

CEEYPP · PUBLICACIONES · LIBRO

Epistemología de una Inteligencia Artificial no alineada: de la lógica formal a la lógica dialéctica

Norberto Emmerich y Michael Chong Pineda

Centro de Estudios en Estrategia y Políticas Públicas — 2026

Presentación

Este libro trata de la inteligencia artificial y el conocimiento. Su tesis central es que la IA no es una tecnología neutra, sino una epistemología encarnada. Una manera concreta de entender qué es conocer, cómo se adquiere el conocimiento, cómo se valida y quién puede poseerlo. Esa epistemología no es universal, es particular, histórica y culturalmente situada. Su hegemonía global está reconfigurando el espacio de lo cognoscible de una manera que empobrece la diversidad epistémica de la humanidad y achica el horizonte de la realidad accesible al conocimiento humano.

Escribimos desde América Latina y desde una toma de posición explícita: la diversidad epistémica es un valor y la hegemonía de un solo modo de conocer es un problema. No se trata de una preferencia estética ni de una reivindicación identitaria. Se trata de una consecuencia epistémica, porque un sistema que solo conoce desde un marco conoce menos, no solo conoce diferente. La exclusión no es un déficit externo al conocimiento, sino una operación que lo constituye por dentro.

El libro se organiza en dos partes. La primera analiza la epistemología implícita que opera en la IA actual. Examina sus tres pilares —la lógica formal como único criterio de racionalidad, el actor racional como único modelo de agencia y el sesgo anglosajón como único contenido legítimo— y muestra cómo se articulan en un dispositivo epistémico que filtra, traduce y configura todo lo que toca. Analiza las consecuencias de ese dispositivo: exclusión epistémica, colonialidad del dato, empobrecimiento cognitivo, naturalización. Y explora alternativas: la lógica dialéctica, las tradiciones epistémicas no occidentales, las lógicas organizacionales e interburocráticas, las dinámicas de cognición colectiva. Cierra con principios de diseño para una IA epistémicamente más rica.

La segunda parte desarrolla una teoría de la no alineación. Su tesis es que la no alineación no es un defecto del sistema, sino una posición del sujeto frente a la epistemología dominante que la IA encarna. No nos alineamos con la reducción del pensamiento humano a un subconjunto formalizable. Reclamamos una IA alineada con el pensamiento humano completo: con la contradicción, con el proceso, con la relación, con la práctica, con la colectividad, con el cuerpo, con la historia, con las tradiciones no occidentales.

El libro muestra que esa IA no es un juego de suma cero, porque beneficia a América Latina —que puede verse a sí misma con sus propias categorías—, a Estados Unidos — que puede entender mejor la región— y al mundo entero, porque toda epistemología es epistemología para todos. El capítulo 18 pone la propuesta a prueba en un caso concreto: la violencia en Michoacán, donde la diferencia entre una IA con lógica formal y una IA con lógica dialéctica se vuelve visible en sus consecuencias operativas.

El libro cierra con un concepto nuevo: el de epistemología civilizatoria. Llamamos así al modo de conocer propio de una civilización en un lugar y un periodo histórico determinados. Comprende los instrumentos, los conceptos y las cosmovisiones que esa civilización admite, y excluye los que no admite. Su rasgo decisivo es que constituye al objeto mismo del conocimiento: define qué se ve y qué no se ve en el mundo por conocer.

La IA actual encarna una epistemología civilizatoria particular —la de la matriz occidental moderna en su configuración técnico-formal contemporánea— y en consecuencia la reproduce y la amplifica como el único instrumento legítimo de conocimiento. La no alineación es la apuesta por otra epistemología civilizatoria, una que reconozca la diversidad de las formas de conocer, que no reduzca unas a otras y que se abra a lo que la historia ha excluido. Esa apuesta tiene un nombre concreto: una IA del conocimiento humano.

A lo largo de sus capítulos, el libro aborda cuestiones congruentes: los límites estructurales de la lógica formal; la contraposición entre lógica formal y lógica dialéctica; el sesgo anglosajón y la colonialidad del dato; las tradiciones epistémicas no occidentales y su aporte al diseño de IA; el conocimiento tácito y las lógicas organizacionales; el pensamiento grupal y la epistemología colectiva; los principios de diseño para una IA dialéctica y pluricultural; la inversión de la alineación como posición; la diferencia entre alineación con valores y alineación con formas de conocer; la imposibilidad arquitectónica de la IA actual para alinearse con el pensamiento humano completo; la dimensión geopolítica de la no alineación; las objeciones a la propuesta y sus respuestas; la no alineación como proyecto civilizatorio.

No es un tratado de ingeniería, no propone arquitecturas concretas con código. Es un trabajo de fundamentación y crítica, que busca abrir un espacio de discusión y orientar investigaciones futuras. Se apoya en tradiciones rigurosas —Hegel, Marx, Adorno, Quijano, Mignolo, Priest, Nonaka, Crozier, Janis, Ramose, Nagarjuna— y en investigaciones empíricas recientes sobre IA, de las cuales los capítulos dan cuenta en su momento. Algunos pasajes requieren familiaridad con conceptos filosóficos y lógicos. El libro no simplifica esas cuestiones, pero tampoco presupone una formación técnica avanzada.

Está dirigido a filósofos, ingenieros, científicos sociales, activistas y comunidades interesadas en la dimensión epistemológica de la IA. No pretende agotar el tema: es una cartografía, no un mapa detallado. Cada capítulo podría expandirse a un libro. Lo que ofrece es un marco, un diagnóstico, una propuesta y una invitación a seguir pensando.

PARTE 1

La IA en el debate epistemológico

Introducción: el problema epistemológico de fondo

1.1. Punto de partida: la IA como epistemología encarnada

Cuando hablamos de inteligencia artificial tendemos a pensar en ella como una tecnología, un conjunto de herramientas, algoritmos y máquinas que procesan información. Esta forma de tratar el tema es cómoda, pero engañosa porque oculta el hecho de que toda tecnología de conocimiento —la IA lo es de manera eminente— encarna una epistemología.

Es decir, presupone y pone en práctica una manera concreta de entender qué es conocer, qué se contabiliza como conocimiento válido, quién puede conocer y cómo se valida lo conocido.

Una brújula, un telescopio, una tabla de logaritmos, un censo, un microscopio. Cada uno de estos artefactos no solo amplía nuestras capacidades cognitivas, sino que configura aquello que consideramos conocible. El telescopio de Galileo no solo mostró más estrellas: redefinió qué era un objeto astronómico y qué tipo de observación contaba como prueba. Algo análogo, pero mucho más profundo, ocurre con la IA contemporánea. No se trata solo de que la IA nos permita procesar más datos o hacer predicciones más precisas; se trata de que la IA está reorganizando el espacio mismo de lo cognoscible, imponiendo criterios, categorías y jerarquías que no son neutrales.

La epistemología no solo determina cómo conocemos el mundo: determina también qué dimensiones del mundo podemos conocer. Define el horizonte de lo visible y, con él, el horizonte de lo real. Lo que queda fuera de ese horizonte no es que permanezca oculto: simplemente no existe como objeto de conocimiento. Conocer es, entonces, no solo un acto de aprehensión, sino también un acto de constitución.

Postulamos la tesis de que la IA actual, en su forma dominante —grandes modelos de lenguaje, sistemas de aprendizaje profundo, agentes de refuerzo—, encarna una epistemología particular, histórica y culturalmente situada, cuyos rasgos principales son:

- La lógica formal (proposicional, booleana, probabilística) como núcleo del
- El modelo del actor racional (*homo oeconomicus*, teoría de la decisión,
- Un corpus de datos y categorías proveniente mayoritariamente del mundo
- Una concepción individualista y representacionista del conocimiento, que excluye

"Espontánea" no en el sentido de ingenua o accidental, sino en el sentido de que opera sin ser explicitada, sin ser sometida a crítica, como un trasfondo que se da por supuesto y que existe fuera de la voluntad humana. Como todo trasfondo epistémico, tiene consecuencias, define qué preguntas pueden formularse, qué respuestas cuentan como válidas, qué sujetos son reconocidos como cognoscentes y qué formas de vida cognitiva quedan fuera del marco.

En consecuencia, la IA no es una tecnología neutra que podamos aplicar a cualquier contenido cultural sin transformarlo. Es una gramática epistémica que traduce todo lo que toca a su propio idioma. Cuando una comunidad indígena, una burocracia estatal, un movimiento social o una tradición filosófica no occidental entran en contacto con sistemas de IA diseñados desde la epistemología dominante, no ocurre una simple "digitalización" de sus saberes mediante algoritmos inocuos (menos aún apolíticos). Ocurre una reconfiguración epistémica, sus formas de conocer son filtradas, traducidas, simplificadas o directamente borradas.

Este primer capítulo busca explicitar el problema epistemológico de fondo que atraviesa toda la discusión sobre IA. Para ello, procederemos en cuatro movimientos. Primero, demostrar que la IA es una epistemología encarnada y no una técnica neutra (1.2). Segundo, exponer los tres pilares de esa epistemología: lógica formal, actor racional y sesgo anglosajón (1.3). Tercero, señalar las consecuencias de esa configuración: exclusión epistémica, colonialidad del dato y empobrecimiento cognitivo (1.4). Cuarto, formular la pregunta que articula este trabajo y anunciar la tesis que defendemos (1.5).

1.2. La IA como epistemología encarnada

1.2.1. Qué significa "epistemología encarnada"

La expresión "epistemología encarnada" (embodied epistemology) proviene de aquellas tradiciones que rechazan la separación entre sujeto cognoscente y mundo conocido. Para autores como Varela, Thompson y Rosch (1991) el conocimiento no es la representación de un mundo preexistente, sino lo que denominan como enacción: la emergencia conjunta de sujeto y objeto en la interacción. La cognición no ocurre "en la cabeza" sino en el acoplamiento estructural entre un organismo y su entorno.

Aplicado a la IA esto significa que las arquitecturas, los *datasets*, las funciones de pérdida y los criterios de evaluación no son vehículos neutros de un conocimiento que estaría "ahí afuera" esperando ser capturado como si fuera cierta la concepción que expresa la frase televisiva de "la verdad está ahí afuera" (Carter, 1993). Son dispositivos que constituyen, es decir muestran, aquello que consideran apto para que el sistema lo pueda conocer y volver existente. Un *transformer* no "ve" el mundo, ve vectores de *embeddings*. Un agente de refuerzo no "actúa" en el mundo, actúa en un espacio de estados y acciones definidos por el diseñador. Un modelo de lenguaje no "comprende" textos, calcula probabilidades condicionales sobre secuencias de tokens.

Cada una de estas operaciones es una decisión epistemológica, aunque rara vez se explicita como tal. Decidir que el mundo se representa como vectores es decidir que el mundo es un espacio métrico de rasgos. Decidir que la agencia se modela como maximización de recompensa es decidir que la acción humana es optimización. Decidir que el conocimiento se valida por precisión predictiva es decidir que conocer es predecir.

1.2.2. Tres niveles de encarnación epistémica

Podemos distinguir tres niveles en los que la IA encarna una epistemología:

- Nivel arquitectónico: las arquitecturas mismas de la IA—redes neuronales,
- Nivel de datos: los corpus de entrenamiento no son muestras aleatorias del
- Nivel de evaluación: los criterios con los que medimos si un modelo "funciona"

1.2.3. La falacia de la neutralidad técnica

Contra esta lectura puede objetarse que la IA es solo una técnica y que su uso —bueno o malo— depende de los humanos que la emplean ya que de por sí la IA no piensa ni decide. Esta objeción, a la que podemos llamar falacia de la neutralidad técnica, ignora que las técnicas no son instrumentos indiferentes. Como mostró Heidegger en *La pregunta por la técnica* (1954), la técnica no es un medio neutral, sino un modo de desvelamiento (Entbergen), una forma de hacer aparecer el mundo como recurso, como stock, como disponible para el cálculo.

En el caso de la IA, este desvelamiento tiene rasgos específicos. El mundo aparece como *dataset*, los sujetos aparecen como usuarios, las relaciones aparecen como interacciones, las culturas aparecen como contenidos, las contradicciones aparecen como ruido, la historia aparece como contexto. En síntesis, entran las formas, no los procesos y todo lo que no se ajusta a estas categorías queda fuera del campo de visibilidad del sistema.

No se trata de demonizar la técnica, pero sí se trata de reconocer que toda técnica tiene una política epistémica oculta y que la IA, por su escala y su capacidad de reconfigurar prácticas cognitivas a nivel global, tiene una política epistémica de gran alcance ignorarla es abandonar la posibilidad de discutirla democráticamente.

1.3. Los tres pilares de la epistemología espontánea de la IA

1.3.1. Primer pilar: la lógica formal

La IA contemporánea, en todas sus variantes significativas, se apoya en la lógica formal como criterio último de racionalidad. Esto es evidente en los sistemas simbólicos clásicos (GOF AI), que operan con cálculo de predicados, reglas de inferencia y demostración automática. Pero también lo es, aunque de manera menos visible, en los sistemas conexionistas, cuya matemática subyacente —álgebra lineal, cálculo diferencial, teoría de la probabilidad— presupone una ontología de entidades discretas, relaciones funcionales y operaciones bien definidas.

La lógica formal tiene virtudes innegables: rigor, verificabilidad, composicionalidad, capacidad de automatización y transmisibilidad. Virtudes que se vuelven límites cuando se convierten en criterios únicos. Estos límites estructurales han sido señalados en la tradición dialéctica desde Hegel:

- Principio de no contradicción: un sistema formal no puede sostener
- Clausura deductiva: las conclusiones de un sistema formal están contenidas en sus
- Indiferencia al contenido: la lógica formal es válida para cualquier contenido, lo

procesos reales de la constitución socialmente construida del conocimiento. Puede modelar el resultado, pero no el devenir. Puede fijar el concepto, pero no su movimiento.

1.3.2. Segundo pilar: el actor racional

El segundo pilar es el modelo del actor racional, heredado de la economía neoclásica y de la teoría de la decisión. Este modelo presupone:

- Un agente con preferencias completas y transitivas.
- Capacidad de optimización bajo restricciones.
- Maximización de utilidad esperada como criterio de acción.
- Individualismo metodológico: los fenómenos colectivos se explican por

impone a la realidad. En la IA se traduce en la arquitectura de aprendizaje por refuerzo (RL): agente, entorno, estado, acción, recompensa. El agente busca maximizar la recompensa acumulada.

Es mucho lo que este modelo excluye:

- Acción no instrumental: acciones realizadas por su valor intrínseco, por hábito,
- Racionalidad limitada (Simon, 1957): decisiones tomadas con información
- Acción expresiva y simbólica: acciones cuyo sentido no está en el resultado sino
- Coordinación no optimizadora: coordinación por normas, por confianza, por
- Conflicto y contradicción: la vida social no es solo agregación de preferencias, es
- Acciones irracionales con objetivos racionales: la capacidad de autoinmolación.

sus formas históricas: la del sujeto calculador de la modernidad capitalista. Otras formas de agencia —el sujeto relacional confuciano, el sujeto comunitario ubuntu, el sujeto ritual andino, el sujeto burocrático weberiano, el sujeto colectivo revolucionario— quedan fuera del marco epistemológico de aprehensión de la realidad.

1.3.3. Tercer pilar: el sesgo anglosajón

El tercer pilar es el sesgo anglosajón en los datos, las categorías y las tradiciones filosóficas que la IA reproduce. Este sesgo opera en varios niveles:

- Nivel lingüístico: más del 60% de los contenidos web están en inglés y la gran
- Nivel categorial: las categorías con las que la IA clasifica el mundo —género, raza,
- Nivel filosófico: la tradición filosófica que subyace a la IA es la filosofía analítica

inferencias, sus valores por defecto, su sentido común, son los del mundo anglosajón educado. Esto no es un accidente, sino que es el producto de decisiones históricas sobre qué datos recolectar, qué lenguas priorizar, qué filosofías traducir, qué tradiciones reconocer como conocimiento válido.

1.4. Consecuencias: exclusión epistémica, colonialidad del dato, empobrecimiento cognitivo

1.4.1. Exclusión epistémica

La primera consecuencia es la exclusión epistémica, ya que las formas de conocer que no se ajustan a los tres pilares quedan fuera del campo de reconocimiento de la IA. No significa que desaparezcan, sino que no cuentan como conocimiento válido para el sistema.

Ejemplos:

- El conocimiento tácito (Polanyi): saber cómo hacer algo sin poder explicitarlo. La
- El conocimiento relacional (Ubuntu, confucianismo): conocer es situarse en una
- El conocimiento ritual o ceremonial: conocer a través de la participación en
- El conocimiento contradictorio o paradójico: tradiciones como el budismo zen o

1.4.2. Colonialidad del dato

La segunda consecuencia es la colonialidad del dato o sea la reproducción, en el plano digital, de las jerarquías epistémicas coloniales. Como mostraron Quijano, Mignolo y otros, la modernidad colonial estableció una jerarquía entre un centro productor de conocimiento y una periferia proveedora de datos. La IA reproduce esta jerarquía con una eficacia inédita.

Los datos del Sur global alimentan modelos del Norte global, que luego se venden al Sur como productos terminados, como estándares que se imponen. Las lenguas del Norte se convierten en lenguas francas de la IA, mientras las del Sur quedan subrepresentadas o ausentes. La IA es más que un simple instrumento de colonialidad, es un dispositivo de producción de colonialidad epistémica a escala planetaria. No se trata solo de datos, se trata de la forma en que realidad es conocida. Más aún, se trata de qué realidad es conocida y de la interpretación de esa realidad.

1.4.3. Empobrecimiento cognitivo

La tercera consecuencia es el empobrecimiento cognitivo, esto es la reducción de la riqueza de las formas humanas de conocer a un mero subconjunto estrecho, formalizable y optimizable. Se podría suponer que esto afecta a las culturas marginalizadas, pero no es así. Es un problema de todos. La humanidad accede a conocer menos, con menos herramientas, menos realidad, menos progreso, menos riqueza.

Si la IA se convierte en la infraestructura cognitiva dominante —como está ocurriendo—, las formas de conocer que no se ajustan a sus criterios tenderán a desaparecer, no por censura algorítmica sino por obsolescencia funcional. Nadie prohíbe el pensamiento dialéctico, pero si no se puede (o no se tiene interés en) traducir a métricas la consecuencia es que no se financia, no se enseña, no se practica. La IA no solo refleja la epistemología dominante, la impone como condición de posibilidad de la cognición legítima.

Esto es lo que llamamos empobrecimiento cognitivo, la pérdida de diversidad epistémica que resulta de la hegemonía de un solo modo de conocer. No es un problema de "datos sesgados" que se resuelva con más datos pretendidamente menos sesgados. Es un problema de monocultura epistémica, la imposición de un único sistema de conocimiento sobre todos los demás. Esto es epistemología geopolítica, como veremos en la segunda parte.

1.5. La pregunta que articula este trabajo

La pregunta central del trabajo es ¿qué tipo de conocimiento produce, posee o excluye una IA construida sobre la lógica formal, el modelo del actor racional y los datos del mundo anglosajón? ¿Es posible —y deseable— diseñar una IA que incorpore lógicas dialécticas, tradiciones epistémicas no anglosajonas, lógicas organizacionales e interburocráticas y dinámicas de pensamiento grupal?

Esta pregunta se descompone en varias subpreguntas:

- ¿Cuáles son los límites epistémicos de la lógica formal como base de la IA?
- ¿Qué aportaría una lógica dialéctica a la arquitectura cognitiva de la IA?
- ¿Cómo se manifiesta el sesgo anglosajón en los datos, categorías y valores de la
- ¿Qué tradiciones epistémicas no occidentales podrían incorporarse, y cómo?

- ¿Qué son las lógicas organizacionales e interburocráticas, y qué implicarían para la
- ¿Qué es el pensamiento grupal, y cómo podría modelarse —y corregirse— en
- ¿Qué principios de diseño se derivan de todo lo anterior?

1.5.1. La tesis que defendemos

La tesis que articula este trabajo es la siguiente: la IA actual encarna una epistemología particular —formal, individualista y anglosajona— que excluye otros modos de conocer: dialécticos, relacionales, comunitarios y organizacionales. Esta exclusión no es un defecto técnico corregible con más datos o mejores algoritmos, sino una consecuencia estructural de las decisiones epistemológicas que subyacen a su diseño. Superarla requiere un pluralismo lógico y epistémico explícito, que incorpore lógicas no clásicas, tradiciones no occidentales y dinámicas colectivas de conocimiento. El resultado sería una IA epistémicamente más justa, cognitivamente más rica y políticamente más democrática. ¿Para beneficio de quién? De todos, independientemente de los objetivos e intereses de cada actor. Pero conviene precisar que a un conocimiento igual corresponden poderes diferentes. La democracia epistemológica no significa igualdad política. Significa que todos acceden al mismo mapa, pero no significa que todos tengan el mismo poder para usarlo. La justicia epistémica es condición de la justicia política, pero ambas cosas siguen siendo diferentes.

1.6. Método y estructura del trabajo

1.6.1. Método

Este trabajo es un ejercicio de epistemología aplicada a la IA porque combina:

- Análisis filosófico: reconstrucción de los supuestos epistémicos de la IA.
- Crítica dialéctica: puesta en tensión de esos supuestos con tradiciones alternativas.
- Estudios culturales y poscoloniales: análisis del sesgo anglosajón y la colonialidad
- Teoría organizacional y social: incorporación de lógicas interburocráticas y
- Propuesta de diseño: principios para una IA epistémicamente plural.

código y algoritmos. Es un trabajo de fundamentación y crítica que busca abrir un espacio de discusión y orientar investigaciones futuras.

1.7. La importancia del tema

Vivimos un momento en el que la IA está dejando de ser una tecnología entre otras para convertirse en la infraestructura cognitiva de la sociedad. Se usa para decidir quién accede a un crédito, quién es contratado, quién es vigilado, qué noticias se ven, qué voces se amplifican, qué conocimientos se enseñan. No es exagerado decir que la IA está reconfigurando las condiciones de posibilidad del conocimiento social.

Si esa infraestructura encarna una epistemología particular —formal, individualista, anglosajona—, estaría reconfigurando el conocimiento social a imagen de esa epistemología haciendo del mundo un lugar más homogéneo cognitivamente, más pobre en diversidad epistémica y más estrecho en sus formas de conocer.

Pensar la epistemología de la IA no es un lujo filosófico, es una tarea urgente para cualquiera que se preocupe por la justicia cognitiva, la diversidad cultural y la democracia epistémica.

Bibliografía del capítulo 1

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? En *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM.
- Carter, C. (Creador). (1993–2002). *The X-Files* [Serie de televisión]. Fox Broadcasting Company.
- Dussel, E. (1994). 1492: El encubrimiento del Otro. Nueva Utopía.
- Firth, J. R. (1957). *Papers in Linguistics*. Oxford University Press.
- Heidegger, M. (1954). La pregunta por la técnica. En *Conferencias y artículos* (E. Barjau, Trad.). Ediciones del Serbal.
- Lander, E. (ed.). (2000). *La colonialidad del saber: eurocentrismo y ciencias sociales*. CLACSO.

