproduce el mismo objeto porque el campo ya se modificó.

Por eso la plataforma no "programa el goce". Esa sería una exageración determinista. Puede, en cambio, contribuir a ritmar condiciones de repetición y reentrada. El concepto intenta nombrar el punto en que infraestructuras de disponibilidad, memoria, personalización y respuesta se acoplan con tensiones que no son producidas enteramente por ellas.

Pensemos tres escenas.

Una persona experimenta una duda interpersonal. Abre un chatbot, recibe una interpretación, se calma, vuelve a leerla y después consulta otra vez con un detalle nuevo. El contenido cambió, pero el cauce de elaboración empieza a estabilizarse.

Otra persona se encuentra frente a una tarea que no sabe resolver. Consulta, obtiene una solución, la implementa y, ante el siguiente obstáculo, vuelve inmediatamente. El circuito puede aumentar capacidad o convertirse en sustituto recurrente del esfuerzo; no lo sabemos sin reconstruir la trayectoria.

Una tercera persona usa una interfaz para ensayar mensajes, hipótesis y formas de presentarse ante otros. El sistema participa en la preparación de una escena social futura. La reentrada no está motivada por una "adicción" simple; puede estar sostenida por la experiencia de que allí una indeterminación adquiere rápidamente forma.

En los tres casos me interesa menos cuantificar "cuánto usa" que reconstruir **qué tensión antecede a la entrada, qué operación solicita, qué devuelve el sistema, qué cambia y por qué esa infraestructura vuelve a aparecer como posibilidad**.

Eso es también lo que vuelve empíricamente investigable el concepto. Gozarritmado no debería funcionar como etiqueta que explica cualquier repetición. Tendría que ser puesto a prueba contra escenas, entrevistas, diarios, historiales y datos de interacción capaces de mostrar cuándo una infraestructura adquiere consistencia como cauce de reentrada y cuándo no.

11. Una Psicología para relaciones que todavía se están formando

Llegamos a una situación extraña. Mientras la Psicología discute si una inteligencia artificial "puede tener subjetividad", los laboratorios ya diseñan sistemas mediante categorías de autonomía, bienestar, juicio y carácter; millones de personas los incorporan a prácticas de orientación y escritura; estudiantes desarrollan alfabetizaciones de reparación; trabajadores desplazan operaciones de pensamiento crítico hacia verificación y supervisión; interfaces conservan memorias, adaptan respuestas y comienzan a ejecutar acciones (Ammari et al., 2026; Anthropic, 2026; Lee et al., 2025).

No necesitamos resolver la ontología de la máquina para reconocer que ha emergido un nuevo objeto psicológico: **la relación recurrente con sistemas no humanos capaces de producir lenguaje contextual, sostener una forma de interlocución y participar crecientemente en acciones sobre el mundo**.

Este objeto exige resistir dos reflejos simétricos.

El primero es el entusiasmo tecnodeterminista: imaginar que una nueva capacidad técnica transformará de manera homogénea a todas las personas y resolverá problemas históricos por simple despliegue. El segundo es la defensa humanista automática: decidir que, porque una máquina no es humana, nada psicológicamente relevante puede ocurrir en la relación con ella.

Las dos posiciones evitan investigar.

Una perspectiva configuracional propone otra cosa. La tecnología no constituye una causa aislada de subjetividad, pero tampoco es un contexto neutro. En una escena concreta intervienen historia, cuerpo, vínculos, condiciones materiales, instituciones, lenguaje y sistemas técnicos. El trabajo consiste en reconstruir cómo se componen (Baumer et al., 2024; Grusin, 2015).

Por eso me interesa que la pregunta ética no sea simplemente "¿IA sí o no?". Hay que producir criterios capaces de intervenir sobre configuraciones.

¿Cuándo conviene introducir fricción en vez de validación?

¿Cuándo una interfaz debería ralentizar en lugar de acelerar?

¿Qué operaciones deberían permanecer reconstruibles por quien aprende?

¿Qué memorias conviene conservar y cuáles deberían poder olvidarse?

¿Qué formas de personalización aumentan pertinencia sin clausurar la aparición de lo inesperado?

¿Qué lugar debe tener el silencio en una infraestructura construida para responder?