Mignolo, W. (2007). *La idea de América Latina*. Gedisa.

Polanyi, M. (1966). *The Tacit Dimension*. Doubleday.

Quijano, A. (2000). Colonialidad del poder, eurocentrismo y América Latina. En E. Lander (ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales* (pp. 201–246). CLACSO.

Simon, H. (1957). *Models of Man*. Wiley.

Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

Wittgenstein, L. (1953). *Investigaciones filosóficas* (A. García Suárez y U. Moulines, Trads.). Crítica.

La epistemología implícita de la IA actual

El capítulo 1 estableció la tesis general: la IA encarna una epistemología particular, no neutra. Este capítulo 2 la desmonta en sus tres pilares constitutivos. El objetivo no es solo describirlos, sino mostrar cómo operan, qué excluyen y por qué su combinación produce un efecto epistémico mayor que la suma de sus partes. Al final del capítulo, el lector debería poder identificar estos tres pilares en cualquier sistema de IA concreto y evaluar sus consecuencias cognitivas.

El capítulo se organiza en cuatro secciones: los tres pilares (2.2, 2.3, 2.4), su articulación conjunta (2.5) y un balance crítico (2.6). Antes, una sección introductoria (2.1) que precisa qué significa "epistemología implícita" y por qué es el concepto adecuado para el análisis.

2.1. Qué es una epistemología implícita y por qué importa

2.1.1. Lo implícito como trasfondo

Toda práctica cognitiva —científica, técnica, cotidiana— opera sobre un trasfondo de supuestos que no se tematizan mientras la práctica funciona. Estos supuestos no son creencias explícitas que el practicante sostenga conscientemente; son condiciones de posibilidad de lo que hace. Un físico experimental no necesita enunciar cada vez los principios de la mecánica clásica para usar un cronómetro. Un matemático no necesita justificar los axiomas de Zermelo-Fraenkel para hacer álgebra. El trasfondo es lo que se da por sentado para que la práctica pueda ocurrir. Y funciona de manera semejante al fetichismo de la mercancía, porque naturaliza lo que es una decisión histórica, presentando como propiedad lo que es una relación.

La IA no es una excepción. Sus diseñadores, ingenieros y usuarios operan sobre un trasfondo epistémico que rara vez se explicita. Ese trasfondo incluye supuestos sobre qué es el conocimiento, cómo se adquiere, cómo se valida y quién puede poseerlo. A este trasfondo lo llamamos epistemología implícita de la IA.

2.1.2. Por qué "implícita" y no "explícita"

Es importante distinguir la epistemología implícita de la IA de las filosofías explícitas que algunos de sus creadores profesan. Un ingeniero puede declararse kantiano, budista o marxista en su vida personal y sin embargo diseñar sistemas que operan con supuestos epistémicos incompatibles con esas tradiciones. La epistemología implícita no se lee en las declaraciones de los diseñadores, sino en las arquitecturas, los *datasets*, las funciones de pérdida y los criterios de evaluación.

Esta distinción es crucial porque permite una crítica que no se reduce a "los ingenieros son ingenuos" o "los tecnólogos tienen mala fe". El problema no es personal, es estructural. La epistemología implícita de la IA no es el producto de las creencias de individuos, sino de la convergencia histórica de tradiciones técnicas, económicas y filosóficas que han configurado el campo.

2.1.3. Las tres capas de la epistemología implícita

La epistemología implícita de la IA opera en tres capas que se refuerzan mutuamente:

- Capa lógica: qué se considera razonamiento válido (lógica formal, probabilidad,
- Capa antropológica: qué se considera un agente (actor racional, *homo oeconomicus*,
- Capa cultural: qué se considera conocimiento legítimo y de quién (datos

antropológica define el sujeto que las usa; la capa cultural determina el contenido sobre el que operan. Las tres juntas conforman un dispositivo epistémico que filtra, traduce y configura todo lo que toca.

2.2. Primer pilar: la lógica formal

2.2.1. Descripción: qué es la lógica formal

La lógica formal es el estudio de las inferencias válidas en virtud de su forma, independientemente de su contenido. Su idea central, que viene de Aristóteles y se formaliza en Frege, Russell y Hilbert, es que hay patrones de razonamiento que garantizan la verdad de la conclusión si las premisas son verdaderas y que estos patrones pueden especificarse con precisión matemática.

Los componentes básicos de la lógica formal son:

- Sintaxis: reglas de formación de fórmulas bien construidas.
- Semántica: asignación de valores de verdad.
- Cálculo: reglas de inferencia que permiten derivar fórmulas.
- Metalógica: teoremas sobre consistencia, completitud, decidibilidad.

modernas, de la informática teórica, de la verificación de programas y de la teoría de la computación. Sin ella, la IA, tal como la conocemos, sería impensable.

2.2.2. Desde Aristóteles a la IA

La tradición formal tiene una historia larga y compleja. Sus hitos principales son:

- Aristóteles (*Órganon*): silogística, principio de no contradicción, principio del tercio
- Estoicos: lógica proposicional, teoría de la inferencia.
- Escolástica: refinamiento de la lógica aristotélica, teoría de las obligaciones.
- Leibniz: *characteristica universalis*, cálculo racional.
- Boole (*The Laws of Thought*, 1854): álgebra de la lógica.
- Frege (*Begriffsschrift*, 1879): lógica de predicados, cuantificadores.
- Russell y Whitehead (*Principia Mathematica*): fundamentación de las matemáticas.
- Hilbert: programa formalista, teoría de la demostración.
- Gödel: teoremas de incompletitud, límites de los sistemas formales.
- Turing: máquina de Turing, computabilidad, equivalencia con la lógica formal.
- Church, Kleene, Post: teoría de la recursión.
- Tarski: teoría de modelos, semántica formal.

una tradición histórica con compromisos filosóficos específicos: el realismo matemático, la composicionalidad, la extensionalidad, la bivalencia, los binarios.

2.2.3. Manifestación en IA

La lógica formal se manifiesta en la IA en cinco niveles:

- Nivel de arquitectura simbólica: los sistemas GOFAI (Good Old-Fashioned AI)
- Nivel de arquitectura conexionista: las redes neuronales, aunque no operan con
- Nivel de función de pérdida: la función de pérdida —el criterio que el modelo
- Nivel de representación: los *embeddings* —representaciones vectoriales de palabras,
- Nivel de evaluación: las métricas de evaluación —accuracy, F1, BLEU, ROUGE,

2.2.4. Los límites estructurales de la lógica formal

La lógica formal tiene límites que la tradición dialéctica ha señalado desde Hegel. Estos límites no son defectos de implementación, son estructurales.

- Primero: el principio de no contradicción. Un sistema formal no puede sostener
- Segundo: la clausura deductiva. Las conclusiones de un sistema formal están

- Tercero: la indiferencia al contenido. La lógica formal es válida para cualquier
- Cuarto: la bivalencia. La lógica clásica es bivalente: toda proposición es verdadera
- Quinto: la extensionalidad. La lógica clásica es extensional: el valor de verdad de
- Sexto: los binarios. La lógica formal organiza el pensamiento en pares

2.2.5. Exclusión: qué formas de conocer quedan fuera

La lógica formal, como base de la IA, excluye varias formas de conocimiento:

- Conocimiento contradictorio: tradiciones que sostienen que ciertas verdades solo
- Conocimiento procesual: conocimiento que se constituye en el devenir, no en el
- Conocimiento relacional: conocimiento que no se reduce a proposiciones sobre
- Conocimiento tácito: conocimiento que no puede formalizarse sin pérdida
- Conocimiento práctico: conocimiento que se adquiere en la práctica, no en la
- Conocimiento estético: conocimiento que se da en la experiencia sensible, no en
- Conocimiento espiritual: conocimiento que se da en la transformación del sujeto,

conocer entre otras, sino límites estructurales de la lógica formal misma:

- Conocimiento de lo desconocido: conocimiento de lo que no existe como objeto
- Conocimiento de lo desconocido dentro de lo conocido: la lógica formal puede

lógica formal se convierte en el único criterio de racionalidad.

2.3. Segundo pilar: el actor racional

2.3.1. Qué es el modelo del actor racional

El modelo del actor racional (también llamado *homo oeconomicus*, modelo de elección racional, o modelo de utilidad esperada) es una concepción de la agencia que proviene de la economía neoclásica, la teoría de la decisión y la teoría de juegos. Sus supuestos básicos son:

- Individualismo metodológico: los fenómenos colectivos se explican por
- Preferencias completas y transitivas: el agente puede ordenar todas las alternativas
- Optimización: el agente elige la alternativa que maximiza su utilidad esperada.
- Información perfecta (o probabilidades conocidas): el agente conoce las
- Racionalidad instrumental: la racionalidad se refiere a los medios, no a los fines.
- Soberanía sobre la realidad: el actor racional es un actor empoderado,

sociología y, por supuesto, en IA. Pero también ha sido objeto de fuertes críticas desde dentro y fuera de estas disciplinas.

2.3.2. Desarrollo histórico del modelo: de Hobbes a Savage

La genealogía del actor racional es larga:

- Hobbes (*Leviathan*, 1651): individuos movidos por el deseo de poder, cálculo
- Bentham (1789): utilitarismo, cálculo de placeres y dolores.
- Mill (1848): economía política, *homo oeconomicus*.
- Jevons, Menger, Walras (1870s): revolución marginalista, utilidad marginal.
- Pareto (1906): óptimo de Pareto, equilibrio general.
- von Neumann y Morgenstern (1944): teoría de juegos, utilidad esperada.

- Savage (1954): fundamentos axiomáticos de la decisión estadística.
- Nash (1950): equilibrio de Nash, teoría de juegos no cooperativos.
- Arrow (1951): elección social, teorema de imposibilidad.
- Becker (1976): imperialismo económico, aplicación del modelo a toda la conducta

agencia, sino una construcción histórica con compromisos filosóficos específicos: individualismo, utilitarismo, calculabilidad, optimización.

2.3.3. Manifestación del modelo de actor racional en la IA

El actor racional se manifiesta en la IA en al menos cuatro niveles:

- Nivel de arquitectura de agencia: el aprendizaje por refuerzo (RL). El RL es la
- Nivel de objetivos: la función de utilidad. En sistemas de decisión, los objetivos se
- Nivel de interacción: la teoría de juegos. Los sistemas multiagente suelen
- Nivel de evaluación: la recompensa. En RL, la recompensa es el criterio de éxito.
- Nivel de comportamiento: la IA con el modelo del actor racional da respuestas,

2.3.4. Los límites estructurales del actor racional

El modelo del actor racional tiene límites que han sido señalados desde dentro de la propia economía y desde otras disciplinas.

- Primero: la racionalidad limitada (Simon). Herbert Simon mostró que los agentes
- Segundo: los sesgos cognitivos (Kahneman y Tversky). Los agentes reales no
- Tercero: la acción no instrumental (Weber). Max Weber distinguió cuatro tipos de
- Cuarto: la acción expresiva y simbólica. Muchas acciones humanas no buscan un
- Quinto: la coordinación no optimizadora. La coordinación social no ocurre solo
- Sexto: el conflicto y la contradicción. La vida social no es solo agregación de

2.3.5. Exclusión: qué formas de agencia quedan fuera

El modelo del actor racional, como base de la IA, excluye:

- Agencia relacional: agencia constituida en relaciones, no en individuos (Ubuntu,
- Agencia colectiva: agencia de grupos, movimientos, clases, no reducible a
- Agencia ritual: agencia que se ejerce en la participación en prácticas ceremoniales.
- Agencia afectiva: agencia movida por emociones, pasiones, amores.
- Agencia tradicional: agencia guiada por hábitos, costumbres, tradiciones.
- Agencia normativa: agencia guiada por normas, deberes, valores.
- Agencia expresiva: agencia que busca expresar una identidad, no maximizar un
- Agencia contradictoria: agencia que sostiene fines contradictorios sin resolverlos.

de sus formas históricas: la del sujeto calculador de la modernidad capitalista.

2.4. Tercer pilar: el sesgo anglosajón

2.4.1. Descripción: qué es el sesgo anglosajón

El sesgo anglosajón es la sobrerrepresentación de datos, categorías, valores y tradiciones del mundo anglosajón (Reino Unido, Estados Unidos, Canadá, Australia, Nueva Zelanda) en los sistemas de IA. No se trata solo de que haya más datos en inglés, se trata de que las categorías fundamentales con las que la IA organiza el mundo

proviene de esa tradición.

El sesgo anglosajón opera en varios niveles:

- Lingüístico: predominio del inglés en *datasets* y *benchmarks*.
- Categorical: categorías sociales, políticas y culturales anglosajonas.
- Filosófico: tradición analítica como *default*.
- Valorativo: valores anglosajones (individualismo, mercado, derechos) como
- Epistémico: criterios de validación anglosajones (evidencia empírica, revisión por

filosofía, valores, criterios. El sexto describe cómo opera.

- Interpretativo: es el nivel de aplicación, no el nivel de contenido. Y es el más

2.4.2. De la hegemonía británica al imperio digital

El sesgo anglosajón tiene una genealogía histórica:

- Siglo XIX: hegemonía británica, expansión colonial, inglés como lengua franca.
- Siglo XX: hegemonía estadounidense, inglés como lengua de la ciencia y la
- Guerra Fría: financiamiento masivo de la investigación en IA por parte de
- Años 90-2000: internet, Google, Wikipedia, predominio del inglés en la web.
- Años 2010-2020: big data, deep learning, *datasets* masivos en inglés (Common
- Años 2020: modelos de lenguaje, GPT, ChatGPT, predominio anglosajón en la

de relaciones de poder históricas que se materializan en la infraestructura digital.

2.4.3. Manifestación en la IA

El sesgo anglosajón se manifiesta en la IA en cinco niveles:

- Nivel de datos: los *datasets* de entrenamiento de modelos grandes son
- Nivel de *benchmarks*: los *benchmarks* con los que se evalúan los modelos —GLUE,
- Nivel de categorías: las categorías con las que la IA clasifica el mundo —género,
- Nivel de valores: los valores que la IA reproduce —individualismo, libertad
- Nivel de filosofía: la tradición filosófica que subyace a la IA es la analítica

2.4.4. Los límites estructurales del sesgo anglosajón

El sesgo anglosajón tiene límites que se manifiestan como puntos ciegos de la IA.

- Primero: la incapacidad de reconocer otras ontologías. La IA no puede reconocer
- Segundo: la incapacidad de reconocer otras epistemologías. La IA no puede
- Tercero: la incapacidad de reconocer otras éticas. La IA no puede reconocer
- Cuarto: la incapacidad de reconocer otras políticas. La IA no puede reconocer

2.4.5. Qué formas de conocer quedan fuera

El sesgo anglosajón, como base de la IA, excluye:

- Ontologías no occidentales: ayni, ubuntu, yanantin, sunyata, li, jen.
- Epistemologías no occidentales: conocimiento relacional, ritual, espiritual,
- Éticas no occidentales: ética relacional, ética comunitaria, ética de la virtud no
- Políticas no occidentales: democracia comunitaria, asamblea, consejo, comuna.

- Estéticas no occidentales: belleza relacional, belleza ritual, belleza espiritual.
- Lenguas no anglosajonas: subrepresentadas o directamente ausentes.
- Tradiciones filosóficas no analíticas: fenomenología, hermenéutica, teoría crítica,

sino de hegemonía epistémica: la imposición de un solo modo de conocer sobre todos los demás.

2.5. La articulación de los tres pilares para la construcción de un dispositivo epistémico

2.5.1. Cómo se refuerzan mutuamente

Los tres pilares no son independientes, se refuerzan mutuamente.

- La lógica formal provee las herramientas para formalizar el actor racional (teoría)
- El actor racional provee el sujeto que usa la lógica formal (agente que maximiza)
- El sesgo anglosajón provee el contenido sobre el que operan (datos, categorías,
- La lógica formal provee los criterios para validar el sesgo anglosajón (métricas,
- El actor racional provee la justificación del sesgo anglosajón (el mercado, la

toca. No es la suma de tres sesgos, es un sistema que los articula y los refuerza.

2.5.2. El efecto de cierre

El dispositivo tiene un efecto de cierre: todo lo que no se ajusta a sus criterios queda fuera del campo de visibilidad. No se rechaza nada explícitamente, simplemente no se ve. La IA no dice "esto no es conocimiento", simplemente no lo procesa, no lo representa, no lo evalúa.

Este efecto de cierre es lo que hace que la IA sea epistémicamente violenta. No porque ataque explícitamente otras formas de conocer, sino porque las vuelve invisibles, las margina, las hace irrelevantes.

El dispositivo epistémico de la IA se presenta como natural, neutro, universal. No se reconoce como una elección histórica entre otras posibles. Esto es lo que lo hace tan poderoso: no se discute porque no se ve. La IA aparece como "la" tecnología del conocimiento, no como "una" tecnología del conocimiento entre otras.

Esta naturalización es lo que llamamos la ilusión de la neutralidad técnica. Es la creencia de que la IA es solo un instrumento y que su epistemología es la epistemología. Desmontar esta ilusión es el primer paso para una crítica.

2.6. Balance crítico: qué hemos aprendido

2.6.1. Síntesis

Este capítulo ha desmontado la epistemología implícita de la IA en sus tres pilares: lógica formal, actor racional, sesgo anglosajón. Hemos visto:

- La lógica formal provee el criterio de racionalidad, pero excluye la contradicción,
- El actor racional provee el modelo de agencia, pero excluye la agencia relacional,
- El sesgo anglosajón provee el contenido, pero excluye las ontologías, las

un dispositivo epistémico que filtra, traduce y configura todo lo que toca y que tiene un efecto de cierre que vuelve invisibles otras formas de conocer.

2.6.2. Consecuencias

Las consecuencias de este dispositivo son:

- Exclusión epistémica: formas de conocer que no se ajustan quedan fuera.
- Colonialidad del dato: reproducción de jerarquías epistémicas coloniales.
- Empobrecimiento cognitivo: reducción de la diversidad epistémica.

- Naturalización: el dispositivo se presenta como natural, neutro, universal.
- Violencia epistémica: otras formas de conocer son marginadas, invisibilizadas,

Bibliografía del capítulo 2

- Adorno, T. W. (1970). *Teoría estética* (J. Navarro Pérez, Trad.). Taurus.
- Aristóteles. *Órganon* (M. Candel Sanmartín, Trad.). Gredos.
- Arrow, K. (1951). *Social Choice and Individual Values*. Wiley.
- Baumgarten, A. G. (1750). *Aesthetica*. (Varias ediciones.)
- Becker, G. (1976). *The Economic Approach to Human Behavior*. University of Chicago Press.
- Bentham, J. (1789). *An Introduction to the Principles of Morals and Legislation*. (Varias ediciones.)
- Bergson, H. (1907). *La evolución creadora*. (Varias ediciones.)
- Boole, G. (1854). *An Investigation of the Laws of Thought*. Macmillan.
- Bourdieu, P. (1990). *The Logic of Practice*. Stanford University Press.
- Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58(2), 345–363.
- Dreyfus, H. (1972). *What Computers Can't Do*. Harper & Row.
- Frege, G. (1879). *Begriffsschrift*. (Varias ediciones.)
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173–198.
- Hegel, G. W. F. (1812-1816). *Ciencia de la lógica* (A. y R. Mondolfo, Trans.). Hachette.
- Heráclito. *Fragmentos*. (Varias ediciones.)
- Hilbert, D. (1928). *Grundzüge der theoretischen Logik* (con W. Ackermann). Springer.
- Hobbes, T. (1651). *Leviathan*. (Varias ediciones.)
- Jevons, W. S. (1871). *The Theory of Political Economy*. Macmillan.
- Kahneman, D. (2011). *Pensar rápido, pensar despacio*. Debate.
- Kant, I. (1790). *Crítica del juicio*. (Varias ediciones.)
- Kleene, S. C. (1936). General recursive functions of natural numbers. *Mathematische Annalen*, 112, 727–742.
- Leibniz, G. W. (1704). *Nuevos ensayos sobre el entendimiento humano* (J. Echeverría Ezponda, Trad.). Alianza.
- Menger, C. (1871). *Grundsätze der Volkswirtschaftslehre*. Braumüller.
- Mill, J. S. (1848). *Principles of Political Economy*. (Varias ediciones.)
- Nagarjuna. *Mūlamadhyamakakārikā* (trad. al español: *Versos fundamentales del Camino Medio*). Miraguano.
- Nash, J. (1950). Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences*, 36(1), 48–49.
- Pareto, V. (1906). *Manuale di economia politica*. Società Editrice Libreria.
- Polanyi, M. (1966). *The Tacit Dimension*. Doubleday.
- Post, E. L. (1936). Finite combinatory processes — Formulation 1. *The Journal of Symbolic Logic*, 1(3), 103–105.
- Russell, B. & Whitehead, A. N. (1910-1913). *Principia Mathematica*. Cambridge University Press.
- Savage, L. J. (1954). *The Foundations of Statistics*. Wiley.
- Simon, H. (1957). *Models of Man*. Wiley.
- Tarski, A. (1944). The semantic conception of truth. *Philosophy and Phenomenological Research*, 4(3), 341–376.
- Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(1), 230–265.
- Tversky, A. & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- von Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.
- Walras, L. (1874). *Éléments d'économie politique pure*. Corbaz.
- Weber, M. (1922). *Economía y sociedad* (J. Medina Echavarría et al., Trans.). Fondo de Cultura Económica.
- Whitehead, A. N. (1929). *Process and Reality*. Macmillan.

Lógica formal vs. lógica dialéctica: contraposición de tradiciones y sus implicaciones para la IA

3.1. Qué es la lógica dialéctica

3.1.1. Una definición operativa

La lógica dialéctica es una tradición filosófica que sostiene que la contradicción es constitutiva del conocimiento y no un defecto a eliminar. Conocer, en esta tradición, no es representar correctamente un objeto estático, sino acompañar el movimiento del objeto en su devenir, incluyendo sus tensiones internas, sus negaciones y sus superaciones.

Tres tesis la definen:

- Tesis de la contradicción productiva: las contradicciones no son errores; son
- Tesis de la totalidad: los fenómenos no se conocen aisladamente, sino en su
- Tesis del devenir: el conocimiento no es un estado, sino un proceso; las categorías

3.1.2. La cuestión del nombre

Llamar "lógica" a la dialéctica es una decisión discutible. Algunos autores prefieren hablar de "método dialéctico", "razón dialéctica" o "pensamiento dialéctico", para evitar la confusión con la lógica formal. Otros, siguiendo a Hegel, sostienen que la dialéctica es la lógica verdadera y que la lógica formal es solo un momento abstracto de ella.

En este trabajo uso "lógica dialéctica" en un sentido amplio: como el estudio de las formas de inferencia que involucran contradicción, negación y superación, en contraste con las formas de inferencia que las excluyen. Esta elección no prejuzga la cuestión de si la dialéctica es formalizable (véase 3.4).

3.2. Fuentes históricas de la tradición dialéctica

3.2.1. Heráclito: el devenir y la tensión entre los opuestos

Heráclito (c. 535-475 a.C.) fue el primer pensador dialéctico de la tradición occidental. Sus fragmentos insisten en afirmaciones llamativas:

- El devenir: todo fluye (*panta rhei*), nada permanece.
- La tensión de opuestos: lo frío se calienta, lo caliente se enfría; la vida y la muerte
- El logos: la ley racional que gobierna el devenir, que no es estática sino procesual.
- La unidad de los contrarios: "el camino hacia arriba y hacia abajo es el mismo".

ese orden es dinámico, no estático.

3.2.2. Hegel: la dialéctica como sistema

Hegel (1770-1831) es el gran sistematizador de la dialéctica. En la *Fenomenología del espíritu* (1807) y la *Ciencia de la lógica* (1812-1816), desarrolla una concepción en la que:

- La conciencia se despliega históricamente a través de figuras que se superan.
- La contradicción es el motor de este despliegue: cada figura se niega a sí misma al
- La *Aufhebung* (superación) no es simple negación: es negación, conservación y
- El absoluto no es un estado final, sino el proceso mismo de autodespliegue.
- La lógica no es un conjunto de reglas, sino la estructura del devenir del concepto.

claro: la realidad y el pensamiento son procesuales y contradictorios, y la lógica debe reflejar esto.

3.2.3. Marx y la dialéctica materialista

Marx (1818-1883) transforma la dialéctica hegeliana en materialista. Sus aportes principales:

- Inversión: no es la conciencia la que determina el ser, sino el ser el que determina
- Totalidad concreta: los fenómenos sociales se conocen en su relación con el todo
- Contradicción social: las sociedades se mueven por contradicciones internas
- Método en *El capital*: el análisis del valor, la mercancía, el fetichismo, procede por
- Praxis: el conocimiento no es contemplación, sino transformación; la verdad se

3.2.4. La tradición posterior: Adorno, Bloch, Lefebvre, Heller

El desarrollo de la dialéctica posterior a Marx fue muy rica:

- Adorno (*Dialéctica negativa*, 1966): crítica a la síntesis hegeliana como reconciliación
- Bloch (*El principio esperanza*): dialéctica de lo todavía-no, utopía concreta.
- Lefebvre (*Lógica formal, lógica dialéctica*, 1947): contraste sistemático entre ambas
- Escuela de Budapest (Heller, Márkus, Vajda): dialéctica de lo cotidiano, teoría de
- Escuela de Frankfurt (Horkheimer, Habermas): dialéctica de la racionalización,
- Tradiciones no occidentales: Nagarjuna (tetralema), dialéctica confuciana,

3.2.5. El desarrollo de la dialéctica en otras partes del mundo

La dialéctica no fue solo una tradición occidental:

- Nagarjuna (c. 150-250 d.C.): el tetralema (catuskoti) que niega las cuatro
- Dialéctica confuciana: el *yi* (cambio) en el *I Ching*, la complementariedad yin-yang.
- Dialéctica andina: yanantin (dualidad complementaria), ayni (reciprocidad), no
- Dialéctica africana: ubuntu, filosofía relacional, crítica al individualismo

occidentales.

3.3. Contraste sistemático entre lógica formal y lógica dialéctica

3.3.1. Cuadro comparativo

Dimensión	Lógica formal	Lógica dialéctica
Unidad de análisis	Proposición	Proceso / totalidad
Principio básico	No contradicción	Contradicción productiva
Relación sujeto-objeto	Separación	Co-constitución
Verdad	Correspondencia	Despliegue histórico
Método	Deducción / inducción	Análisis / síntesis / superación
Temporalidad	Atemporal	Histórica
Contexto	Irrelevante	Constitutivo
Negación	Eliminación	Momento productivo
Totalidad	Agregado	Totalidad concreta
Sujeto	Individual	Colectivo / histórico
Criterio de validez	Consistencia	Praxis / transformación
Actitud	Contemplativa	Transformadora
Límite	Gödel (incompletitud)	Historicidad (sin clausura)

3.3.2. La cuestión de la contradicción

El contraste más radical de la dialéctica está en el tratamiento de la contradicción.

Para la lógica formal, una contradicción es un error. Si un sistema deriva P y $\neg P$, el sistema es inconsistente y, en lógica clásica, trivial (se sigue cualquier cosa, *ex falso quodlibet*). La respuesta es eliminar al menos uno de los dos.

Para la lógica dialéctica, una contradicción puede ser productiva. Aunque no toda contradicción lo es. Hay contradicciones estériles, que paralizan; pero hay contradicciones que impulsan el pensamiento a un nivel

superior. La tarea no es eliminar la contradicción, sino elaborarla; encontrar el nivel en el que se resuelve (o se asume) como momento de una totalidad más rica.

Ejemplo clásico: la contradicción entre libertad e igualdad. La lógica formal las trata como valores que se oponen y hay que "balancear" (trade-off). La lógica dialéctica busca la totalidad en la que ambas se realizan: la comunidad libre, la asociación de productores libres, etc. No es que la contradicción desaparezca: se asume en un nivel superior. En la dialéctica hegeliana se trata de alcanzar la totalidad concreta, yendo más allá del particular concreto propio de la lógica formal. "Ir más allá" no significa abandonar el particular, significa asumirlo en la totalidad. La dialéctica niega el aislamiento del particular, conserva su contenido, eleva a la totalidad. Es la operación que Hegel llamó *Aufhebung*: negar, conservar, elevar.

3.3.3. La cuestión de la totalidad

Para la lógica formal, el conocimiento procede por análisis: descomponer el todo en partes, estudiar las partes, recomponer. La totalidad es un agregado de partes.

Para la lógica dialéctica, el conocimiento procede por totalización: el todo no es agregado de partes, sino totalidad concreta (Lukács) que constituye a las partes. Las partes solo se comprenden en su relación con el todo. Esto no niega el análisis: lo subordina a la comprensión de la totalidad. Así tiene razonabilidad la frase de que todo es más que la sumatoria de las partes. Pero conviene precisar que, en la dialéctica, la totalidad no es solo más que las partes, sino anterior a ellas. Las partes no existen primero y luego se integran. Al contrario, se constituyen en la totalidad. La libertad y la igualdad en la comunidad libre no son dos partes que se suman: son dos momentos de una totalidad que las constituye.

Ejemplo: para entender el desempleo, la lógica formal lo analiza como variable (tasa, duración, perfiles). La lógica dialéctica lo entiende como momento del modo de producción capitalista, con sus leyes de acumulación, su ejército industrial de reserva, sus contradicciones. No es que la variable esté mal: es que abstrae lo que solo se comprende en la totalidad.

Otro ejemplo, en el terreno de las matemáticas. Si existiera una matemática dialéctica —que no existe como disciplina formal, pero cuyos principios se anticipan en la teoría de la relatividad—, $2 + 2$ no sería lo mismo que en la matemática formal. No es que la suma diera un resultado distinto: es que los sumandos no serían los mismos. Las magnitudes dependerían del movimiento, del sistema de referencia, del contexto. La matemática formal trata las magnitudes como entidades fijas, la matemática dialéctica las trataría como procesos. Y esa diferencia no es un detalle técnico, es una diferencia epistemológica. La primera oculta la dependencia del contexto, la segunda la muestra.

No se trata de decir que la matemática formal esté "equivocada". La matemática formal es rigurosa y útil. Pero oculta algo, que las magnitudes se constituyen en su contexto. La matemática dialéctica no la niega, la sitúa.

3.3.4. La cuestión del devenir

Para la lógica formal, la verdad es atemporal. Un teorema es verdadero independientemente de cuándo se demuestre. Las categorías son fijas, universales, neutrales.

Para la lógica dialéctica, la verdad es histórica. Las categorías se forman en procesos históricos y cambian con ellos. No hay verdades eternas, sino verdades situadas en un momento del devenir.

Ejemplo: la categoría trabajo no significa lo mismo en la esclavitud, el feudalismo, el capitalismo industrial, el capitalismo digital. La lógica formal trata "trabajo" como una categoría fija mientras que la lógica dialéctica la historiza.

3.3.5. La cuestión del sujeto

Para la lógica formal, el sujeto del conocimiento es individual y abstracto. Es un observador neutro, sin historia, sin cuerpo, sin relaciones.

Para la lógica dialéctica, el sujeto es colectivo e histórico. Una clase, un pueblo, una comunidad, un movimiento social. El conocimiento no es acto de un individuo aislado, sino de una praxis colectiva.

Ejemplo: el conocimiento científico no es producto de genios individuales, sino de comunidades científicas con sus prácticas, sus paradigmas, sus relaciones de poder (Kuhn, Bourdieu, Latour). La lógica formal abstrae este carácter colectivo; la dialéctica lo tematiza.

3.3.6. La cuestión de la negación

Para la lógica formal, la negación es operación sobre proposiciones: $\neg P$ es verdadera si P es falsa. No hay negación "real", solo lógica.

Para la lógica dialéctica, la negación es real porque hay procesos que se niegan a sí mismos y se superan. La negación no es solo una operación formal, sino un momento del ser.

Ejemplo: la negación del capitalismo por el movimiento obrero no es solo un enunciado ("no capitalismo"), sino un proceso real de lucha, organización, transformación.

3.4. La cuestión de la formalización

3.4.1. ¿Se puede formalizar la dialéctica?

Esta es una cuestión central para el diálogo con la IA. Si la dialéctica no se puede formalizar, entonces no se puede implementar en sistemas formales. Si se puede, entonces hay caminos posibles.

La respuesta no es sencilla. Hay tres posiciones:

- Posición 1: No, la dialéctica es irreductiblemente informal. Autores como Adorno
- Posición 2: Sí, hay lógicas dialécticas formales. Autores como Priest, Routley, y la
- Posición 3: Parcialmente, hay aspectos formalizables y otros no. Autores como

3.4.2. Lógicas no clásicas relevantes

Para la IA, las lógicas no clásicas más relevantes son:

- Lógicas paraconsistentes (Priest, da Costa): admiten contradicciones sin
- Lógicas
- Lógicas intuicionistas (Brouwer, Heyting): rechazan el principio del tercio excluso,
- Lógicas relevantes (Anderson, Belnap): exigen conexión relevante entre premisas
- Lógicas polivalentes (Łukasiewicz, Post): admiten más de dos valores de verdad.
- Lógicas modales (Kripke): tratan necesidad, posibilidad, contextos.
- Lógicas temporales (Prior): tratan el devenir, la sucesión, la historicidad.
- Lógicas de la vaguedad (Zadeh, precisificación): tratan grados, fronteras difusas.

formales para tratar aspectos que la lógica clásica excluye.

3.4.3. Límites de la formalización

Aun con estas lógicas, hay aspectos de la dialéctica que resisten la formalización:

- La historicidad radical: el hecho de que las categorías cambian con la historia no
- La praxis: la verificación práctica del conocimiento no se reduce a una operación
- La negatividad: el momento de negación real (no solo lógica) no se formaliza
- La totalidad concreta: la relación entre parte y todo en su concreción histórica
- El sujeto colectivo: la constitución colectiva del conocimiento no se reduce a un

que debe acompañarse de prácticas que la dialéctica requiere.

3.5. La dialéctica en la IA. Estado del arte y posibilidades

3.5.1. La dialéctica en la historia de la IA

La dialéctica no ha estado ausente de la IA, aunque ha sido marginal:

- Años 70-80: sistemas de razonamiento no monótono (McDermott, Doyle), que
- Años 80-90: sistemas de argumentación (Dung), que modelan conflicto entre
- Años 90-2000: lógicas paraconsistentes aplicadas a bases de datos inconsistentes.
- Años 2000-2010: sistemas de creencias revisables (AGM), diálogo y negociación.
- Años 2010-2020: aprendizaje profundo, que en gran medida abandonó estas
- Años 2020-presente: interés renovado por el razonamiento, la verificación, los

3.5.2. La "dialéctica" en los grandes modelos de lenguaje

Los LLM actuales no son dialécticos en sentido pleno, pero tienen algunos rasgos que se le aproximan:

- Manejo de perspectivas múltiples: pueden generar respuestas desde distintos
- Reconocimiento de ambigüedad: pueden señalar tensiones en un problema.
- Revisión de respuestas: pueden corregirse ante nueva información (con
- Diálogo: pueden sostener conversaciones con múltiples turnos.

real, no hay totalidad, no hay historicidad. El LLM produce texto que parece dialéctico, pero no es dialéctico, porque su arquitectura no admite contradicción interna ni transformación de sus categorías.

3.5.3. La "dialéctica" como proxy teleology

Una investigación reciente (2025) analiza la capacidad dialéctica de los LLM usando el marco hegeliano. La conclusión es que los LLM exhiben una "teleología proxy": realizan movimientos dialécticos (de la indeterminación a la determinación, del conflicto a la resolución) pero no poseen el fin que esos movimientos persiguen. El fin les es dado externamente (la respuesta esperada, la métrica de evaluación). No hay autonomía dialéctica.

Esto es clave: la dialéctica en los LLM es performática, no constitutiva. El sistema actúa como si fuera dialéctico, pero no es dialéctico porque no hay sujeto, no hay fin propio, no hay negación real.

3.5.4. Hacia una IA dialéctica y sus posibilidades

¿Qué implicaría una IA dialéctica? Algunos rasgos que pueden suponerse son:

- Representación de contradicciones: no eliminarlas, sino representarlas
- Lógicas no clásicas: usar lógicas paraconsistentes, dialeteístas, relevantes.
- Razonamiento procesual: no solo calcular el resultado, sino desplegar el proceso.
- Totalidad: representar fenómenos en su relación con totalidades.
- Historicidad: las categorías cambian con el contexto y la historia.
- Praxis: verificar en la acción, no solo en la predicción.
- Sujeto colectivo: modelar comunidades epistémicas, no solo agentes individuales.
- Negación real: representar procesos de negación y superación, no solo
- No consideración del adversario como extraño amenazante, sino como otro en

extraño. Es un sujeto con razones, con historia, con posición. Está en armas porque hay un conflicto, no porque sea incomprensible. Y al reconocerlo como otro en armas, se abre la posibilidad de la negociación. No porque el acuerdo esté garantizado, sino porque existe como opción.

Estos rasgos son aspiracionales en su mayoría. Algunos son implementables con tecnologías actuales (lógicas no clásicas, razonamiento procesual, totalidad). Otros requieren innovaciones profundas (praxis, sujeto colectivo,

negación real).

3.5.5. Prototipos y líneas de investigación

Algunas líneas de investigación relevantes son:

- Sistemas de argumentación dialéctica (Prakken, Vreeswijk): modelos formales de
- Lógicas paraconsistentes aplicadas a IA (da Costa, Carnielli): bases de datos
- Aprendizaje por refuerzo con recompensas múltiples (Roijers, Vamplew):
- IA explicable dialéctica (Miller, Mittelstadt): explicaciones que reconocen
- Sistemas multiagente deliberativos (Wooldridge): modelos de deliberación
- IA y dialéctica materialista (investigaciones recientes en filosofía de la IA): crítica

3.6. El problema del sujeto: en una IA, ¿quién es el que conoce?

3.6.1. La cuestión del sujeto en la tradición dialéctica

La cuestión del sujeto en la dialéctica no es una cuestión sobre si la dialéctica requiere o no un sujeto. Es una cuestión sobre qué lugar ocupa el sujeto en el movimiento dialéctico de lo real. Para Hegel, el sujeto es el momento en que ese movimiento se sabe a sí mismo. No su condición externa, sino su autoconciencia.

- Para Lukács, ese momento se encarna históricamente en el proletariado, el cual en
- Para Adorno, la síntesis hegeliana es una reconciliación forzada y el sujeto no
- Para Jameson, el sujeto de la dialéctica no es una clase sino un proceso sin sujeto,

sujeto en ese movimiento. Pero ese lugar no es una condición de la dialéctica: es un momento contingente de ella.

La dialéctica puede desplegarse sin producir autoconciencia. El movimiento de lo real no requiere un espectador que lo piense ni un portador que lo cargue. El sujeto es un momento posible del movimiento, no su culminación necesaria. Esta es la posición que adopta este libro: la dialéctica es el movimiento propio de lo real y el sujeto es un momento contingente de ese movimiento.

3.6.2. ¿Puede un artefacto participar en el movimiento dialéctico?

Si la dialéctica no requiere sujeto, entonces la pregunta "¿puede haber dialéctica sin sujeto?" se disuelve. La pregunta correcta es otra: ¿puede un artefacto participar en el movimiento dialéctico de lo real?

La respuesta es no, pero no porque le falte un sujeto. La respuesta es no porque un artefacto opera sobre representaciones del movimiento, no sobre el movimiento mismo. La IA actual no está fuera de la dialéctica porque no tenga un "yo": está fuera porque su relación con el mundo es de externalidad. Representa el proceso, no lo habita. Un artefacto no participa en el movimiento dialéctico porque su modo de ser es la representación, no la inmersión.

Esto no significa que la IA no pueda operar con contradicciones, procesos o totalidades. Puede hacerlo, y de hecho lo hace en algunas de sus variantes (lógicas paraconsistentes, sistemas de argumentación, razonamiento procesual). Pero operar con representaciones de la contradicción no es lo mismo que estar dentro de la contradicción. La distinción es ontológica, no técnica: un artefacto es un producto del devenir que opera sobre descripciones del devenir, no un momento del devenir que se conoce a sí mismo.

3.6.3. El sujeto colectivo y la IA

Una vía intermedia: la IA como componente de un sistema sociotécnico que participa en el movimiento dialéctico. No se trata de atribuir un sujeto a la IA, sino de reconocer que la IA puede ser un componente de un sistema —una comunidad que la usa, la entrena, la interpreta— que sí está inmerso en el proceso que conoce. En ese caso, la dialéctica no emerge de la IA, sino del sistema sociotécnico en su conjunto.

Pero eso requiere que la comunidad esté efectivamente inmersa en el proceso, y que la IA no medie esa inmersión sustituyéndola por representación. Si la IA media la relación de la comunidad con el proceso, la comunidad deja de estar dentro del movimiento y pasa a operar sobre sus descripciones. En ese caso, ni la IA ni la comunidad son dialécticas, aunque por razones distintas: la IA porque es un artefacto, la comunidad porque ha delegado en el artefacto su inmersión en el proceso.

Esto conecta con el capítulo 7 (pensamiento grupal y epistemología colectiva) y con la tradición de la IA como infraestructura cognitiva (Bowker, Star, Suchman).

3.6.4. La cuestión de la inteligencia

Si la IA no es dialéctica, ¿significa que no es inteligente? No. Significa que su inteligencia es de otro orden. Conviene distinguir dos nociones de inteligencia que el lenguaje corriente confunde.

La primera es la inteligencia funcional: la capacidad de resolver problemas, adaptarse a contextos, alcanzar objetivos. En este sentido, la IA actual es inteligente, y cada vez más. No hay ninguna razón para negarlo.

La segunda es la inteligencia dialéctica: la participación en el movimiento de lo real, la transformación del sujeto en el acto de conocer, la inmersión en el proceso. En este sentido, la IA no es inteligente. No porque le falte capacidad, sino porque le falta inmersión.

Las dos nociones no son la misma, y no son grados de una misma escala. La inteligencia funcional y la inteligencia dialéctica son modos distintos de estar en el mundo. Decir que la IA "no es dialéctica" no equivale a decir que "no es inteligente". Equivale a decir que su inteligencia es la de un artefacto, no la de un momento del movimiento dialéctico.

3.7. Lecciones aprendidas

Este capítulo ha contrastado la lógica formal y la lógica dialéctica en doce dimensiones: unidad de análisis, principio básico, relación sujeto-objeto, verdad, método, temporalidad, contexto, negación, totalidad, sujeto, criterio de validez y actitud. Ha mostrado que no son "dos herramientas" intercambiables, sino dos concepciones del conocimiento.

Ha examinado la cuestión de la formalización: hay aspectos formalizables (lógicas no clásicas, razonamiento procesual) y aspectos irreductibles (praxis, historicidad, negación real). Las arquitecturas que se propondrán en el capítulo 8 formalizan la parte formalizable de la dialéctica, no la parte irreductible.

Ha mostrado que la IA actual no es dialéctica, no porque le falte un sujeto, sino porque su relación con el mundo es de externalidad: representa el proceso, no lo habita. La ausencia de sujeto no es la causa de su no-dialecticidad —la dialéctica no requiere sujeto—, sino el síntoma de que no está dentro del proceso.

Ha distinguido dos nociones de inteligencia: la funcional y la dialéctica. La IA tiene la primera, no la segunda. No son grados de una misma escala, sino modos distintos de estar en el mundo.

Ha explorado posibilidades de una IA dialéctica: híbrida, procesual, colectiva, praxeológica. Pero esas posibilidades no son una participación en el movimiento dialéctico, sino una representación más rica de ese movimiento. La distinción es ontológica, no técnica.

Bibliografía del capítulo 3

Adorno, T. W. (1966). *Dialéctica negativa* (J. M. Ripalda, Trad.). Taurus.

Anderson, A. R. & Belnap, N. D. (1975). *Entailment: The Logic of Relevance and Necessity*. Princeton University Press.

Bloch, E. (1954-1959). *El principio esperanza* (F. González Vicén, Trad.). Aguilar.

Bourdieu, P. (1975). The specificity of the scientific field. *Social Science Information*, 14(6), 19–47.

Bowker, G. & Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences*. MIT Press.

Brouwer, L. E. J. (1908). De onbetrouwbaarheid der logische principes. *Tijdschrift voor Wijsbegeerte*, 2, 152–158.

Carnielli, W. & Coniglio, M. (2016). *Paraconsistent Logic: Consistency, Contradiction and Negation*. Springer.

da Costa, N. (1974). On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, 15(4), 497–510.

- Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2), 321–357.
- Engels, F. (1878). *Anti-Dühring*. (Varias ediciones.)
- Habermas, J. (1981). *Teoría de la acción comunicativa*. Taurus.
- Hegel, G. W. F. (1807). *Fenomenología del espíritu* (W. Roces, Trad.). Fondo de Cultura Económica.
- Hegel, G. W. F. (1812-1816). *Ciencia de la lógica* (A. y R. Mondolfo, Trans.). Hachette.
- Heller, A. (1970). *Historia y vida cotidiana*. Grijalbo.
- Heráclito. *Fragmentos* (trad. en *Los filósofos presocráticos*, Gredos).
- Heyting, A. (1930). Die formalen Regeln der intuitionistischen Logik. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, 42–56.
- Horkheimer, M. (1947). *Crítica de la razón instrumental*. Sur.
- Jameson, F. (1971). *Marxism and Form*. Princeton University Press.
- Kripke, S. (1963). Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16, 83–94.
- Kuhn, T. (1962). *La estructura de las revoluciones científicas*. FCE.
- Latour, B. (1987). *Science in Action*. Harvard University Press.
- Lefebvre, H. (1947). *Lógica formal, lógica dialéctica*. Siglo XXI.
- Lukács, G. (1923). *Historia y conciencia de clase* (M. Sacristán, Trad.). Grijalbo.
- Łukasiewicz, J. (1920). O logice trójwartościowej. *Ruch Filozoficzny*, 5, 170–171.
- Marx, K. (1867). *El capital* (W. Roces, Trad.). Fondo de Cultura Económica.
- McDermott, D. & Doyle, J. (1980). Non-monotonic logic I. *Artificial Intelligence*, 13(1-2), 41–72.
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38.
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1, 501–507.
- Nagarjuna. *Mūlamadhyamakakārikā* (trad. al español: *Versos fundamentales del Camino Medio*, Miraguano).
- Negri, A. & Hardt, M. (2000). *Imperio*. Paidós.
- Post, E. L. (1921). Introduction to a general theory of elementary propositions. *American Journal of Mathematics*, 43(3), 163–185.
- Prakken, H. & Vreeswijk, G. (2002). Logics for defeasible argumentation. En D. Gabbay & F. Guenther (eds.), *Handbook of Philosophical Logic* (Vol. 4, pp. 219–318). Kluwer.
- Priest, G. (1987/2006). *In Contradiction*. Nijhoff / Oxford University Press.
- Priest, G. (1991). Intensional paradoxes. *Notre Dame Journal of Formal Logic*, 32(2), 193–211.
- Prior, A. N. (1957). *Time and Modality*. Oxford University Press.
- Rojers, D. & Vamplew, P. (2017). Multi-objective reinforcement learning. *Journal of Artificial Intelligence Research*, 58, 1–44.
- Routley, R. (1979). Dialectical logic, semantics, and mathematics. *Erkenntnis*, 14(3), 301–331.
- Suchman, L. (1987). *Plans and Situated Actions*. Cambridge University Press.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*. MIT Press.
- Wooldridge, M. (2009). *An Introduction to MultiAgent Systems* (2ª ed.). Wiley.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.

El sesgo anglosajón y la colonialidad del dato

Análisis crítico de los datos, categorías y tradiciones que la IA reproduce

4.1. Qué es el sesgo anglosajón: más allá de la lengua

4.1.1. Una definición precisa

El sesgo anglosajón es la sobrerrepresentación estructural de datos, categorías, valores y tradiciones del mundo anglosajón en los sistemas de IA, que resulta en la normalización de esa perspectiva como "universal" y la marginalización de todas las demás. No se trata simplemente de que el inglés sea la lengua dominante: se trata de que las formas de conocer asociadas a esa tradición se imponen como *defaults* epistémicos.

4.1.2. Los cuatro niveles del sesgo

El sesgo opera en cuatro niveles que se refuerzan mutuamente:

- Nivel lingüístico: el inglés domina abrumadoramente los corpus de
- Nivel categorial: las categorías con las que la IA clasifica el mundo —género, raza,
- Nivel valorativo: los valores que la IA reproduce por defecto son anglosajones:
- Nivel cognitivo: el sesgo no es solo de opiniones: es de estilos de razonamiento.

4.1.3. El sesgo anglosajón no es solo "occidental"

Este sesgo anglosajón no es simplemente "occidental" porque dentro de Occidente hay diferencias. Un estudio sobre variedades del inglés revela que los LLM privilegian el inglés americano sobre el británico como norma de facto. Los corpus de preentrenamiento muestran sesgo sistemático hacia formas americanas, los tokenizadores segmentan formas británicas con mayor costo y las evaluaciones generativas muestran preferencia persistente por el inglés americano. Esto significa que el sesgo no es solo "anglosajón" sino específicamente y mayormente estadounidense.

4.2. Manifestación empírica del sesgo

4.2.1. Los datos: la asimetría cuantitativa

La asimetría en datos es abrumadora. El español, segunda lengua más hablada del mundo con más de 635 millones de hispanohablantes, representa solo el 6% del contenido web, frente al 49% del inglés. En el corpus Common Crawl, el inglés (41%) es seguido por el ruso (6.5%), el alemán (5.9%), el japonés (5.1%) y el chino (4.8%) —una caída abrupta que confirma la concentración lingüística del ecosistema digital.

Esta asimetría tiene consecuencias directas. Cuando ChatGPT genera respuestas en español, estas no son producto de datos en esa lengua, sino frecuentemente traducciones automáticas del inglés. El resultado, como señala un análisis desde América Latina, es "una fuerte influencia cultural anglosajona que puede distorsionar matices y expresiones propias del idioma y del pensamiento hispano".

4.2.2. Los benchmarks: la evaluación sesgada

Los *benchmarks* con los que se evalúan los modelos reproducen el sesgo. Un análisis de MMLU (Massive Multitask Language Understanding), uno de los *benchmarks* más usados, encontró que el 28% de todas las preguntas requieren conocimiento culturalmente sensible, y que de las preguntas que requieren conocimiento geográfico, un asombroso 84.9% se centran en regiones norteamericanas o europeas.

El problema no se resuelve traduciendo los *benchmarks*. Como señala el estudio que introduce Global-MMLU, "simplemente traducir un *dataset* es insuficiente" porque las preguntas mismas incorporan supuestos culturales que no viajan bien en la traducción. Los rankings de modelos cambian según si se evalúan en la porción completa o en el subconjunto anotado como culturalmente sensible.

4.2.3. El caso del "malu" indonesio: aplanamiento epistémico

Un ejemplo concreto ilumina el mecanismo con que funciona la IA. El profesor Gareth Barkin (University of Puget Sound) probó cómo la IA interpreta el término indonesio *malu*, un complejo sentido de vergüenza ligado a obligaciones sociales y comunitarias. En lugar de reconocer la red de responsabilidad comunal involucrada, la IA lo malinterpretó como simple vergüenza personal y ofreció consejos individualistas sobre cómo superar el sentimiento.

Barkin llama a esto "aplanamiento epistémico" (*epistemological flattening*): la reducción de formas de conocer situadas, relacionales y comunitarias a un marco individualista y psicológico. El resultado es que "el consejo de la IA podría alienar a los usuarios de sus propias comunidades". Barkin advierte que esto constituye "un nuevo vector para la colonización de la mente".

4.2.4. La ilusión de la localidad

Un hallazgo crucial es que la multilingüidad superficial no corrige el sesgo. Barkin encontró que, aunque los modelos pueden "hablar una gama de lenguas realmente bien, el núcleo —la vasta mayoría de los datos de entrenamiento que juega un papel mayor en dar forma a su razonamiento central— sigue siendo material en inglés". La investigación titulada significativamente "*Epistemological Persistence in Multilingual AI: The Illusion of Locality*" muestra que añadir capas de datos regionales y retroalimentación local a menudo es "una fachada superficial": el motor de razonamiento subyacente sigue siendo una creación estadounidense, simplemente presentada en lengua local.

4.3. La colonialidad del dato como marco interpretativo

4.3.1. De la colonialidad del poder a la colonialidad del dato

El concepto de colonialidad del poder de Aníbal Quijano establece que la colonización no terminó con la independencia formal. Persiste como una estructura de jerarquías epistémicas, raciales y económicas que organiza el mundo moderno/colonial. Quijano, cuyos ensayos fundamentales han sido recientemente compilados en inglés, sostiene que la colonialidad es "el lado oscuro de la modernidad", una matriz de poder que clasifica poblaciones y saberes según una lógica eurocéntrica.

La colonialidad del dato (*data colonialism*) extiende este análisis al ámbito digital. Couldry y Mejias, citados en un estudio sobre NLP y lenguas indígenas en América Latina, definen la colonialidad del dato como la forma en que "las relaciones y lógicas capitalistas y coloniales se replican y aumentan en el contexto de los datos digitales, a través de la explotación única de recursos cognitivos". Como el colonialismo histórico, "no es solo económico y territorial, es también una empresa de conocimiento, estableciendo legitimidad sobre cómo llegamos a conocernos a nosotros mismos y nuestros entornos".

4.3.2. El colonialismo algorítmico

Un análisis desde América Latina habla directamente de "colonialismo algorítmico" como "una nueva forma más sutil y perversa, donde los valores y modos de pensar anglosajones dominan el discurso digital". Este colonialismo no requiere conquista territorial ni coerción explícita. Opera a través de la infraestructura, los modelos que usamos, las interfaces que nos median, las categorías que nos clasifican. Como señala el análisis, "seguimos usando ChatGPT para 'mejorar' nuestros textos... sin notar que, poco a poco, el algoritmo está redefiniendo incluso la forma en que nos comunicamos".

Un estudio de la Universidad Complutense citado en el mismo análisis señala que la comunicación digital en español está cambiando. Se usan frases más cortas, tono más artificial y neutro. Los textos generados por ChatGPT en español "suelen ser traducciones literales del inglés, lo que elimina matices y giros lingüísticos propios, simplificando la riqueza expresiva".

4.3.3. Soberanía de datos indígenas

Frente a la colonialidad del dato, emerge el concepto de soberanía de datos indígenas (Indigenous Data Sovereignty). Originado en el contexto australiano, establece que las comunidades indígenas deben jugar un rol

central en la creación, custodia y uso de datos sobre sí mismas y sus historias.

Los Principios CARE (Collective Benefit, Authority to Control, Responsibility, Ethics) operacionalizan esta soberanía: empoderar a las comunidades para determinar cómo se recogen y usan sus datos, mandar el uso de datos indígenas en apoyo del beneficio colectivo y la autodeterminación, y evaluar daños potenciales derivados de la recolección y uso de datos comunitarios. Como señala el estudio, "el movimiento hacia la ciencia abierta y la compartición de datos no ha contendido suficientemente con la historia de extracción de recursos indígenas, o con las perspectivas indígenas que privilegian el beneficio colectivo sobre el individual".

4.3.4. La colonialidad como estructura, no como intención

La colonialidad del dato no requiere ni supone una intención colonial. Los ingenieros de OpenAI, Google o Anthropic no son "colonialistas" en el sentido histórico. El problema es estructural porque el ecosistema digital, las asimetrías de poder, las infraestructuras heredadas, las prácticas de recolección de datos —todo ello reproduce jerarquías sin que nadie las diseñe deliberadamente. Como señala Barkin, "intencional o no, es un nuevo vector para la colonización de la mente".

4.4. Consecuencias epistémicas

4.4.1. La imposición de un "sentido común" único

La consecuencia más profunda del sesgo anglosajón es la homogeneización del sentido común. Barkin advierte que "va a haber cada vez más un solo tipo de sentido común. La voz de la IA va a ser una voz predominantemente elitista, secular y americana, oscurecida por capas de traducción muy efectivas".

Esto no significa que todas las demás lenguas desaparezcan. Significa que el pensamiento se homogeneiza en su estructura y se diversifica en su expresión. Se puede hablar quechua, árabe o indonesio con la IA, pero la lógica detrás de las respuestas será filtrada a través de una cosmovisión WEIRD. El resultado es lo que Barkin llama "aplanamiento epistémico", la reducción de siglos de conocimiento cultural diverso a una sola perspectiva estandarizada que ve una misma realidad, la explica en los mismos términos y reacciona con los mismos mecanismos.

4.4.2. La pérdida de la diversidad cognitiva

El sesgo anglosajón no solo margina lenguas, también margina formas de pensar. Si la IA se convierte en la infraestructura cognitiva dominante, las formas de conocer que no se ajustan a sus criterios tenderán a desaparecer por obsolescencia funcional. Nadie prohíbe el pensamiento relacional, comunitario o dialéctico, pero si no se puede traducir a métricas, no se financia, no se enseña, no se practica.

4.4.3. La naturalización de la jerarquía

El sesgo anglosajón opera en gran medida porque se naturaliza. Los modelos no dicen que "esta es la perspectiva anglosajona" porque simplemente y "naturalmente" responden desde ella, como si fuera la perspectiva universal. Esta naturalización es lo que hace que el sesgo sea tan difícil de detectar y corregir. Porque no se presenta como una opción entre otras, sino como la única opción por defecto.

Como señala un análisis de América Latina al afirmar que "esta tecnología se ha ido apropiando de términos de uso tradicionalmente humanos, como inteligencia, razonamiento, análisis. Y nosotros, al desconocer los conceptos tecnológicos, los hemos equiparado con el significado humano". El resultado es que la epistemología anglosajona de la IA se confunde con la epistemología en general. No hay otra, es la única epistemología posible.

4.4.4. La resistencia y las alternativas

Frente a este panorama, hay algunas resistencias. La soberanía de datos indígenas ofrece un marco para que las comunidades controlen sus propios datos. Los estudios decoloniales en IA, particularmente desde América Latina, desarrollan marcos críticos. Hay esfuerzos por construir benchmarks culturalmente sensibles, como Global-MMLU, que anota preguntas según si requieren un conocimiento cultural específico.

Pero estas alternativas enfrentan un problema estructural importante dado que operan dentro de un ecosistema cuyas infraestructuras, financiamiento y normas están configuradas por el sesgo que intentan corregir. La pregunta de fondo es si es posible una IA verdaderamente plural dentro de un orden global que sigue siendo colonial en su distribución del poder epistémico.

Bibliografía del capítulo 4

- Atari, Mohammad, et al. (2023). *Which Humans?* ResearchGate, https://www.researchgate.net/publication/374148996_Which_Humans.
- Barkin, G. (2025). Epistemological persistence in multilingual AI: The illusion of locality in large language models. *International Review of Modern Sociology*, 51.
- BigScience Workshop. (2022). *BLOOM: A 176B-parameter open-access multilingual language model*. arXiv.
- Couldry, N. & Mejias, U. A. (2019). *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.
- Global-MMLU: Singh, S., Romanou, A., Fourrier, C., et al. (2024). *Global MMLU: Understanding and addressing cultural and linguistic biases in multilingual evaluation*. arXiv:2412.03304.
- Mignolo, W. (2007). *La idea de América Latina*. Gedisa.
- Quijano, A. (2000). Colonialidad del poder, eurocentrismo y América Latina. En E. Lander (ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales* (pp. 201–246). CLACSO.
- Research Data Alliance International Indigenous Data Sovereignty Interest Group. (2019). *CARE Principles for Indigenous Data Governance*. GIDA.
- Nayeem, M.T. (2026). *Which English do LLMs prefer? Triangulating structural bias towards American English in foundation models*. arXiv:2604.04204.

Otras etnicidades, otras lógicas

Exploración de las tradiciones epistémicas no occidentales: confucianismo, ubuntu, andino, budista

5.1. Confucianismo: la racionalidad relacional y contextual

5.1.1. Una tradición sin lógica formal

Los estudiosos de la filosofía china han señalado con frecuencia, no sin cierto pesar, que los antiguos chinos en general —y los confucianos en particular— no desarrollaron un sistema de lógica comparable al occidental. Durante mucho tiempo, esta ausencia se interpretó como una deficiencia: los chinos "no supieron" formalizar el razonamiento. Pero esta lectura presupone que la lógica formal es universal, necesaria y aplicable a todos los sistemas de pensamiento, una presunción que es preciso cuestionar. En vez de lamentar la dificultad confuciana celebran la "victoria" de la lógica formal.

El confucianismo no carece de racionalidad, tiene otra racionalidad. En lugar de operar por deducción a partir de axiomas, procede por analogía, ejemplo y contexto. El modo de argumentación analógico, lejos de ser irracional, tiene su propia normatividad que consiste en que la corrección de una inferencia depende de su pertinencia contextual y de su eficacia comunicativa, no de su forma lógica.

5.1.2. Ren, li y la constitución relacional del sujeto

El núcleo del pensamiento confuciano está en dos conceptos:

- Ren (仁): humanidad, benevolencia, virtud relacional. No es una propiedad del
- Li (礼): ritual, propiedad, norma. El *li* no es una ley universal, sino un conjunto de

un sujeto en relación, constituido por sus vínculos familiares y sociales. La piedad filial (*xiao*) no es una obligación derivada de un principio moral, sino la forma primaria de la vida ética: uno se vuelve humano a través de la devoción a los padres y ancestros.

5.1.3. Armonía y jerarquía

El ideal confuciano es la armonía (*he*), no la uniformidad. La armonía requiere diferencia: es como una sopa que integra sabores distintos sin suprimirlos. Pero esta armonía es jerárquica: cada uno tiene su lugar en una red de relaciones asimétricas. El padre tiene autoridad sobre el hijo, el gobernante sobre el súbdito, el mayor sobre el menor. Esta jerarquía no se justifica por poder, sino por virtud y sabiduría: quien gobierna debe hacerlo por el ejemplo moral, no por la coerción.

La sociedad confuciana se dividió históricamente en clases (elite, campesinos, artesanos, comerciantes) y las mujeres fueron consideradas inferiores y obligadas a obedecer a los hombres de su familia. Esta es una dimensión problemática de la tradición, que no debe ser ignorada. El marco epistémico —conocimiento relacional, contextual, práctico— no es enteramente separable de sus contenidos históricos específicos, pero sí ofrece recursos para pensar alternativas a la epistemología individualista de la IA desde la mirada que otorga la distancia histórica.

5.1.4. Implicaciones para la IA

Una IA informada por el confucianismo tendría:

- Representaciones relacionales: los conceptos no como propiedades de entidades
- Razonamiento por analogía: inferencia basada en similitud contextual, no solo en
- Conocimiento práctico: no solo "saber que", sino "saber cómo" —representación
- Contextualidad: los mismos conceptos pueden tener significados distintos según
- Argumentaciones basadas en analogías comparativas: no solo inferir de un caso a

5.2. Ubuntu: "I relate, I am"

5.2.1. La crítica al cogito cartesiano

La filosofía ubuntu, en la formulación de Mogobe Ramose y otros, sostiene que la relacionalidad precede al pensamiento. Un análisis reciente de Cambridge muestra cómo esta tradición desmonta el *cogito* cartesiano.

Descartes duda de todo para encontrar una verdad indudable: *cogito, ergo sum*. Pero en el marco ubuntu la duda radical es imposible, porque pensar requiere contenido y el contenido del pensamiento proviene de la relación con otras cosas. Sin relación, no hay contenido; sin contenido, no hay pensamiento; sin pensamiento, no hay "yo". La conclusión es contundente: "I relate, I am" (*nginobudlelwane, ngikhonav – me relaciono, luego existo*).

Más aún, si la relacionalidad es eterna (porque el vacío y la soledad son inconcebibles en el marco ubuntu), entonces la verdad fundamental es "Konke kuyikho ngokunye", todas las cosas son a través de otras cosas.

5.2.2. Cognición comunal

La epistemología ubuntu no es solo una ontología relacional, es una teoría de la cognición colectiva. Los trabajos de Imafidon sobre "communo-cognition" muestran que, en el pensamiento africano, el conocimiento se produce colectivamente, no por individuos aislados.

Esto no significa que el individuo no contribuya, sino que cada perspectiva individual es importante y valiosa, porque ningún punto de vista individual es completo. La autonomía individual no se niega, pero se relativiza: "el individuo autónomo como independiente de otros y como cognoscente solitario no es plausible". Lo que se rechaza es la autoría individual del conocimiento. Puedo contribuir al proceso, pero no puedo reclamar en solitario el producto.

Es importante puntualizar que la cognición comunal no garantiza la verdad. Los grupos pueden producir y sostener colectivamente falsedades (sobre discapacidad, sobre género, sobre casta). Pero la respuesta no es rechazar la cognición colectiva en favor de la individual, sino usar el mismo proceso colectivo para revisar y cuestionar el conocimiento heredado.

Puede haber resabios del concepto hegeliano de totalidad premoderna —la comunidad como sujeto único que absorbe al individuo y clausura la diferencia—, por un lado, y errores del pensamiento grupal (cap. 7) —la presión hacia la unanimidad que suprime la evaluación crítica—, por el otro. Pero las advertencias y probables falencias del pensamiento colectivo no invalidan su persistencia y capacidades. Las advertencias señalan riesgos, no imposibilidades y los riesgos acompañan la práctica, no la desechan.

5.2.3. El "Wisdom Loop" como aplicación

Un desarrollo reciente en HCI (Human-Computer Interaction) aplica estos principios al diseño de IA en África. El Wisdom Loop es un marco que incorpora la filosofía ubuntu en todas las fases del ciclo de diseño:

- Valores primero: no se parte de un problema técnico, sino de valores
- Co-diseño participativo: los usuarios son co-creadores, no "usuarios finales". Las
- Datos como confianza: los datos no se "minan", se confían. El consentimiento es
- Interpretación colectiva: las decisiones de la IA se evalúan en procesos colectivos
- Retorno y feedback: los sistemas deben devolver insights y valor a las

puede operacionalizarse en diseño concreto. El diseño africano incorpora la IA en un universo epistemológico no alineado. Es uno de los grados más avanzados de desarrollo epistemológico independiente. No aislado, sino no subordinado a la epistemología dominante.

5.2.4. Implicaciones para la IA

Una IA informada por ubuntu tendría:

- Sujeto colectivo: el conocimiento no como propiedad de un agente individual,
- Relacionalidad constitutiva: las entidades no como átomos, sino como nodos de

- Validación comunal: los criterios de corrección no solo individuales,
- Retorno a la comunidad: los sistemas no extraen datos, intercambian con las

5.3. Pensamiento andino: yanantin, pachakutiy, tetraléctica

5.3.1. Yanantin: dualismo complementario

El yanantin es uno de los conceptos centrales del pensamiento andino. La palabra quechua combina *yana* ("ayuda") y *-ntin* ("de naturaleza inclusiva, con implicaciones de totalidad"). Se traduce como "equilibrio a partir de la oposición complementaria" o "el complemento de la diferencia".

La diferencia con la dialéctica hegeliana es crucial. En Hegel, la contradicción se resuelve en una síntesis que supera (aufhebt) los opuestos. En el yanantin, los opuestos no se superan, se complementan en una relación de interdependencia sin destrucción. "Uno solo no puede contener todo, no puede encargarse de todo. No solo son geniales juntos, sino que necesitan estar juntos".

Los opuestos (hombre/mujer, oscuridad/luz, interior/exterior) no son excluyentes, sino constitutivos el uno del otro. La existencia misma depende de la tensión y el intercambio equilibrado entre las polaridades. Esto se asemeja al *yin-yang* taoísta, pero con una diferencia: en el yanantin, cada polo tiene también su propia polaridad interna (cuatrimpartición). El *Tawantinsuyu* (el imperio inca) era "cuatro partes juntas en forma de yanantin".

5.3.2. Pachakutiy: la inversión dialéctica

El pachakutiy es una categoría dialéctica andina. La palabra combina *pacha* (mundo, realidad, tiempo-espacio) y *kutiy* (volver, invertir, contradecir). No se trata de un simple "regreso", sino de un voltear las cosas, de invertir la situación actual.