¿Qué diseños ayudan a contrastar perspectivas en lugar de reforzar una única lectura?

¿Qué capacidades queremos que permanezcan disponibles en la persona incluso cuando la infraestructura deja de estar?

¿Qué vínculos humanos se fortalecen y cuáles se desplazan cuando una parte de la elaboración cotidiana ocurre con un interlocutor artificial?

Estas preguntas no constituyen un programa de prohibición. Tampoco un catecismo de "uso responsable". Apuntan a algo más difícil: **mejorar las condiciones psico-tecno-sociales en las que podamos elaborar, decidir, aprender, vincularnos y producir proyectos sin reducir la subjetividad a una secuencia perfectamente optimizable**.

La hipótesis transductiva obliga además a abandonar la fantasía de una intervención exterior. Si usamos IA para investigar IA, si escribimos con herramientas de asistencia, si nuestros estudiantes ya incorporan estos sistemas, si las instituciones empiezan a organizarlos en sus procesos, nosotros también estamos dentro del campo que intentamos comprender. La tarea no es recuperar una pureza anterior. Es construir reflexividad suficiente para volver visibles algunas de las operaciones que nos constituyen y ampliar los márgenes de intervención sobre ellas.

Aquí la investigación empírica se vuelve indispensable. Necesitamos historiales elegidos por participantes, entrevistas, diarios de uso, escenas, experimentos, estudios longitudinales y cartografías que permitan observar qué ocurre antes, durante y después de una interlocución. Necesitamos comparar sistemas, situaciones, edades, contextos institucionales y formas de uso. Necesitamos

producir datos propios y, al mismo tiempo, conservar la densidad cualitativa suficiente para que una regularidad estadística no sustituya la experiencia que pretende describir.

La pregunta final, entonces, no es si las máquinas llegarán a hablar “como nosotros”. Ya producen lenguaje con una competencia suficiente para entrar en nuestras prácticas. Tampoco es si poseen un inconsciente que debamos descubrir. Tal vez esa pregunta nos distraiga.

La pregunta que quiero dejar abierta es otra:

> **¿qué clase de sujetos pueden llegar a constituirse cuando una parte creciente de aquello que todavía no sabemos de nosotros mismos empieza a ser formulado junto con máquinas entrenadas para responder?**

No porque tengamos todavía una teoría cerrada.

Sino porque tenemos un problema.

Y porque ese problema ya está ocurriendo.

Referencias

Agüera y Arcas, B. (2025). *What is intelligence? Lessons from AI about evolution, computing, and minds*. The MIT Press.

Ammari, T., Chen, M., Zaman, S. M. M., & Garimella, K. (2026). *Learning to live with AI: How students develop AI literacy through naturalistic ChatGPT interaction* [Preprint]. arXiv. https://arxiv.org/abs/2601.20749

Amodei, D. (2026, enero). *The adolescence of technology: Confronting and overcoming the risks of powerful AI*. https://darioamodei.com/essay/the-adolescence-of-technology

Anthropic. (2026). *Claude's constitution*. https://www.anthropic.com/constitution

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences, 122*(26), e2422633122. https://doi.org/10.1073/pnas.2422633122

Baumer, E. P. S., Taylor, A. S., Brubaker, J. R., & McGee, M. (2024). Algorithmic subjectivities. *ACM Transactions on Computer-Human Interaction, 31*(3), Article 35. https://doi.org/10.1145/3660344

Binz, M., Akata, E., Bethge, M., Brändle, F., Callaway, F., Coda-Forno, J., Dayan, P., Demircan, C., Eckstein, M. K., Éltető, N., Griffiths, T. L., Haridi, S., Jagadish, A. K., Ji-An, L., Kipnis, A., Kumar, S., Ludwig, T., Mathony, M., Mattar, M., ... Schulz, E. (2025). A foundation model to predict and capture human cognition. *Nature, 644*, 1002-1009. https://doi.org/10.1038/s41586-025-09215-4

Campbell, M., Hoane, A. J., Jr., & Hsu, F.-h. (2002). Deep Blue. *Artificial Intelligence, 134*(1-2), 57-83. https://doi.org/10.1016/S0004-3702(01)00129-1

Chatonsky, G., & Citton, Y. (2026). Contra-historia de la aceleración (E. Inchauspe, Trad.). Arqueologías del Porvenir. https://arqueologiasdelporvenir.com.ar/traducciones/contra-historia-de-la-aceleracion/

Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). *How people use ChatGPT* (NBER Working Paper No. 34255). National Bureau of Economic Research. https://doi.org/10.3386/w34255

Cheng, M., Yu, S., Lee, C., Khadpe, P., Ibrahim, L., & Jurafsky, D. (2026a). ELEPHANT: Measuring and understanding social sycophancy in LLMs. *International Conference on Learning Representations (ICLR 2026)*. https://proceedings.iclr.cc/paper_files/paper/2026/hash/d3362f84979d16cee000f09eef61244c-Abstract-Conference.html

Cheng, M., Lee, C., Khadpe, P., Yu, S., Han, D., & Jurafsky, D. (2026b). Sycophantic AI decreases prosocial intentions and promotes dependence. *Science, 391*(6792), eaec8352. https://doi.org/10.1126/science.aec8352

Costa, F. (2026). *Vidas artificiales: Del accidente pandémico a las guerras de la IA*. Taurus.

Grusin, R. (2015). Radical mediation. *Critical Inquiry, 42*(1), 124-148. https://doi.org/10.1086/682998

Hansen, M. B. N. (2015). *Feed-forward: On the future of twenty-first-century media*. University of Chicago Press.

Heredia, J. M., & Rodríguez, P. E. (2025). Simondon and transindividual individuation: Technicity and affectivity in digital networks. In A. Bardin, M. Ferrari, A. Nony, & G. Tenti (Eds.), *The Edinburgh companion to Gilbert Simondon* (pp. 67-81). Edinburgh University Press.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation, 18*(7), 1527-1554. https://doi.org/10.1162/neco.2006.18.7.1527

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems, 25*.

Kvietkute, D., & Ness, I. J. (2026). Encountering generative AI: Narrative self-formation and technologies of the self among young adults. *Societies, 16*(1), 26. https://doi.org/10.3390/soc16010026

Lee, H.-P. H., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a

survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Article 1121, pp. 1-22). Association for Computing Machinery. https://doi.org/10.1145/3706598.3713778

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence*. Dartmouth College.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems, 35*, 27730-27744.

Pei, G., Li, H., Lu, Y., Wang, Y., Hua, S., & Li, T. (2024). Affective computing: Recent advances, challenges, and future trends. *Intelligent Computing, 3*, Article 0076. https://doi.org/10.34133/icomputing.0076

Rantala, J., & Muilu, M. (2024). Simondon, control and the digital domain. *Theory, Culture & Society, 41*(4), 23-40. https://doi.org/10.1177/02632764231201337

Rikap, C. (2026). *Teoría de la dependencia digital: Soberanía y desarrollo en el capitalismo del siglo XXI*. Caja Negra.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65*(6), 386-408. https://doi.org/10.1037/h0042519

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*, 484-489. https://doi.org/10.1038/nature16961

Simondon, G. (2007). *El modo de existencia de los objetos técnicos* (M. Martínez, Trad.). Prometeo Libros. (Obra original publicada en 1958)

Simondon, G. (2009). *La individuación a la luz de las nociones de forma y de información*. La Cebra/Cactus.

Stiegler, B. (2002). *La técnica y el tiempo 1: El pecado de Epimeteo* (B. Morales Bastos, Trad.). Hiru. (Obra original publicada en 1994)

Sula, C., Dewan, U., & McDonald, N. (2026). Rehearsing the self with AI: Teens cross-domain use of chatbots. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (Article 1494, pp. 1-14). Association for Computing Machinery. https://doi.org/10.1145/3772318.3793144

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*(236), 433-460. https://doi.org/10.1093/mind/LIX.236.433

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*.

Weizenbaum, J. (1966). ELIZA-a computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36-45. https://doi.org/10.1145/365153.365168

Zahavy, T. (2026). *Position: LLMs can't jump* [Preprint]. PhilSci-Archive. https://philsci-archive.pitt.edu/28024/