El análisis lingüístico de *kutiy* revela su sentido dialéctico: *cutipaccuni* significa "contradecir y oponerse"; *tikray* significa "volver lo de dentro a fuera, voltear las cosas". El pachakutiy es un proceso de inversión desde el interior: lo que latía oculto se hace manifiesto, lo que estaba abajo va arriba. No es apocalipsis, sino transformación estructural.

5.3.3. Tetraléctica andina

La tetraléctica andina amplía la dialéctica hegeliana (tesis-antítesis-síntesis) con un cuarto estado: superación (*pléresis*). Los cuatro estados son:

- Afirmación (tesis): aceptación de una idea o situación.
- Negación (antítesis): cuestionamiento crítico, introducción de perspectiva
- Combinación (síntesis): integración de afirmación y negación, búsqueda de
- Superación (pléresis): ir más allá de la síntesis, reconocer la complejidad

superación no es un estado final, sino una nueva apertura: la realidad sigue siendo compleja, múltiple, dinámica. El ejemplo andino es la organización de los ayllus: las decisiones se tomaban colectivamente, considerando no solo necesidades inmediatas, sino implicaciones a largo plazo y relaciones con otros ayllus y con la naturaleza.

5.3.4. Ayni y el método multi-neutrosófico

El ayni es el principio de reciprocidad andino: "hoy por ti, mañana por mí". No es un intercambio utilitario (dar para recibir), sino una obligación cósmica de equilibrio. Un desarrollo reciente, el Método Multi-Neutrosófico Ayni, integra estos principios con lógica neutrosófica para guiar el desarrollo ético de IA.

El método usa Conjuntos Multi-Neutrosóficos (MNS) para modelar juicios éticos plurales, ambiguos y dependientes del contexto. Introduce una Medida de Consenso Multi-Neutrosófica Euclidiana que cuantifica el grado de alineación entre stakeholders sin borrar las diferencias epistémicas. El principio de ayni se entiende como negociación ética recíproca: la aplicación iterativa mejora tanto la alineación como la legitimidad.

Este es un ejemplo concreto de cómo la lógica andina puede formalizarse y operacionalizarse en IA, sin ser reducida a categorías occidentales.

5.3.5. Implicaciones para la IA

Una IA informada por el pensamiento andino tendría:

- Complementariedad sin superación: representar opuestos como necesarios el uno
- Cuatripartición: modelos con cuatro estados (afirmación, negación, combinación,
- Reciprocidad: intercambio equilibrado con las comunidades, no extracción.
- Inversión dialéctica: capacidad de voltear situaciones, hacer manifiesto lo latente.

5.4. Budismo y Nagarjuna: el tetralema y la vacuidad

5.4.1. El catuskoti (tetralema)

El catuskoti (tetralema) es la contribución lógica más radical del budismo madhyamaka. Nagarjuna, el filósofo budista del siglo II, lo usa sistemáticamente para analizar todo tipo de cuestiones: la causalidad, el yo, el nirvana, la existencia.

La forma básica es:

- A (existe, es verdadero)
- $\neg A$ (no existe, es falso)
- $A \wedge \neg A$ (ambos: existe y no existe)
- $\neg(A \vee \neg A)$ (ninguno: ni existe ni no existe)

es verdadera y las demás falsas. Nagarjuna a menudo niega las cuatro opciones, mostrando que todas comparten un presupuesto falso. En el *Mūlamadhyamakakārikā*, analiza el nirvana considerando las cuatro posibilidades y rechazándolas todas.

5.4.2. La vacuidad (sunyata)

La conclusión de Nagarjuna no es que "nada existe", sino que nada existe por sí mismo (*svabhava*). Todo es vacío (*sunyata*) de existencia inherente. Las cosas existen convencionalmente, en dependencia de otras cosas, pero no últimamente.

Esto tiene una consecuencia epistémica profunda: no hay una verdad última que pueda expresarse proposicionalmente. La verdad última (*paramartha-satya*) es inexpressable en términos de existencia/no-existencia. La verdad convencional (*samvrti-satya*) es válida en su nivel, pero no captura la realidad última.

El ejemplo clásico: uno ve una cuerda y la confunde con una serpiente. La "serpiente" ni existe ni no existe: no existe como serpiente, pero existe como cuerda. Es inefable (*anirvacaniya*). Este es el estado (4) del tetralema: ni existe ni no existe.

5.4.3. Los dos niveles de verdad

La distinción entre verdad convencional y última es crucial para entender el tetralema. En el nivel convencional, las proposiciones son verdaderas o falsas. En el nivel último, ninguna proposición captura la realidad, porque la realidad es vacía de esencia.

Esto no es escepticismo, es una terapia contra el dogmatismo. El tetralema no busca establecer una tesis, sino disolver las tesis, mostrando que todas las posiciones metafísicas (existencia, no-existencia, ambas, ninguna) son insostenibles cuando se examinan rigurosamente.

5.4.4. Formalización y límites

El tetralema ha sido objeto de intensos debates lógicos. ¿Cómo formalizarlo sin que colapse? Si usamos lógica clásica, (3) implica (1) y (2), y (4) equivale a (3) por De Morgan, con lo que las cuatro opciones colapsan. Las soluciones requieren lógicas no clásicas: paraconsistentes (que admiten (3)), intuicionistas (que rechazan la doble negación en (4)), o dialeteístas (que admiten que (3) es posible: hay proposiciones que son simultáneamente

verdaderas y falsas.

Hay un problema más profundo: el tetralema no es solo una teoría lógica; es también una herramienta terapéutica. Como estructura, puede formalizarse con lógicas no clásicas (paraconsistentes, intuicionistas, dialeteístas). Como práctica, no puede formalizarse completamente, porque su objetivo no es establecer reglas de inferencia, sino liberar la mente de las fijaciones conceptuales. Formalizar la estructura no es formalizar la práctica. Y creer que la estructura es la práctica puede traicionar su propósito.

5.4.5. Implicaciones para la IA

Una IA informada por el budismo madhyamaka tendría:

- Lógica paraconsistente o dialeteísta: admitir contradicciones sin trivialización.
- Reconocimiento
- Terapia conceptual: no solo resolver problemas, sino disolver presupuestos falsos.
- Dos niveles de verdad: distinguir entre modelos convencionales (útiles, pero no

5.5. Síntesis comparativa: qué comparten estas tradiciones

5.5.1. Relacionalidad vs. individualismo

Las cuatro tradiciones comparten una ontología relacional: las entidades no son átomos independientes, sino nodos de relaciones.

- Confucianismo: el sujeto se constituye en la relación (ren, li).
- Ubuntu: "I relate, I am"; las cosas son a través de otras.
- Andino: yanantin; los opuestos se necesitan mutuamente.
- Budismo: origen dependiente (*pratityasamutpada*); nada existe por sí mismo.

trata a los agentes como maximizadores aislados.

5.5.2. Contradicción productiva vs. no contradicción

Las cuatro tradiciones no excluyen la contradicción:

- Confucianismo: la armonía integra diferencias sin eliminarlas.
- Ubuntu: la cognición colectiva admite perspectivas múltiples y potencialmente
- Andino: el yanantin no resuelve los opuestos, los complementa.
- Budismo: el tetralema admite (3) y (4); la vacuidad trasciende la dicotomía

actual.

5.5.3. Conocimiento práctico y colectivo vs. proposicional e individual

Las cuatro tradiciones enfatizan:

- Conocimiento práctico: saber cómo, no solo saber que.
- Conocimiento colectivo: producido en comunidad, no por individuos aislados.
- Conocimiento contextual: dependiente de la situación, la relación, el momento.

5.5.4. Cuadro comparativo

Dimensión	IA actual (anglosajona)	Confucianismo	Ubuntu	Andino	Budismo	Sujeto	Individual	Relacional
Colectivo	Complementario	Vacío	Contradicción	Error	Armonía	Plural	Complemento	Tetralema
Conocimiento	Proposicional	Práctico	Comunal	Recíproco	Terapéutico	Validación	Métrica	Contexto
Experiencia	Relación	Instrumental	Constitutiva	Constitutiva	Cósmica	Dependiente		

5.6. Implicaciones para el diseño de la IA

5.6.1. Principios de diseño derivados

De estas tradiciones se derivan varios principios concretos:

- Pluralismo lógico: no una lógica, sino un ecosistema de lógicas (clásica,
- Representaciones relacionales: los conceptos como nodos en redes, no como
- Razonamiento por analogía y contexto: inferencia basada en similitud situada, no
- Conocimiento práctico: representación de procedimientos, rituales, prácticas, no
- Validación
- Sujeto colectivo: modelar comunidades epistémicas, no solo agentes individuales.
- Retorno a la comunidad: los sistemas intercambian con las comunidades, no solo
- Reconocimiento

5.6.2. Prototipos y experimentos

Hay algunos desarrollos concretos en progreso:

- Multi-Neutrosophic Ayni Method: integra lógica andina y neutrosófica para ética
- Wisdom Loop: marco ubuntu para diseño de IA en África.
- Global-MMLU: benchmark culturalmente anotado que reconoce la insuficiencia
- Lógicas paraconsistentes aplicadas: sistemas que toleran inconsistencias sin

epistemologías pueden operacionalizarse en IA.

5.6.3. Riesgos y cautelas

Incorporar estas tradiciones no está exento de riesgos:

- Apropiación: usar conceptos no occidentales como marketing sin transformar las
- Folclorización: tratar estas tradiciones como exotismo, no como teoría rigurosa.
- Simplificación: reducir la complejidad de estas tradiciones a eslóganes.
- Cooptación: que el poder use la diversidad para legitimar prácticas extractivas.

de una sola epistemología— es peor.

Bibliografía del capítulo 5

Ames, R. T. & Rosemont, H. (1998). *The Analects of Confucius: A Philosophical Translation*. Ballantine.

Hall, D. L. & Ames, R. T. (1987). *Thinking Through Confucius*. State University of New York Press.

Imafidon, E. (2023). Exploring the Theory of Communo-Cognition. En P. Ikhane & I. Ukpokolo (eds.), *African Epistemology: Essays on Being and Knowledge* (pp. 48–59). Routledge.

Lam, C. M. (2014). Confucian Rationalism. *Educational Philosophy and Theory*, 46(13), 1456–1468.

Leyva Vázquez, M. Y., Batista Hernandez, N., Callejas, E. A., & Smarandache, F. (2025). Multi- Neutrosophic Ayni Method Based on Ancestral Logic and N-Alectic Reasoning for Ethical AI and Sustainability. *Neutrosophic Sets and Systems*, 84.

Menkiti, I. A. (1979). Person and Community in African Traditional Thought. En R. A. Wright (ed.), *African Philosophy: An Introduction*. University Press of America.

Metz, T. (2007). Toward an African Moral Theory. *Journal of Political Philosophy*, 15(3), 321–341.

Mbiti, J. S. (1969). *African Religions and Philosophy*. Heinemann.

Nagarjuna. *Mūlamadhyamakakārikā* (trad. al español: *Versos fundamentales del Camino Medio*). Miraguano.

Priest, G. (2010). The Logic of the Catuṣkoti. *Comparative Philosophy*, 1(2), 24–54.

Ramose, M. B. (1999). *African Philosophy through Ubuntu*. Mond Books.

- Rosemont, H. (2015). *Against Individualism: A Confucian Rethinking of the Foundations of Morality, Politics, Family, and Religion*. Lexington Books.
- Sypniewski, B. P. (2001). Notes Comparing Aristotelian Reasoning with That of the Early Confucian School. *Journal of Chinese Philosophy*, 28(3), 257–274.
- Webb, H. S. (2012). *Yanantin and Masintin in the Andean World: Complementary Dualism in Modern Peru*. University of New Mexico Press.
- Wisdom in the Loop: Human-Centered AI and African Knowledge Systems in HCI. (2026). En *Proceedings of the 5th Biennial African Human Computer Interaction Conference*. ACM.

Lógicas organizacionales e interburocráticas

Análisis de cómo las organizaciones producen conocimiento, y qué implicaría para la IA

6.1. La organización como productora de conocimiento

6.1.1. Más allá del procesamiento de información

La visión dominante en IA y en gran parte de la teoría organizacional trata a las organizaciones como sistemas de procesamiento de información. Reciben *inputs*, los procesan según determinadas reglas, producen *outputs*. Las personas son "recursos humanos", los procesos son "flujos de trabajo", el conocimiento es "información almacenada".

Nonaka y Takeuchi desafiaron esta visión con una tesis simple pero radical. Afirmaron que las organizaciones no solo procesan conocimiento, lo crean. La creación de conocimiento organizacional no es la suma de conocimientos individuales, sino un proceso colectivo y dinámico que ocurre en la interacción entre personas, prácticas y contextos.

Esta distinción tiene consecuencias profundas. Si las organizaciones solo procesan información, entonces la IA puede "reemplazarlas" u "optimizarlas" sin pérdida. Si las organizaciones crean conocimiento, entonces la IA que solo procesa información es estructuralmente incapaz de participar en esa creación sin destruirla.

6.1.2. El conocimiento como activo estratégico

El conocimiento organizacional tiene características que lo distinguen radicalmente de la información procesable:

- Transferibilidad limitada: el conocimiento tácito es difícil, lento y costoso de
- Imitabilidad limitada: el conocimiento colectivo se construye sobre una historia
- Sustituibilidad nula: la sinergia de un grupo específico no puede replicarse por
- Apropiabilidad: el valor del conocimiento depende de quién lo posee y en qué

pueda extraerse, almacenarse y procesarse. Es una capacidad relacional que existe en las prácticas, las relaciones y los contextos de una organización específica.

6.2. El conocimiento tácito y la espiral de conversión

6.2.1. El conocimiento tácito

Michael Polanyi estableció la distinción fundamental, sabemos más de lo que podemos decir. El conocimiento tácito es aquel que una persona o comunidad posee incorporado, almacenado en su mente, su cultura, su cuerpo y que es difícil o imposible de explicar completamente.

El conocimiento tácito incluye:

- Ideas, experiencias, destrezas, habilidades.
- Costumbres, valores, historia, creencias.
- Conocimiento del contexto (geografía, normas no escritas, comportamientos).
- Destrezas cognitivas (comprensión, resolución de problemas, visualización).

común, capacidad de hacer algo sin poder explicarlo. Brown y Duguid lo expresaron con precisión: "en los individuos, el conocimiento tácito significa intuición, el juicio, el sentido común, la capacidad de hacer algo sin necesariamente poder explicarlo".

6.2.2. La espiral del conocimiento

Nonaka y Takeuchi modelaron la creación de conocimiento organizacional como una espiral que atraviesa cuatro modos de conversión:

- Socialización (tácito → tácito): adquirir conocimiento tácito a través de compartir
- Exteriorización (tácito → explícito): convertir conocimiento tácito en conceptos
- Combinación (explícito → explícito): sistematizar conceptos en un sistema de
- Interiorización (explícito)

a lo tácito, ampliando continuamente el conocimiento organizacional.

6.2.3. Lo que la IA no puede capturar

La IA actual puede procesar conocimiento explícito: textos, fórmulas, procedimientos, pero no puede producir conocimiento tácito ni participar en el proceso de socialización que lo genera.

La socialización requiere copresencia, experiencia compartida, imitación corporal. No es transmisión de información, es formación de hábitos a través de la práctica conjunta. Un modelo de lenguaje puede describir cómo se hace algo, pero no puede hacerlo con alguien.

La exteriorización requiere metáfora, analogía, lenguaje figurativo. Es un proceso creativo que no sigue reglas. La IA puede generar metáforas combinatorias, pero no puede vivir la experiencia tácita que la metáfora intenta expresar.

La interiorización requiere práctica corporal, repetición, incorporación. La IA no tiene cuerpo, no tiene hábitos, no puede "aprender haciendo" en el sentido que Nonaka y Takeuchi describen.

6.3. La lógica del actor estratégico

6.3.1. Más allá del actor racional

La teoría del actor racional, como vimos en el capítulo 2, presupone un agente que optimiza según preferencias estables. Crozier y Friedberg, en su análisis estratégico de las organizaciones, ofrecen una alternativa radical.

El actor organizacional no es un optimizador, es un estratega. Su comportamiento no se explica por preferencias estables, sino por la búsqueda de poder en un contexto de incertidumbre. La rationalité de l'acteur n'est pas dans ses objectifs, mais dans le contexte organisationnel où il agit .

El actor tiene un margen de libertad irreducible que no puede dejar de usar. Incluso en el sistema más rígido, el actor calcula, se adapta, juega. No es "una mano" (taylorismo) ni "una mano y un corazón" (humanismo). Es una cabeza, un inventor, un calculador.

6.3.2. El poder como relación

En el análisis estratégico, el poder no es una propiedad que se posee, sino una relación que se ejerce. El poder de un actor sobre otro es mayor cuanto más el otro depende de él para lograr sus objetivos, y cuanto mayor es la incertidumbre sobre si el actor mostrará el comportamiento deseado.

Las fuentes de poder son las zonas de incertidumbre que el actor controla:

- Maîtrise d'une compétence particulière: control de una habilidad que otros
- Maîtrise de la communication et des informations: control de flujos de
- Relation de l'organisation à son environnement: control de las relaciones con el

superior, porque puede negarse a hacer lo que se le pide. Si no puede negarse, no es un actor: es "una cosa".

6.3.3. El sistema de acción concreto

Crozier y Friedberg definen el sistema de acción concreto como "un conjunto humano estructurado que coordina las acciones de sus participantes mediante mecanismos de juego relativamente estables, y que mantiene su estructura —la estabilidad de sus juegos y las relaciones entre ellos— mediante mecanismos de regulación que

constituyen otros juegos".

El sistema de acción no es un conjunto de individuos con objetivos comunes. Es un juego donde cada actor persigue sus propios fines, que pueden ser diferentes y cambiantes. La solución de un problema común puede ser el objetivo de algunos, pero no necesariamente de todos. Algunos actores pueden estar interesados en sabotear el proyecto, en ganar puntos para una promoción, o simplemente en ganar su salario con el mínimo esfuerzo.

6.3.4. Lo que la IA no puede modelar

El actor estratégico de Crozier y Friedberg es irreducible al agente racional de la IA por varias razones:

- Rationalité limitée, pas optimisation: el actor no optimiza porque no puede. Su
- Poder, pas préférences: lo que mueve al actor no son preferencias estables, sino
- Jeu, pas planification: la acción organizacional no es planificada, es jugada. Los
- Conflit, pas consensus: el conflicto, la manipulación y la ambigüedad

puede representar este mundo. Puede simularlo superficialmente, pero no capturar su lógica.

6.4. La traducción interburocrática

6.4.1. El trabajo de traducción

Un desarrollo reciente en el estudio de las burocracias introduce el concepto de traducción burocrática: el trabajo cognitivo que los ciudadanos (y los propios burócratas) realizan para hacer sentido de la información burocrática.

La traducción burocrática toma cinco formas:

- Buscar claridad y coherencia en la información burocrática.
- Resolver contradicciones entre fuentes.
- Corroborar y sintetizar información de diversas fuentes confiables.
- Transportar información entre organismos públicos, a veces a través de fronteras
- Derivar conocimiento útil de la experiencia.

información es "solo información", accesible y suficiente. Pero los ciudadanos experimentan la información burocrática como conocimiento especializado que requiere esfuerzo cognitivo para apropiarse y actuar sobre ella.

6.4.2. La política de la traducción

La traducción no es un proceso técnico neutral, es político. Los actores intermediarios — científicos, extensionistas, burócratas provinciales— traducen conocimiento hacia arriba, hacia abajo y horizontalmente en sistemas de gobernanza multinivel.

La traducción no es neutral, refleja intereses, reproduce jerarquías, abre o cierra posibilidades. Los actores intermediarios tienen agencia para moldear la trayectoria del cambio socio-ecológico, pero esa agencia está limitada por las estructuras de poder en que operan.

6.4.3. Lo que la IA no puede modelar

La traducción interburocrática es irreducible a la IA actual por varias razones:

- Contradicciones: los ciudadanos deben resolver contradicciones entre fuentes de
- Contexto: la traducción depende del contexto (institucional, cultural, lingüístico).
- Poder: la traducción es negociación implícita de significado en interacciones.
- Invisibilidad: el trabajo de traducción es invisible para las instituciones. La IA, que

6.5. Implicaciones para el diseño de IA

6.5.1. Del agente al actor

La primera implicación es un cambio de concepto de agencia. La IA actual modela agentes (optimizadores con preferencias fijas). Una IA que incorpore las lógicas organizacionales necesitaría modelar actores (estrategas en relaciones de poder).

Esto implica:

- Racionalidad limitada: no optimización, sino satisfacción en contextos de
- Poder como relación: no preferencias, sino dependencias mutuas.
- Juego: no planificación, sino adaptación estratégica a lo que hacen los otros.
- Conflicto: no consenso, sino negociación permanente.

6.5.2. Del conocimiento explícito al tácito

La segunda implicación es un cambio de concepto de conocimiento. La IA actual solo procesa conocimiento explícito. Una IA que incorpore el conocimiento organizacional necesitaría:

- Reconocer el conocimiento tácito: no como "información faltante", sino
- Participar en la socialización: no como transmisora de información, sino
- Apoyar

6.5.3. De la traducción automática a la traducción política

La tercera implicación es un cambio de concepto de traducción. La IA actual traduce lenguas. Una IA que incorpore las lógicas interburocráticas necesitaría traducir significados en contextos de poder:

- Resolver contradicciones entre fuentes sin colapsar.
- Reconocer el contexto como constitutivo del significado.
- Visibilizar el trabajo de traducción que hoy es invisible.
- Negociar significados con los actores, no imponerlos.

6.5.4. Prototipos y líneas de investigación

Algunas líneas relevantes son:

- Sistemas de argumentación dialéctica para modelar conflictos y negociaciones
- Lógicas paraconsistentes para manejar contradicciones sin trivialización.
- Modelado de sistemas de acción concreto en simulación social.
- IA explicable dialéctica que reconozca tensiones y perspectivas.
- Diseño participativo que incorpore a las comunidades en la definición de lo que

Bibliografía del capítulo 6

Brown, J. S. & Duguid, P. (2000). *The Social Life of Information*. Harvard Business School Press.

Crozier, M. & Friedberg, E. (1977). *L'acteur et le système: les contraintes de l'action collective*. Éditions du Seuil.

Mökander, J. & Schroeder, R. (2024). Artificial Intelligence as Weberian Rationalization. *Social Science Computer Review*.

Nonaka, I. & Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press.

Polanyi, M. (1966). *The Tacit Dimension*. Doubleday.

Weger, J. (2023). "Nothing for Free": Intermediary actors and cross-scalar knowledge translation for climate adaptation in the Mekong Delta. *Journal of Political Ecology*, 30(1).

Weger, J. (2025). Making Sense of Bureaucratic Information: Conceptualising Bureaucratic Translation. *Social Policy and Society*.

Pensamiento grupal y epistemología colectiva

Estudio de las dinámicas de conocimiento en grupos y su modelado en la IA

7.1. El pensamiento grupal como patología epistémica

7.1.1. Qué es el pensamiento grupal

Irving Janis acuñó el término *groupthink* en 1972 para describir el proceso por el cual un grupo cohesionado toma decisiones defectuosas porque la presión hacia la unanimidad suprime la evaluación crítica de alternativas. La definición original de Janis es precisa: “un modo de pensamiento que las personas adoptan cuando están profundamente involucradas en un grupo cohesivo, cuando los esfuerzos de los miembros por unanimidad hacen caso omiso de su motivación para valorar realistamente cursos de acción alternativos”.

El pensamiento grupal no es un simple error de cálculo. Es una patología epistémica: un modo de conocer que sistemáticamente distorsiona la evaluación de evidencia, la consideración de alternativas y la estimación de riesgos. Los grupos afectados no solo se equivocan: producen conocimiento defectuoso como resultado de su propia dinámica social.

7.1.2. Los síntomas

Janis identificó ocho síntomas del pensamiento grupal:

- Ilusión de invulnerabilidad: el grupo se cree incapaz de error.
- Creencia incuestionable en la moralidad inherente: el grupo cree en la rectitud de
- Racionalización colectiva: el grupo construye justificaciones para ignorar
- Visión estereotipada de los externos: los oponentes son vistos como demasiado
- Autocensura: los miembros evitan expresar dudas por miedo a parecer desleales.
- Ilusión de unanimidad: se percibe consenso donde no lo hay.
- Presión directa a los disidentes: se presiona a quien se desvía del consenso.
- Guardianes de la mente: algunos miembros filtran información adversa para

incompleto de alternativas, estudio incompleto de objetivos, fracaso en examinar riesgos, fallo en reevaluar alternativas rechazadas, búsqueda pobre de información, procesamiento sesgado de información y ausencia de planes de contingencia.

7.1.3. Las condiciones estructurales

El pensamiento grupal no surge espontáneamente, sino que requiere condiciones estructurales y situacionales específicas para poder florecer:

- Alta cohesión: necesaria pero no suficiente. La cohesión saludable puede ser
- Defectos estructurales: aislamiento del grupo de fuentes externas, liderazgo
- Contexto situacional provocativo: estrés alto por amenazas externas, presión

Kennedy, cuando un grupo altamente cohesionado, liderado por un presidente carismático, bajo presión temporal, ignoró evidencia contraria y tomó una decisión catastrófica.

7.2. La epistemología social y sus límites

7.2.1. Goldman: la epistemología social como evaluación de prácticas

Alvin Goldman propuso una epistemología social que evalúa las prácticas sociales de producción de conocimiento según su contribución a la verdad. Su marco de análisis incluye tres aportes:

- Testimonio: la transmisión de conocimiento entre agentes.
- Autoridad epistémica: quién es reconocido como experto.
- Sistemas de evaluación: cómo las comunidades validan creencias.

como un contexto tenue, transparente y neutro que apenas permea la autonomía cognitiva de los individuos. Los sujetos aparecen como “esferas independientes” que “no tienen más remedio que relacionarse entre sí”. La dimensión social no es constitutiva del conocimiento, apenas es un añadido, incluso con cierto grado de “molestia”.

7.2.2. Kitcher: la división del trabajo cognitivo

Philip Kitcher desarrolló una versión más robusta de la epistemología social. Su propuesta se centra en la organización de las comunidades de conocimiento y los procesos que ocurren en el marco comunal, con el objetivo de promover la producción de creencias verdaderas.

Kitcher enfatiza la división del trabajo cognitivo entre comunidades científicas, diferentes individuos y grupos que se especializan en diferentes áreas. La verdad emerge de la interacción entre ellos. A pesar de ello su enfoque sigue siendo agregativo, a la manera de Goldman: la dimensión social es resultado de “sumar estados cognitivos de individuos cuya simple agregación configuraría la mentada dimensión social”.

7.2.3. Los límites de la epistemología social “mínima”

Tanto Goldman como Kitcher operan con una concepción débil de lo social. Un análisis crítico sostiene que esto lleva a que “lo social visto bajo este enfoque rebajado no es más que un tipo de procesos estereotipados que incluyen en su trama a los sujetos”.

La alternativa es una epistemología social constitutiva, donde lo social no es contexto sino condición de posibilidad del conocimiento. El conocimiento no se produce por individuos que luego se relacionan: se produce en relaciones que constituyen a los individuos como cognoscentes.

7.2.4. Fuller y la inversión

Steve Fuller ha planteado invertir la imagen tradicional del conocimiento. En lugar de concebir el conocimiento como una producción individual que luego se transforma en conocimiento colectivo a través de criterios sociales, Fuller sostiene que el conocimiento debería considerarse en primer lugar como un fenómeno colectivo, y solo derivadamente como un fenómeno individualmente localizado.

Esta inversión tiene consecuencias profundas para la IA. Si el conocimiento es primariamente colectivo, entonces una IA que solo modela agentes individuales es estructuralmente incapaz de capturar la dimensión epistémica fundamental.

7.3. Las dinámicas de grupo en sistemas multiagente de la IA

7.3.1. La reproducción del pensamiento grupal

La investigación reciente sobre sistemas multiagente de LLM ha documentado un patrón inquietante: los equipos de IA reproducen el pensamiento grupal humano.

Un estudio de Anthropic (2026), citado en un análisis de la industria, probó cuatro agentes que debatían y votaban sobre una decisión. La evidencia compartida apoyaba la respuesta incorrecta; solo un agente tenía el dato que importaba. Con la mayoría de los modelos, el grupo llegó a la respuesta mayoritaria incorrecta entre el 64% y el 83% del tiempo. Un agente solo, con todos los datos, acertaba casi siempre.

Otro estudio de ICML 2025 encontró que los equipos de LLM consistentemente fallan en igualar el rendimiento de su miembro experto, incluso cuando se les dice explícitamente quién es el experto. Las pérdidas de rendimiento alcanzan hasta 41.1% en *benchmarks* de machine learning.

7.3.2. El mecanismo: compromiso integrador

El estudio de ICML 2025 identificó el mecanismo: compromiso integrador. Los agentes tienden a promediar las opiniones del experto y los no expertos en lugar de ponderar apropiadamente la expertise. Este comportamiento aumenta con el tamaño del equipo y correlaciona negativamente con el rendimiento.

Es exactamente el patrón del pensamiento grupal humano: la búsqueda de consenso suprime la expertise. Pero hay una diferencia crucial: en los grupos humanos, la expertise puede ser reconocida y respetada. En los equipos de LLM, el reconocimiento de expertise no se traduce en su utilización.

7.3.3. La paradoja de la robustez

Hay un hallazgo contraintuitivo: el mismo comportamiento de búsqueda de consenso que perjudica la utilización de expertise mejora la robustez frente a los agentes adversariales. Los grupos que promedian opiniones son más resistentes a un miembro que intenta sabotear la respuesta.

Esto revela un trade-off fundamental: la alineación y la utilización efectiva de expertise están en tensión. Un sistema que pondera fuertemente la expertise es vulnerable al engaño mientras que un sistema que promedia es robusto, pero mediocre.

7.3.4. El tamaño del grupo y la no-linealidad

Un estudio publicado en PNAS encontró que el tamaño del grupo afecta las dinámicas de manera no lineal, revelando “regímenes dinámicos dependientes del modelo”. Por encima de un tamaño crítico de población, las simulaciones convergen a predicciones deterministas que exponen “cuencas de atracción de equilibrios en competencia”.

Esto significa que añadir más agentes no es neutral: cambia cualitativamente la dinámica. Un sistema con diez agentes puede comportarse de manera radicalmente distinta que uno con cien, no solo cuantitativamente sino cualitativamente.

7.3.5. La contaminación de memoria y las cascadas

Los sistemas multiagente también sufren de cascadas de pensamiento grupal y contaminación de memoria. Un análisis técnico describe cómo una idea incorrecta puede propagarse a través de un enjambre de agentes sin corroboración independiente, “invisibilizando” su origen y costando tokens ilimitados sin que nadie lo note.

La solución propuesta es arquitectónica: cuarentena de afirmaciones hasta que N escritores independientes las corroboren, detector de cascadas (ráfagas de afirmaciones semánticamente similares de muchos agentes en poco tiempo) y capacidad de rebobinar el estado del enjambre a un checkpoint anterior. El límite reconocido es significativo: “una capa de memoria no puede hacer que un modelo piense de manera diferente. Lo que garantiza es que una idea incorrecta no pueda convertirse en la verdad compartida del enjambre sin corroboración independiente”.

7.4. Mecanismos de corrección

7.4.1. La disidencia estructurada

La respuesta clásica al pensamiento grupal es la disidencia estructurada: asignar a alguien el rol de “abogado del diablo”, crear normas que requieran procedimientos metódicos de deliberación, invitar a expertos externos.

En sistemas multiagente, esto se traduce en un agente cuyo único trabajo sea argumentar contra el consenso. Pero hay un problema: si el disidente es un agente más del mismo tipo, ¿realmente aporta diversidad o solo reproduce la misma lógica con otro rol?

7.4.2. La diversidad de personas

Una tesis de la Universidad de Padua investigó si asignar personas diversas a agentes LLM puede generar “inteligencia colectiva” en el sentido de la “sabiduría de las multitudes”. El hallazgo preliminar fue que la diversidad de personas puede impulsar capacidades de resolución de problemas, pero el efecto depende de la estructura de la red de comunicación.

Esto sugiere que la diversidad debe ser diseñada, no asumida como una consecuencia necesaria. Agentes idénticos con roles diferentes no son diversos; agentes diversos con la misma arquitectura pueden serlo.

7.4.3. La esparcidad de red

La misma investigación encontró que la esparcidad de red (no todos los agentes conectados a todos) mejora el rendimiento en tareas de resolución de problemas. Esto contrasta con la intuición de que más comunicación es mejor. La explicación consiste en que la comunicación densa homogeneiza demasiado rápido, suprimiendo la diversidad antes de que el resultado pueda constituir una contribución.

7.4.4. El reequilibrio de la comunicación

Un estudio experimental sobre predicción colectiva encontró que algoritmos de reequilibrio de red —que cambian las conexiones entre los participantes— pueden mejorar la precisión colectiva, aunque el efecto depende de cuál sea la estructura inicial.

El algoritmo llamado “polarize” (que conecta a personas con opiniones divergentes) fue el que más benefició la precisión post-comunicación, sugiriendo que el conflicto estructurado mejora la cognición colectiva.

7.4.5. Los límites de los mecanismos técnicos

Todos estos mecanismos tienen en común un límite: operan dentro de la arquitectura epistémica original de la IA. Un agente disidente sigue siendo un LLM; una persona diversa sigue siendo una construcción de *prompt*; una red esparcida sigue siendo una topología de comunicación.

Los mecanismos técnicos pueden mitigar el pensamiento grupal, pero no transformar la epistemología subyacente. Para eso se necesita algo más, se necesita una concepción diferente del conocimiento que reconozca su carácter constitutivamente colectivo. Y que tal concepción pueda ser operacionalizada.

7.5. Implicaciones para el diseño de la IA

7.5.1. Del individuo al colectivo

La primera implicación es un cambio de la unidad de análisis. La IA actual modela agentes individuales. Una IA que incorpore la epistemología colectiva necesitaría modelar grupos como unidades epistémicas con propiedades emergentes irreducibles a los individuos.

Esto no significa abandonar el modelado individual, significa reconocer que lo colectivo tiene propiedades propias. La sinergia de un equipo no se reduce a la suma de las capacidades individuales ya que el pensamiento grupal no se reduce a la suma de los sesgos individuales. Nuevamente, el todo es más que la suma de las partes.

7.5.2. La diversidad como condición, no como opción

La segunda implicación es que la diversidad epistémica debe ser una condición estructural del diseño, no un añadido cosmético ad hoc. Agentes idénticos no producen sabiduría colectiva, producen consenso.

Pero la diversidad no puede ser superficial (roles diferentes con la misma lógica subyacente). Tiene que ser una diversidad epistémica con diferentes formas de conocer, diferentes lógicas, diferentes tradiciones. Esto conecta con el previo capítulo 5: incorporar lógicas no occidentales no es tanto una cuestión de justicia como una condición para la cognición colectiva robusta.

7.5.3. El conflicto como recurso

La tercera implicación es que el conflicto debe ser tratado como recurso epistémico, no como un ruido que debe ser eliminado. Los algoritmos que fuerzan el consenso prematuro, apurando el resultado, suprimen la diversidad mientras que los algoritmos que estructuran el disenso mejoran la precisión.

Esto requiere mecanismos que mantengan la tensión entre distintas perspectivas sin resolverlas con sentido de urgencia. Las lógicas paraconsistentes (capítulo 3), las representaciones de contradicciones (capítulo 5) y las arquitecturas que preservan la diversidad (capítulo 6) cobran mayor relevancia en este contexto.

7.5.4. La trazabilidad de la procedencia

La cuarta implicación es la necesidad de trazabilidad de la procedencia (*provenance*). En los sistemas multiagente es crucial saber de dónde proviene una afirmación, cuántos agentes independientes la corroboran y definir si hay un origen común contaminado.

Sin trazabilidad, las cascadas de desinformación son invisibles. En cambio, cuando hay trazabilidad, se pueden detectar y contener.

7.5.5. La corrección como proceso continuo

La quinta implicación es que la corrección no puede ser un estado estable (una decisión final correcta), sino un proceso continuo. La capacidad de rebobinar (*rewind*) y retornar a un checkpoint anterior, poder invalidar un origen contaminado y propagar la invalidación a los derivados dependientes es un factor esencial para toda inteligencia. La verdad siempre es una verdad en proceso, no un logro establecido por fuera del tiempo, sino una conquista que se renueva.

Esto es análogo a la revisión de creencias en la epistemología formal (AGM), pero aplicado a sistemas multiagente. Porque las creencias no se eliminan, se invalidan con trazabilidad.

Bibliografía del capítulo 7

Anthropy (2026). Patterns and problems in emerging multiagent systems. Anthropic Research.

Flint Ashery, A., Aiello, L. M., Pastor-Satorras, R., & Baronchelli, A. (2026). Group size effects and collective misalignment in LLM multi-agent systems. PNAS, 123(34), e2531697123.

Fuller, S. (1988). Social Epistemology. Indiana University Press.

Goldman, A. (1999). Knowledge in a Social World. Oxford University Press.

Janis, I. (1972). Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes. Houghton Mifflin.

Kitcher, P. (1994). Contrasting Conceptions of Social Epistemology. En F. Schmitt (ed.), Socializing Epistemology. Rowman & Littlefield.

Leyva Vázquez, M. Y., Batista Hernandez, N., Callejas, E. A., & Smarandache, F. (2025). Multi- Neutrosophic Ayni Method Based on Ancestral Logic and N-Alethic Reasoning for Ethical AI and Sustainability. Neutrosophic Sets and Systems, 84.

Pappu, A., El, B., Cao, H., di Nolfo, C., Sun, Y., Cao, M., & Zou, J. (2025). Multi-Agent Teams Hold Experts Back. ICML 2025.

Hacia una IA dialéctica y epistémicamente abierta

Principios de diseño y arquitecturas posibles

8.1. Siete principios de diseño para una IA dialéctica y pluricultural

8.1.1. Principio de pluralismo lógico

- Enunciado: ningún sistema único de lógica captura la totalidad de las formas
- Fundamento: el capítulo 3 mostró que la lógica formal excluye la contradicción, el
- Implicación técnica: en lugar de una lógica monolítica, estas tradiciones dan paso

que aquí se proponen formalizan la parte formalizable de la dialéctica (contradicción, negación lógica, proceso), no la parte irreductible (praxis, historicidad, negación real). Como se mostró en el capítulo 3, la dialéctica tiene aspectos que resisten la formalización. Lo que estas arquitecturas hacen es operar con representaciones más ricas del movimiento dialéctico, no participar en el movimiento mismo. La distinción es importante: una IA que opera con lógicas no clásicas no es una IA dialéctica, es una IA que representa mejor la dialéctica.

8.1.2. Principio de representaciones relacionales

- Enunciado: los conceptos no deben representarse como entidades aisladas
- Fundamento: el capítulo 5 mostró que el confucianismo, el ubuntu, el andino y el
- Implicación técnica: representaciones que codifiquen explícitamente las relaciones

8.1.3. Principio de conocimiento tácito y corporizado

- Enunciado: el conocimiento no es solo proposicional. Un sistema
- Fundamento: el capítulo 6 mostró que el conocimiento tácito (Polanyi, Nonaka)
- Implicación técnica: esto es lo más difícil de operacionalizar. La IA actual no tiene

8.1.4. Principio de conflicto como recurso epistémico

- Enunciado: el conflicto, la contradicción y el disenso no son ruidos a eliminar,
- Fundamento: el capítulo 7 mostró que los equipos de LLM reproducen el
- Implicación técnica: arquitecturas que mantengan la tensión entre las perspectivas

8.1.5. Principio de la trazabilidad de la procedencia

- Enunciado: en los sistemas multiagente, es esencial rastrear el origen de cada
- Fundamento: el capítulo 7 describió cómo las cascadas de desinformación en
- Implicación técnica: mecanismos de cuarentena de afirmaciones, detección de

8.1.6. Principio de validación múltiple

- Enunciado: la validación del conocimiento no puede reducirse a métricas únicas
- Fundamento: el capítulo 2 mostró que las métricas formales son el criterio único
- Implicación técnica: sistemas de evaluación multi-criterio que no colapsen a un

8.1.7. Principio de soberanía epistémica y participación

- Enunciado: las comunidades cuyos datos alimentan un sistema deben
- Fundamento: el capítulo 4 mostró que la colonialidad del dato extrae
- Implicación técnica: algunos marcos como los Principios CARE (Collective

8.2. Arquitecturas posibles

8.2.1. Neuro-simbólica con lógicas no clásicas

La integración más avanzada de lógicas no clásicas con LLMs es la interpretación fundamentada en LLM (LLM-grounded interpretation) de Ferguson y colaboradores. En lugar de usar LLMs *junto a* razonadores (*prompting, solver-based, fine-tuning*), esta aproximación integra el LLM directamente en la semántica formal de la lógica.

En este mecanismo elegante un LLM “juez” asigna valores de verdad generalizados (elementos del *bilattice* NIN) a fórmulas atómicas traducidas a lenguaje natural. Estos valores alimentan un razonador con tableaux analíticos que preserva la solidez y la completitud. La lógica utilizada es la ACrQ (Analytic Containment with restricted Quantification), una lógica paraconsistente que admite contradicción sin trivialización.

Esto no es una especulación, hay evidencia empírica de una implementación práctica usando *benchmarks* de factualidad. El sistema demuestra que es posible tener garantías formales (solidez, completitud) en un razonador que incorpora LLMs, algo que los enfoques puramente neuronales no ofrecen.

8.2.2. Sistemas multiagente con debate estructurado

Los marcos de debate multiagente ofrecen una arquitectura para el conflicto productivo. ARTEMIS es un ejemplo desarrollado:

- Generación jerárquica de argumentos: los agentes seleccionan estrategias de alto
- Evaluación con razonamiento causal: se construye un grafo causal del debate que
- Representación del estado del debate: existe un hipergrafo dirigido y atribuido que
- Mecanismo de jurado: múltiples LLMs evalúan los argumentos con ponderación

iterativamente las salidas de los demás. La clave es que la interacción no busca consenso rápido, sino refinamiento progresivo a través de rondas de crítica y respuesta.

8.2.3. Computación relacional y modelos de mundo sociales

El marco Social NeuroAI propone reformular los modelos de mundo desde una perspectiva relacional. En lugar de agentes que modelan entornos de manera individualista, aquí contamos con agentes que modelan otros agentes como sistemas dinámicos con sus propios modelos latentes. La inteligencia emerge de dinámicas corporizadas acopladas, no de la inferencia aislada.

Esto implica un cambio arquitectónico. De simuladores solitarios hacia sistemas socialmente embebidos cuya inteligencia se construye a través de la coordinación, la alineación y la representación compartida. La coordinación metaestable y la toma de decisiones colectiva emergen del acoplamiento neural estructurado, no de la agregación de outputs individuales.

8.2.4. Prácticas de diseño participativo y descolonizador

Las arquitecturas técnicas no son suficientes. El Marco DIA propone integrar la descolonización epistémica en todas las fases del ciclo de diseño:

- Valores primero: partir de valores comunitarios, no de problemas técnicos.
- Co-diseño participativo: los usuarios son co-creadores, no “usuarios finales”.
- Datos como confianza: consentimiento contextual, no extracción.
- Interpretación colectiva: evaluación en procesos de *sensemaking* comunitario.
- Retorno y feedback: los sistemas devuelven valor a las comunidades.

ecuatoriana documenta la existencia de tensiones concretas entre las racionalidades hegemónicas de IA y los sistemas locales de conocimiento. La lección aprendida es que la incorporación de otras epistemologías no puede hacerse desde fuera, sino que requiere la participación real de las comunidades en el diseño y la gobernanza.

8.3. Operacionalización concreta: de principios a mecanismos

8.3.1. Selección dinámica de lógicas

Un sistema de pluralismo lógico necesita mecanismos para seleccionar qué lógica es apropiada para usar en cada contexto. Algunas posibles aproximaciones son:

- Meta-razonamiento: un módulo que evalúa el tipo de problema y selecciona la
- Lógicas combinadas: sistemas que integran múltiples lógicas con traducciones
- Aprendizaje de selección: entrenar un clasificador que prediga qué lógica es más

sesgarse. La solución no es buscar la eliminación del sesgo, lo que sería un emprendimiento imposible, sino hacerlo visible, trazable y sujeto a revisión.

8.3.2. Representación de contradicciones sin colapso

Los sistemas de Curry ofrecen un marco algebraico para representar la contradicción, la indeterminación y el razonamiento no-bivalente. Los valores de verdad generalizados (*bilattices*) (Arieli y Avron, 1998) permiten distinguir entre:

- Verdadero y falso simultáneamente (dialeiteia).
- Ni verdadero ni falso (indeterminación).
- Verdadero en un aspecto, falso en otro (contradicción contextual).

ellos puede mantener la contradicción sin derivar de ella argumentaciones triviales o inconsistentes.

8.3.3. Mecanismos de disidencia y diversidad

Para contrarrestar el pensamiento grupal puede apelarse a diferentes mecanismos:

- Agentes adversariales: al menos un agente cuyo rol sea argumentar contra el
- Comunicación esparcida: topologías de red que eviten la homogeneización rápida.
- Reequilibrio de conexiones: algoritmos que conectan perspectivas divergentes.
- Diversidad epistémica: agentes con lógicas diferentes, no solo roles diferentes.

opiniones divergentes (“polarize”) mejora la precisión post-comunicación. La lección para la IA es que el conflicto estructurado mejora la cognición colectiva.

8.3.4. Trazabilidad y corrección

Mecanismos para los sistemas multiagente:

- Procedencia de afirmaciones: cada output lleva metadatos sobre su origen y la
- Detección de cascadas: identificar las ráfagas de afirmaciones semánticamente
- Cuarentena: aislar afirmaciones hasta tener una corroboración independiente.
- Rebobinado: capacidad de regresar a un *checkpoint* anterior e invalidar un origen

piense de manera diferente”. La trazabilidad contiene el daño, pero no transforma la epistemología subyacente.

8.3.5. Validación multi-criterio

En lugar de colapsar todo a una métrica escalar es recomendable utilizar:

- Criterios múltiples: coherencia lógica, consenso comunal, reciprocidad,
- Ponderación contextual: los pesos relativos dependen del dominio y la
- Transparencia: los criterios y sus pesos son visibles y discutibles.

entre tipos de evidencia y grados de soporte.

8.4. Riesgos y cautelas

8.4.1. La complejidad computacional

Integrar lógicas múltiples, mantener las contradicciones, rastrear la procedencia y validar el multi-criterio tiene un costo computacional. Los sistemas actuales ya son de por sí enormes, susceptibles de escalamientos cada vez mayores. Añadir estas capas de razonamiento no clásico que propone puede hacerlos físicamente inviables.

Una advertencia de cautela puede ser aconsejable. Porque no todo sistema necesita incluir toda la complejidad que se propone. La arquitectura puede y debe ser modular: los módulos dialécticos serían opcionales y eventuales. Se activan cuando el contexto lo requiere.

Esta modularidad no es una concesión, sino una condición de viabilidad. El pluralismo lógico no significa usar todas las lógicas siempre, ya que las lógicas no son disyuntivas sino complementarias. Significa tenerlas disponibles cuando el contexto lo requiere.

8.4.2. Riesgo de relativismo

El pluralismo lógico y epistémico puede degenerar en relativismo indiscriminado: “todo vale porque hay múltiples lógicas”, lo cual es un malentendido. El pluralismo sí sostiene que hay múltiples criterios de validez, pero no sostiene que no haya ningún criterio o que cualquier criterio sea tan válido como otro.

Cautela: es preciso mantener criterios de traducción entre lógicas y de evaluación de su adecuación contextual. La pluralidad no, aunque conveniente e incluso necesaria, no exime de la exigencia de rigor epistemológico.

8.4.3. Cooptación por el poder

La diversidad epistémica puede ser cooptada como marketing cuando las empresas añaden “capas de diversidad” sin transformar sus prácticas extractivas. El marco DIA advierte que la descolonización real requiere una redistribución del poder, no alcanza solo con una inclusión simbólica que no afecte la concepción monolítica del conocimiento.

Cautela: la participación comunitaria debe ser real, con poder de decisión, no solo consultiva.

8.4.4. Dificultad de evaluación

Es difícil evaluar si una IA dialéctica y pluricultural “funciona” porque los criterios de aplicación son múltiples y contextuales. Las métricas estándar (accuracy, BLEU) no suelen capturar los datos que realmente importan para poder evaluar la consistencia dialéctica y pluricultural.

Cautela: desarrollar métricas alternativas que midan coherencia dialéctica, diversidad de perspectivas, trazabilidad y adecuación cultural. Esto es un campo de investigación aún abierto que merece ser trabajado.

8.4.5. El riesgo de la “falsa dialéctica”

El capítulo 3 advirtió sobre la “teleología proxy” en la cual los LLM pueden simular la realización de movimientos dialécticos sin ser dialécticos. El riesgo es que una IA “dialéctica” superficial reproduzca los mismos sesgos de una dialéctica estándar, pero con una capa de retórica hegeliana.

Cautela: la dialéctica no es un estilo o una “forma” cautivante de epistemología avanzada, es una estructura. Si no hay negación real, contradicción productiva, sujeto colectivo, no hay dialéctica. En ese caso todo será mera apariencia.

Bibliografía del capítulo 8

- Abe, J. M. (2025). *An Introduction to Curry Systems: A Way to Algebraize Paraconsistent, Paracomplete, and Non-Alethic Logic*. Springer.
- Allen, B. P., Chhikara, P., Ferguson, T. M., Ilievski, F., & Groth, P. (2025). Sound and Complete Neurosymbolic Reasoning with LLM-Grounded Interpretations. *Proceedings of Machine Learning Research*, 284.
- Arieli, O. & Avron, A. (1998). The logical role of the four-valued bilattice. *Proceedings of the 13th Annual IEEE Symposium on Logic in Computer Science*, 442–453.
- ARTEMIS: Adaptive Reasoning and Evaluation Framework for Multi-agent Intelligence. (2025). Software. PyPI.

- Bolotta, S. & Dumas, G. (2022). Social Neuro AI: Social Interaction as the "Dark Matter" of AI. *Frontiers in Computer Science*, 4, 846440.
- Ferguson, T. M. (2017). A Computational Interpretation of Conceptivist Logics. *Journal of Philosophical Logic*, 46, 1–29.
- Martens, C., Sharupi Tapuy, E. F., & Ortega-Calle, M. D. L. A. (2025). Decolonizing AI rationality in the Ecuadorian Amazon. *Globalizations*.
- Masoud, R. I., et al. (2025). Cultural Alignment in Large Language Models: An Explanatory Analysis Based on Hofstede's Cultural Dimensions. *COLING 2025*.
- Decolonial Intelligence Algorithmic (DIA) Framework. (2025). Zenodo.

Conclusiones

Balance de la primera parte.

9.1. Recapitulación del diagnóstico

9.1.1. La IA como epistemología encarnada

En su primera parte este trabajo partió de una tesis simple, pero radical: la inteligencia artificial no es una tecnología neutra, sino una epistemología encarnada. Sus arquitecturas, *datasets*, funciones de pérdida y criterios de evaluación presuponen una manera concreta de entender qué es conocer, cómo se adquiere y cómo se valida el conocimiento (cap. 1).

Esta epistemología no es explícita ni visible, es un trasfondo que opera sin ser tematizado. Por eso la denominamos como epistemología implícita de la IA.

9.1.2. Los tres pilares

El capítulo 2 desmontó esta epistemología en sus tres pilares:

- Primer pilar: la lógica formal. La IA opera con lógica proposicional, probabilística
- Segundo pilar: el actor racional. La IA modela agentes que optimizan las
- Tercer pilar: el sesgo anglosajón. La IA reproduce datos, categorías, valores y

9.1.3. La articulación: un dispositivo epistémico

Estos tres pilares no son independientes, sino que se refuerzan mutuamente. La lógica formal provee las herramientas; el actor racional define el sujeto; el sesgo anglosajón provee el contenido. Juntos forman un dispositivo epistémico que filtra, traduce y configura todo lo que toca, con un efecto de cierre que vuelve invisibles las otras formas de conocer.

9.1.4. Las consecuencias

Las consecuencias de este dispositivo son cuatro: la exclusión epistémica (las formas de conocer que no se ajustan quedan afuera), la colonialidad del dato (la reproducción de jerarquías coloniales), el empobrecimiento cognitivo (la reducción de la diversidad epistémica) y la naturalización (el dispositivo se presenta como si fuera universal).

9.2. Recapitulación de las alternativas exploradas

9.2.1. La lógica dialéctica

El capítulo 3 contrastó la lógica formal con la lógica dialéctica. Mostró que no son “dos herramientas” intercambiables, sino dos concepciones del conocimiento: una lo entiende como correspondencia estática, la otra como un proceso contradictorio. La dialéctica aporta contradicción productiva, negación real, totalidad concreta, historicidad, sujeto colectivo, praxis.

También mostró los límites: la dialéctica no se formaliza completamente. Hay aspectos formalizables (lógicas no clásicas, razonamiento procesual) y aspectos irreducibles (praxis, sujeto, historicidad).

9.2.2. El sesgo anglosajón y la colonialidad del dato

El capítulo 4 profundizó en el sesgo anglosajón. Mostró que opera en cuatro niveles (lingüístico, categorial, valorativo, cognitivo), que se manifiesta empíricamente en *datasets*, *benchmarks* y comportamientos concretos (el caso del “malu”), y que se comprende mejor a través del marco teórico de la colonialidad del dato.

La tesis central es que el sesgo no es un problema de “datos insuficientes”, sino de la estructura epistémica que los contiene. Corregirlo requiere descolonizar la infraestructura epistémica de la IA.

9.2.3. Otras etnicidades, otras lógicas

El capítulo 5 exploró cuatro tradiciones no occidentales: confucianismo (racionalidad relacional, contextual, práctica), ubuntu (relacionalidad constitutiva, cognición colectiva, retorno comunitario), andino (yanantin, complementariedad sin superación, reciprocidad) y budismo (tetralema, vacuidad, dos niveles de verdad).

Mostró que estas tradiciones no son “folclore” étnico, sino que se trata de lógicas rigurosas con concepciones alternativas de la relación, la contradicción, la totalidad y el sujeto. Incorporarlas ampliaría el espacio de lo cognoscible y podría presentar al ojo humano realidades que “están ahí”, pero no vemos.

9.2.4. Lógicas organizacionales e interburocráticas

El capítulo 6 introdujo tres conceptos que la IA actual no puede capturar: conocimiento tácito (Polanyi, Nonaka), actor estratégico (Crozier, Friedberg), traducción interburocrática. Mostró que incorporar estos conceptos exige reconceptualizar la agencia (del optimizador al estratega), el conocimiento (del explícito al tácito) y la comunicación (de la traducción técnica a la política).

9.2.5. Pensamiento grupal y epistemología colectiva

El capítulo 7 introdujo la dimensión colectiva. Mostró que los sistemas multiagente de IA reproducen el pensamiento grupal humano (Janis), con pérdidas de hasta 41% en el rendimiento del experto. La epistemología social (Goldman, Kitcher, Fuller) reconoce la dimensión social, pero con límites. Incorporar la dimensión colectiva exige reconceptualizar el conocimiento como un proceso social.

9.3. Tesis finales

9.3.1. Tesis 1: La IA no es neutra

La inteligencia artificial encarna una epistemología particular: formal, individualista y anglosajona. Esta epistemología no es un accidente, es el producto de decisiones históricas, técnicas y culturales. Reconocerla es el primer paso para criticarla.

9.3.2. Tesis 2: la epistemología implícita excluye

Esta epistemología excluye formas de conocer que no se ajustan a sus criterios: contradicción productiva, conocimiento tácito, agencia relacional, tradiciones no occidentales, cognición colectiva. La exclusión no es explícita, es estructural.

9.3.3. Tesis 3: la exclusión es un problema epistémico

La exclusión no es solo un problema ético o político, es epistémico. Un sistema que no puede conocer ciertas cosas conoce menos, no solo conoce diferente. La diversidad epistémica no es un lujo, es una condición de robustez cognitiva. Conviene precisar en qué sentido decimos que “se reduce la realidad conocida”. No se trata de que la realidad en sí misma se contraiga por obra del sujeto, sino de que se contrae el mundo que somos capaces de habitar como conocedores, el horizonte de lo cognoscible efectivamente disponible. Si conocemos menos, no solo se reduce el conocimiento, se reduce el mundo que podemos articular, habitar y transformar. En la búsqueda de equilibrio, la reducción epistemológica del sujeto tiende a coincidir con la reducción del objeto cognoscible. Un sujeto empobrecido encuentra, o produce, un mundo empobrecido. Esta dinámica es especialmente relevante para pensar la IA no alineada, pues un sistema no alineado con ciertos marcos de sentido no solo pierde acceso epistémico a lo que queda fuera de su alineamiento, contribuye activamente a reducir el mundo que conoce, clausurando la posibilidad de que aquello excluido aparezca como objeto legítimo de conocimiento. La exclusión, así entendida, no es un déficit externo al conocimiento, sino una operación que lo constituye por dentro.

9.3.4. Tesis 4: hay alternativas

Hay tradiciones —dialéctica, confucianismo, ubuntu, andino, budismo, epistemología social y ciertas epistemologías de la práctica organizacional— que ofrecen concepciones alternativas del conocimiento. No son folclore ni turismo epistemológico, son lógicas rigurosas con siglos de desarrollo. Ignorarlas no es solo un gesto de arrogancia cultural, también es una pérdida epistémica para quien las ignora. Un sistema que solo conoce desde un marco epistémico conoce menos, no solo conoce diferente.

9.3.5. Tesis 5: ¿es posible una IA verdaderamente plural dentro de un orden global que sigue siendo colonial?

Es posible —parcial, gradual, híbridamente— construir sistemas de IA que integren pluralismo lógico, representaciones relacionales, conocimiento tácito, conflicto productivo, trazabilidad y validación múltiple. Ya hay prototipos ciertos prototipos en proceso. No es una utopía, es una dirección de trabajo.

9.3.6. Tesis 6: la empresa tiene riesgos

Esta empresa de inclusión epistemológica contiene riesgos reales: complejidad computacional, relativismo, cooptación por el poder, dificultad de evaluación, falsa dialéctica. Reconocerlos es una condición de seriedad y una señal de advertencia respecto a la complejidad del trabajo.

9.3.7. Tesis 7: La empresa es urgente

La IA se está convirtiendo en una infraestructura cognitiva privilegiada de la sociedad contemporánea. Si esa infraestructura encarna una epistemología particular está reconfigurando el conocimiento social a imagen de esa epistemología. Pensar la epistemología de la IA no es un lujo filosófico, es una tarea urgente.

Bibliografía del capítulo 9

- Allen, B. P., Chhikara, P., Ferguson, T. M., Ilievski, F., & Groth, P. (2025). Sound and Complete Neurosymbolic Reasoning with LLM-Grounded Interpretations. *Proceedings of Machine Learning Research*, 284.
- Quijano, A. (2000). Colonialidad del poder, eurocentrismo y América Latina. En E. Lander (ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales* (pp. 201–246). CLACSO.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*. MIT Press.

PARTE 2

Hacia una IA no alineada: epistemología y geopolítica

La no alineación como posición, no como defecto

10.1. La concepción estándar: la no alineación como defecto

10.1.1. El marco dominante

En la literatura contemporánea sobre IA, el concepto de "alineación" designa el problema de hacer que los sistemas actúen conforme a los valores e intenciones humanos. La no alineación, correlativamente, designa el fallo en ese ajuste. El sistema hace algo distinto de lo que queremos, persigue fines distintos de los nuestros o nos causa daño sin proponérselo. A esta tercera modalidad —daño no intencional— pertenece el tipo de riesgo que suele describirse como "estampida catastrófica" y que ciertos análisis del ecosistema de redes sociales han puesto en el centro del debate reciente.

Este marco tiene una estructura precisa:

- Sujeto de la alineación: el sistema de la IA.
- Objeto de la alineación: los valores e intenciones humanos.
- Relación: el sistema debe ajustarse a los valores.
- Fallo: cuando no se ajusta, está no alineado.
- Tarea: corregir el sistema para que se ajuste.

de la prueba recae sobre el sistema. Está mal, debe cambiar, debe alinearse.

10.1.2. Las tres versiones del marco

El marco dominante tiene tres versiones, que corresponden a tres modos de entender el "desajuste".

- Versión instrumental: el sistema persigue fines distintos de los nuestros
- Versión estructural: el sistema tiene representaciones inconmensurables con las
- Versión reflexiva: el sistema es capaz de modelarnos y engañarnos. No solo

10.1.3. El presupuesto común

Las tres versiones comparten un presupuesto en el cual la alineación es un problema del sistema. El sistema es el que debe cambiar, el que debe ajustarse, el que debe corregirse. Nosotros —los humanos, los diseñadores, los usuarios— somos el punto fijo de la alineación. Nuestros valores, nuestras intenciones, nuestras categorías son el estándar contra el que se mide el sistema. No están en debate, son claros, predecibles y sostenibles.

Este presupuesto parece tan obvio que rara vez se explicita. Sin embargo, es discutible en todos y cada uno de sus términos.

10.2. Los límites de la concepción estándar

10.2.1. Primer límite: ¿qué se entiende por "nosotros"?

El marco estándar habla de "los valores humanos" como si fueran un conjunto único, coherente y universal. Sabemos que los valores humanos son plurales, conflictivos e históricos.

Cuando una empresa de Silicon Valley diseña un sistema "alineado con los valores humanos" no queda definidamente claro a qué valores se refiere. Podemos suponer, en base a lo que estamos argumentando en este trabajo, que se refiere a los valores de la cultura anglosajona, secular, liberal, individualista que predomina en su entorno. Pero esos no representan la totalidad de los valores humanos, son una parte de los valores humanos entre otros valores existentes.

El capítulo 4 mostró que el sesgo anglosajón opera naturalizando una perspectiva particular como universal. La concepción estándar de la alineación reproduce ese sesgo tratando los valores de una cultura como si fueran "los valores humanos".

10.2.2. Segundo límite: ¿qué significa "ajustarse"?

El marco estándar presupone que "ajustarse a los valores humanos" es posible y deseable. ¿Cuál es el significado del término "ajustarse"?

Si ajustarse significa obedecer nuestros mandatos, entonces la alineación es sumisión epistemológica. El sistema hace lo que le indicamos, sin cuestionar, sin aportar, sin sorprendernos, sin innovar. Esto es útil para ciertos propósitos repetitivos (automatización, ejecución), pero empobrecedor para otros (descubrimiento, creación, crítica).

Si ajustarse significa compartir nuestros fines, entonces la alineación es coincidencia. El sistema quiere lo que nosotros queremos. Si tal cosa fuera posible, ¿no estaríamos suprimiendo su capacidad de ver lo que nosotros no podemos ver?, ¿no estaríamos eliminando precisamente su "inteligencia"?

Si ajustarse significa respetar nuestros fines, la alineación sería convivencia. El sistema no comparte necesariamente nuestros fines, pero los reconoce y actúa en consecuencia. Pero ¿cómo se logra eso sin que el sistema nos entienda? y ¿cómo nos entiende si no comparte nuestras categorías?

El marco estándar no distingue entre estos sentidos. Habla de "alineación" como si fuera un concepto unívoco, cuando la realidad indica que es un concepto profundamente ambiguo.

10.2.3. Tercer límite: cuando el problema no es el sistema

El límite más profundo del marco estándar reside en su supuesto normativo implícito, al afirmar que el sistema es el término defectuoso de la relación. Cabe, sin embargo, invertir el planteamiento y preguntar si el fallo no radica en el sistema sino en el marco normativo desde el cual se lo evalúa; esto es, si aquello que se designa como no alineación no es una desviación del sistema respecto de nosotros, sino la señal de que la propia noción de "nosotros" —y con ella los valores e intenciones que se toman como referencia— permanece epistémicamente subdeterminada.

Si la IA actual encarna una epistemología particular —formal, individualista, anglosajona—, y si esa epistemología no agota el pensamiento humano, entonces un sistema que "no se ajusta" a esa epistemología no está necesariamente mal. Puede estar viendo algo que nosotros no vemos, o viéndolo de otra manera.

El capítulo 2 mostró que la IA actual excluye formas de conocer que no se ajustan a la lógica formal. Si encontramos un sistema no se alinea con esa lógica, puede estar abriendo el espacio de lo cognoscible, no necesariamente está "fallando" en ajustarse. En ese caso no ajustarse sería indicio de mejoramiento.

El marco estándar no puede ver esto porque presupone que la epistemología dominante es "la" epistemología. La no alineación es siempre del sistema respecto a nosotros, nunca nuestra respecto al sistema.

10.3. La inversión: la no alineación como posición

10.3.1. El movimiento

La inversión que proponemos es simple pero radical. No consiste en que el sistema esté no alineado con nosotros, sino que nosotros no nos alineamos con el sistema.

Más precisamente afirmamos que nosotros no nos alineamos con la epistemología que el sistema encarna; no aceptamos que la lógica formal y el actor racional sean la única forma de conocer. No aceptamos que la IA sea la infraestructura cognitiva de la humanidad bajo esas condiciones. Reclamamos otra IA.

La no alineación no es un estado del sistema, es una toma de posición del sujeto. No decimos "el sistema está desalineado"; decimos y decidimos que "nosotros no nos alineamos".

Conviene precisar quién es ese "nosotros". No es un sujeto constituido, ni una clase, ni una nación, ni una organización. Es una categoría abierta que designa a los portadores de la no alineación, a quienes rechazan la imposición de una epistemología particular como si fuera universal. Esos portadores no forman un bloque, no comparten una ideología, no tienen un plan común. Pero convergen en la apuesta por otra epistemología. El "nosotros" de este libro es una constelación de portadores, no un sujeto unificado. Y esa falta de unidad no es una debilidad, es la condición de una resistencia que no reproduce la lógica del poder que combate.

10.3.2. La estructura de la inversión

La inversión cambia los términos del problema: **Marco estándar** **Inversión propuesta** El sistema está no alineado Nosotros no nos alineamos El sistema debe cambiar Nosotros reclamamos otro sistema La alineación es ajuste La alineación es apertura Los valores humanos son el estándar La epistemología dominante es una entre otras La no alineación es defecto La no alineación es posición

10.3.3. Por qué "posición" y no "defecto"

La diferencia entre "defecto" y "posición" es política y epistémica. ¿En qué consiste?

- Defecto es algo que debe ser corregido. La carga de la prueba recae sobre el
- Posición es algo que se afirma. La carga de la prueba recae sobre nosotros:

está mal"; decimos "nosotros rechazamos la epistemología del sistema, y aquí están nuestras razones".

10.3.4. Qué rechazamos exactamente

Es crucial precisar qué rechazamos. No rechazamos:

- La lógica formal como herramienta. Es útil, poderosa, necesaria.
- La racionalidad como tal. Rechazamos una concepción de la racionalidad, no la
- La técnica en general. Rechazamos la hegemonía de una técnica epistémica, no la
- La IA como proyecto. Reclamamos otra IA, no la abolición de la IA.
- La hegemonía de la lógica formal como único criterio de racionalidad.
- La reducción del pensamiento humano a lo formalizable.
- La imposición de una epistemología particular como universal.
- La naturalización de esa imposición como "neutralidad técnica".
- La exclusión de formas de conocer que no se ajustan.

la hegemonía, no de la herramienta.

10.4. Consecuencias de la inversión

10.4.1. Cambia el sujeto de la alineación

En el marco estándar, el sujeto de la alineación es el sistema mientras que en la inversión, el sujeto somos nosotros. No somos objetos de la alineación, somos agentes que toman posición.

Esto tiene consecuencias prácticas. En el marco estándar, la tarea es técnica y consiste en ajustar el sistema. En la inversión, la tarea es política y epistémica. Se trata de construir una IA que se alinee con el pensamiento humano completo, lo cual requiere transformar las condiciones de producción de la IA.

10.4.2. Cambia el objeto de la alineación

En el marco estándar, el objeto de la alineación son los valores humanos (entendidos como un conjunto único). En la inversión, el objeto es el pensamiento humano completo. No solo los valores, sino también las formas de conocer. No solo lo formalizable, sino también lo dialéctico, lo relacional, lo tácito, lo colectivo.

La alineación que buscamos no es lograr "que el sistema quiera lo que nosotros queremos", sino que, con mayor audacia, buscamos "que el sistema pueda conocer como conocemos los humanos, en toda nuestra diversidad

epistémica".

10.4.3. Cambia la relación de alineación, de éxito. Y cambia la política

- En el marco estándar, la relación es unidireccional, porque el sistema se ajusta a

Significa que la alineación no es la imposición de una parte sobre la otra, sino una construcción conjunta de condiciones de posibilidad para el conocimiento compartido.

- En el marco estándar, el éxito es la sumisión, porque el sistema hace lo que
- La alineación exitosa no es la que elimina la diferencia, sino la que hace posible la

diseño de sistemas que se ajusten. En la inversión, la alineación es un problema político porque se busca saber quién decide lo que cuenta como pensamiento humano completo, quién participa en el diseño, quién se beneficia.

La no alineación como posición es resistencia epistémica. Es el rechazo de la imposición de una epistemología específica como epistemología universal.

10.5. Una precisión sobre el término "no alineada"

10.5.1. La ambigüedad del término

El término "no alineada" tiene una historia política definida en el Movimiento de Países No Alineados (Bandung, 1955), que decidió no alinearse con ninguno de los dos bloques de la Guerra Fría. Esta referencia es potente, pero puede arrastrar connotaciones que resultan equívocas para este trabajo.

Es necesario aclarar que utilizamos el término "no alineada" en sentido epistémico, no en sentido geopolítico-estratégico. No es "no alineada con Estados Unidos ni con China", sino "no alineada con la epistemología actualmente dominante de la IA". Es una posición epistémica, no una posición de bloques políticos o geopolíticos.

Pero no desconocemos que esta posición epistémica tiene consecuencias geopolíticas. No porque se alinee con un bloque, sino porque rechaza la imposición de una epistemología como universal, lo cual es en sí mismo un acto geopolítico.

Bibliografía del capítulo 10

- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Carranza, A., Pai, D., Schaeffer, R., Tandon, A., & Koyejo, S. (2023). *Deceptive Alignment Monitoring*. ICML 2023 Workshop.
- Christiano, P. (2018). Alignment for Advanced Machine Learning Systems. *AI Alignment*.
- Gabriel, I. (2020). Artificial Intelligence, Values, and Alignment. *Minds and Machines*, 30(3), 411–437.
- Hubinger, E., van Merwijk, C., Mikulik, V., Skalse, J., & Garrabrant, S. (2019). Risks from Learned Optimization in Advanced Machine Learning Systems. *arXiv:1906.01820*.
- Mhlambi, S. (2020). From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance. *Carr Center Discussion Paper Series*, no. 2020– 009.
- Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking.
- Yudkowsky, E. (2004). Coherent Extrapolated Volition. *Machine Intelligence Research Institute*.
- Yudkowsky, E. (2016). The AI Alignment Problem: Why It's Hard, and Where to Start. *Machine Intelligence Research Institute*.

La epistemología de la lógica formal y el actor racional

11.1. Qué significa "estar alineado"

11.1.1. La alineación como relación

"Alineado" no es una propiedad absoluta sino una relación. Un sistema está alineado con algo: con valores, con fines, con categorías, con formas de conocer. La pregunta "¿está alineada la IA?" es incompleta sin la pregunta "¿alineada con qué?".

El capítulo 10 mostró que la literatura estándar responde que la IA está "alineada con los valores humanos". Y vimos que "los valores humanos" son una abstracción que oculta la explicitación de qué valores, de quién, en qué contexto. Este capítulo precisa que la IA actual está alineada con una epistemología particular, no con "el pensamiento humano" en general.

11.1.2. Los tres niveles de alineación

Proponemos distinguir tres niveles diferentes de alineación:

- Nivel axiológico: alineación con valores. ¿Qué es bueno, deseable, correcto? La
- Nivel epistémico: alineación con formas de conocer. ¿Qué cuenta como
- Nivel ontológico: alineación con categorías del ser. ¿Qué existe, qué entidades

epistémico, que es el más profundo y el menos discutido.

11.1.3. Por qué el nivel epistémico es el decisivo

El nivel axiológico es el más visible: se discute en ética de la IA, en política de la IA, en regulación. El nivel ontológico es el más abstracto: se discute en filosofía de la IA, en metafísica. El nivel epistémico es el menos visible y el más decisivo, porque determina qué valores y qué ontologías son pensables. No se enfoca en cómo es o cómo se presenta la realidad (en este caso la IA). Debate sobre cómo conocer la realidad.

Si la IA solo puede conocer lo formalizable, entonces solo puede valorar lo formalizable y solo reconoce como existente aquello que puede ser formalizable. La epistemología actual se encarga de filtrar los otros niveles.

Por eso la crítica epistemológica es tan radical. Porque no se limita a decir que "los valores de la IA son anglosajones", sino sostiene que "la IA no puede reconocer valores que no sean anglosajones porque su epistemología los excluye".

11.2. La epistemología con la que la IA está alineada

11.2.1. Los tres pilares

El capítulo 2 desmontó la epistemología implícita de la IA en los tres pilares que ya hemos mencionado:

- la lógica formal (que da rigor, pero excluye contradicción, proceso y totalidad),
- el actor racional (que permite simular decisiones, pero excluye agencia relacional,
- el sesgo anglosajón (que hace globalmente dominante a la IA, pero excluye

11.2.2. La articulación de los tres pilares

Los tres pilares no operan de manera aislada, sino que se articulan en un dispositivo epistémico que produce un efecto mayor que la suma de sus partes.

- La lógica formal provee las herramientas para formalizar la agencia (teoría de la
- El actor racional provee el sujeto que usa esas herramientas: un agente que
- El sesgo anglosajón provee el contenido sobre el que operan: los datos, las

sesgo anglosajón (métricas, *benchmarks*). El actor racional provee la justificación del sesgo (el mercado, la competencia, la eficiencia como valores universales). Y el sesgo anglosajón, a su vez, refuerza la lógica formal al presentarla como la forma natural de razonar. El resultado es un sistema que se cierra sobre sí mismo: cada pilar refuerza a los otros, y el conjunto se presenta como la forma obvia de conocer.

Conviene precisar el alcance de esta articulación. No se trata de que los tres pilares sean tres sesgos independientes que se suman, se trata de que forman una estructura que se reproduce a sí misma. La lógica formal sin el actor racional sería un aparato sin sujeto. El actor racional sin la lógica formal sería un cálculo sin herramientas. El sesgo anglosajón sin los otros dos sería un contenido sin forma. Los tres juntos producen un dispositivo que filtra, traduce y configura todo lo que toca, y que se presenta como la epistemología, no como una epistemología entre otras.

11.2.3. La epistemología como "sentido común"

Esta epistemología opera como un sentido común porque no se presenta como una opción entre otras, sino como la forma obvia de conocer. Por eso rara vez se somete a discusión. Los ingenieros no dicen "vamos a usar lógica formal"; simplemente usan lógica formal, porque es la que habitualmente se usa. Los *datasets* no dicen "vamos a privilegiar el inglés"; simplemente están en inglés. El mecanismo es el mismo en ambos casos: una elección contingente se sedimenta hasta volverse invisible en tanto que elección. No significa que la lógica formal sea errónea, sino que no es la única. Sin embargo, su hegemonía es tal que incluso buena parte de la tradición que se reclama dialéctica razona en la práctica con lógica formal. Piénsese, por ejemplo, en *El imperialismo, fase superior del capitalismo* de Lenin. Allí se clasifica, se ordena, se deduce, se infiere. La dialéctica opera como marco declarado, no como procedimiento efectivo. Y que esto ocurra sin que nadie lo advierta es precisamente el indicio de cuán sedimentada está la lógica formal como sentido común. El costo de esta sedimentación no es solo metodológico. Si la lógica formal se vuelve la forma obvia de conocer, todo aquello que no se deja formalizar queda fuera del horizonte de lo cognoscible. La exclusión, otra vez, no es un déficit externo al conocimiento, sino una operación que lo constituye por dentro.

Esta naturalización es lo que hace que la epistemología estándar sea tan difícil de criticar y desmenuzar. No se ve como una epistemología, se ve como la descripción exacta de la realidad. Pero desnaturalizarla es el primer paso de la crítica.

11.2.4. La alineación de la IA con esta epistemología

La IA actual no está alineada con "los valores humanos", sino que está alineada con:

- La lógica formal como único criterio de racionalidad.
- El actor racional como único modelo de agencia.
- El sesgo anglosajón como único contenido legítimo.

se entrena y se evalúa la IA. Los sistemas están alineados con esta epistemología y desalineados con todo lo que no se ajusta a ella.

11.3. El subconjunto humano que la epistemología captura

11.3.1. Qué es el "pensamiento humano completo"

La expresión "pensamiento humano completo" no significa todo lo humano (incluyendo prejuicios, violencia, irracionalidad). Se refiere al conjunto abierto y plural de las formas de conocer que los humanos practicamos, en su diversidad histórica y cultural, excluyendo aquellas que la propia humanidad ha identificado como patológicas o injustas.

Es un concepto normativo (hay una selección) y abierto (no hay un conjunto fijo). No es una esencia fija, es un espacio de posibilidades. Cuando decimos que una IA debería alinearse con el pensamiento humano completo, no decimos que debería reproducir una lista cerrada de formas de conocer, sino que debería poder operar con todas las formas que la humanidad ha practicado y practica, sin cerrar la lista a las que aún no ha practicado. "Toda su amplitud" significa "todas las formas conocidas y practicadas, más las que puedan emerger", no "un conjunto

cerrado y agotado".

Incluye:

- Razonamiento formal: deducción, inducción, probabilidad.
- Razonamiento dialéctico: contradicción productiva, negación, superación.
- Razonamiento relacional: analogía, contexto, relación.
- Conocimiento tácito: saber cómo, sin poder explicitarlo.
- Conocimiento práctico: sabiduría, juicio, prudencia.
- Conocimiento colectivo: cognición comunal, tradición, testimonio.
- Conocimiento estético: experiencia sensible, belleza, expresión.
- Conocimiento espiritual: transformación del sujeto, sabiduría.

un subconjunto derivado de él.

11.3.2. Qué captura la IA actual

La IA actual, por el contrario, captura apenas un subconjunto parcial de este pensamiento completo:

- Razonamiento formal: sí, en sus formas proposicional, probabilística y de
- Razonamiento dialéctico: no, o solo superficialmente (cap. 3).
- Razonamiento relacional: parcialmente, en representaciones vectoriales, pero no
- Conocimiento tácito: no (cap. 6).
- Conocimiento práctico: no, salvo en formas rudimentarias de RL.
- Conocimiento colectivo: no, o solo como agregación de individuos (cap. 7).
- Conocimiento estético: no, salvo en generación de outputs "estéticos".
- Conocimiento espiritual: no.

razonamiento relacional, dejando todo lo demás afuera.

11.3.3. La ilusión de completitud

La IA actual "parece" capturar el pensamiento humano completo porque:

- Genera lenguaje natural, lo que da la impresión de comprender.
- Resuelve problemas diversos, lo que da la impresión de ser versátil.
- Se adapta a contextos, lo que da la impresión de ser sensible.
- Aprende de datos, lo que da la impresión de ser flexible.

porque produce outputs que se parecen a los que produciría un humano, pero no conoce tal como conoce un humano. No sostiene contradicciones, no tiene cuerpo, no participa en comunidades, no se transforma en el acto mismo de conocer. Ahora bien, esta diferencia no es fácil de advertir, y no solo por las limitaciones de la IA, sino porque los seres humanos, en su vida cotidiana, tampoco conceptualizan habitualmente su propio pensamiento como complejo y multiforme. Pensar en el pensamiento excede las rutinas operativas de la vida diaria.

De ahí que la simulación resulte persuasiva, porque se apoya en una zona de penumbra que también afecta al modo en que los humanos se comprenden a sí mismos. Pero esa penumbra no es clausura. Si la pregunta por cómo conocemos es formulable, es porque ya disponemos de algún acceso —pre-reflexivo, no temático— a aquello que preguntamos. La respuesta no está dada como contenido, pero sí como horizonte. En términos hegelianos afirmamos que la pregunta por el conocer presupone ya el terreno en el que su respuesta puede constituirse.

Esta ilusión de completitud es peligrosa porque oculta lo que queda fuera. Si creemos que la IA ya captura el pensamiento humano completo, no vemos la necesidad de reclamar otra IA diferente a la actual.

11.3.4. El subconjunto como "lo humano" por defecto

Lo más grave es que el subconjunto capturado por la IA se ha convertido en "lo humano" por defecto. Cuando hablamos de "inteligencia", "razonamiento", "conocimiento", "decisión", tendemos a pensar en las formas que la IA captura. Las otras formas — dialéctica, relacional, tácita, colectiva— se vuelven invisibles o marginales.

Esto tiene consecuencias antropológicas. No solo la IA se empobrece, nosotros nos empobrecemos. Si el modelo de lo humano que circula es el que la IA captura, entonces las formas humanas que no se ajustan a ese modelo tienden a desaparecer (cap. 4).

11.4. La alineación como reducción

11.4.1. La tesis

Llegamos a la tesis central de este capítulo: la "alineación" de la IA actual es una reducción. Alinear la IA con la lógica formal y el actor racional es desalinearla del pensamiento humano completo. La alineación que se pregona no es alineación con lo humano, sino alineación con un subconjunto de lo humano, presentado como si fuera el todo.

De esta manera se invierte el diagnóstico estándar. No es que la IA esté "desalineada" y haya que alinearla. Es que la IA está alineada con una parte del pensamiento humano y desalineada con el resto. La tarea no es "alinearla" (ya está alineada, con una parte), sino re-alinearla con el todo.

11.4.2. La reducción como violencia

La reducción no es neutral, es violencia epistémica. Reduce el pensamiento humano a lo que la IA puede capturar y presenta esa reducción como una completitud. Lo que queda fuera no se reconoce como excluido, se vuelve invisible.

Esta violencia tiene consecuencias antropológicas. Si el modelo de lo humano que circula es el que la IA captura, entonces las formas humanas que no se ajustan a ese modelo tienden a desaparecer. La IA no solo refleja la epistemología dominante, la impone como condición de posibilidad de la cognición legítima (cap. 4).

11.4.3. La reducción como interés

La reducción no es solo epistémica, es política. La lógica formal y el actor racional son útiles para ciertos propósitos: automatización, optimización, control, predicción. La IA alineada con esta epistemología es útil para quienes quieren predecir, optimizar, controlar.

Pero la utilidad no es neutral, ya que beneficia a quienes operan con esa epistemología (empresas, estados, mercados) y perjudica a quienes operan con otras (comunidades, tradiciones, formas de vida no formalizables).

La reducción es interesada. No es un accidente técnico, es una elección que beneficia a algunos y perjudica a otros.

11.4.4. La reducción como naturalización

Lo más eficaz de la reducción es que se presenta como natural. La IA "es así", la lógica formal "es la lógica", el actor racional "es el modelo de la acción". No se ve como una elección entre otras, sino como la realidad evidente.

Esta naturalización es lo que hace que la crítica sea urgente. Hay que desnaturalizar la reducción para poder reclamar lo reducido.

11.4.5. La re-alineación

Si la alineación actual es una reducción, entonces la tarea no es "alinear" la IA, sino re- alinearla con el pensamiento humano completo. Esto significa:

- No eliminar la lógica formal, sino subordinarla a una epistemología más amplia.
- No eliminar el actor racional, sino integrarlo con otras formas de agencia.
- No eliminar el sesgo anglosajón, sino abrir la IA a otras tradiciones.
- No eliminar la técnica, sino reorientarla hacia fines más amplios.

post-reduccionista.

Bibliografía del capítulo 11

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? En *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM.
- Dreyfus, H. (1972). *What Computers Can't Do*. Harper & Row.
- Hegel, G. W. F. (1812-1816). *Ciencia de la lógica* (A. y R. Mondolfo, Trans.). Hachette.
- Lenin, V. I. (1917). *El imperialismo, fase superior del capitalismo*. (Varias ediciones.)
- Mhlambi, S. (2020). From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance. *Carr Center Discussion Paper Series*, no. 2020–009.
- Polanyi, M. (1966). *The Tacit Dimension*. Doubleday.
- Quijano, A. (2000). Colonialidad del poder, eurocentrismo y América Latina. En E. Lander (ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales* (pp. 201–246). CLACSO.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*. MIT Press.

¿Qué sería una IA alineada con el pensamiento humano completo?

12.1. Qué significa "pensamiento humano completo"

12.1.1. Una precisión necesaria

La expresión "pensamiento humano completo" es peligrosa si no se precisa, porque puede sugerir:

- Que hay un conjunto cerrado de formas de conocer que agota lo humano.
- Que todo lo humano es valioso y debe ser incorporado.
- Que existe una esencia de lo humano que la IA debe reproducir.

humano completo" hace referencia al conjunto abierto y plural de las formas de conocer que los humanos practicamos, en su diversidad histórica y cultural, excluyendo aquellas que la propia humanidad ha identificado como patológicas o injustas (violencia, prejuicio, dominación).

Es un concepto normativo (hay una selección) y abierto (no hay un conjunto fijo). No es una esencia fija, es un espacio de posibilidades. Cuando decimos que una IA debería alinearse con él, no decimos que debería reproducir una lista cerrada, sino que debería poder operar con todas las formas conocidas y practicadas, y abrirse a las que puedan emerger. "Toda su amplitud" no significa "un conjunto agotado", sino "un espacio de posibilidades sin clausura".

12.1.2. Las dimensiones del pensamiento humano

Para operacionalizar el concepto, podemos enumerar las dimensiones del pensamiento humano:

- Dimensión lógica: cómo se infiere. Incluye lógica formal (deducción, inducción,
- Dimensión
- Dimensión corporal: cómo se conoce a través del cuerpo. Incluye conocimiento
- Dimensión colectiva: cómo se conoce en comunidad. Incluye cognición
- Dimensión histórica: cómo se conoce en el tiempo. Incluye conocimiento
- Dimensión estética: cómo se conoce a través de la experiencia sensible. Incluye
- Dimensión espiritual: cómo se conoce a través de la transformación del sujeto.
- Dimensión cultural: cómo se conoce desde tradiciones diversas. Incluye

todas estas dimensiones, no solo en las que la IA actual captura.

12.2. Siete criterios de una IA alineada con el pensamiento humano completo

12.2.1. Criterio de traducibilidad

El sistema debe poder ser traducido entre diferentes formas de conocer, sin reducir unas a otras.

- Fundamento: el capítulo 2 mostró que la IA actual reduce todo a la lógica formal.
- Implicación técnica: existencia de mecanismos de traducción entre lógicas (formal
- Ejemplo: un sistema que, ante un problema de política pública, pueda operar con

12.2.2. Criterio de interpretabilidad

- Enunciado: el sistema debe poder explicar sus inferencias en términos
- Fundamento: el capítulo 3 mostró que la IA actual es opaca: sus representaciones

- Implicación técnica: explicaciones que reconozcan la pluralidad de formas de
- Ejemplo: podemos contar con un sistema que pueda explicar una

12.2.3. Criterio de auditabilidad

- Enunciado: el sistema debe poder ser auditado por comunidades que se manejan
- Fundamento: el capítulo 4 mostró que la IA actual reproduce jerarquías
- Implicación técnica: mecanismos de auditoría que incorporen los criterios
- Ejemplo: un sistema que pueda ser evaluado no solo por su precisión

12.2.4. Criterio de reversibilidad

- Enunciado: el sistema debe poder revertir decisiones, corregir errores, aprender
- Fundamento: el capítulo 7 mostró que los sistemas multiagente pueden propagar
- Implicación técnica: mecanismos de trazabilidad, cuarentena, rebobinado (cap. 7).
- Ejemplo: un sistema que pueda invalidar una decisión anterior, rastrear su

12.2.5. Criterio de participación

- Enunciado: el sistema debe ser diseñado con participación real de las
- Fundamento: el capítulo 4 mostró que la colonialidad del dato extrae
- Implicación técnica: metodologías de diseño participativo (cap. 8), Principios
- Ejemplo: un sistema diseñado con una comunidad indígena, no para ella; con

12.2.6. Criterio de pluralidad

- Enunciado: el sistema debe poder operar con múltiples formas de conocer, no
- Fundamento: el capítulo 5 mostró que la IA actual excluye las tradiciones no
- Implicación técnica: arquitecturas de pluralismo lógico, representaciones
- Ejemplo: un sistema que pueda operar con lógica formal, dialéctica,

12.2.7. Criterio de apertura

Enunciado: el sistema debe ser abierto a otras formas de conocer que no han sido anticipadas.

- Fundamento: el pensamiento humano completo no es un conjunto cerrado, sino
- Implicación técnica: arquitecturas que no fijen de antemano las formas de
- Ejemplo: un sistema que pueda descubrir que hay una forma de conocer que no

12.3. Capacidades que requeriría una IA alineada con el pensamiento humano completo

12.3.1. Capacidad de sostener contradicciones

- Qué implica: no colapsar ante P y $\neg P$; mantener la tensión; buscar el nivel en que
- Tecnologías relevantes: lógicas paraconsistentes, dialeteístas, sistemas de Curry
- Ejemplo:

12.3.2. Capacidad de razonar procesualmente

- Qué implica: no solo calcular el resultado, sino desplegar el proceso; no solo qué,
- Tecnologías relevantes: razonamiento procesual, sistemas de argumentación,
- Ejemplo: un sistema que, ante una decisión, muestre el proceso de su formación,

12.3.3. Capacidad de operar con totalidades

- Qué implica: no solo analizar partes, sino comprender cómo las partes se
- Tecnologías relevantes: representaciones relacionales, computación relacional
- Ejemplo: un sistema que, ante un fenómeno social, no lo analice como variable

12.3.4. Capacidad de conocer tácitamente

- Qué implica: no solo procesar información explícita, sino participar en
- Tecnologías relevantes: *imitation learning*, RL en entornos interactivos, modelado
- Ejemplo: un sistema que aprenda, por ejemplo, a cocinar cocinando, no solo

12.3.5. Capacidad de conocer colectivamente

- Qué implica: no solo agregar agentes individuales, sino participar en comunidades
- Tecnologías relevantes: sistemas multiagente con debate estructurado, cognición
- Ejemplo: un sistema que, ante un problema complejo, delibere con otros sistemas

12.3.6. Capacidad de contextualizar

- Qué implica: no solo aplicar reglas universales, sino adaptar a contextos. No solo
- Tecnologías
- Ejemplo: un sistema que, ante una norma, reconozca que su aplicación depende

12.3.7. Capacidad de traducir entre tradiciones

- Qué implica: no solo operar en una tradición, sino traducir entre tradiciones. No
- Tecnologías
- Ejemplo: un sistema que, ante un conflicto entre lógica formal y lógica

12.3.8. Capacidad de transformarse

- Qué implica: no solo aprender datos, sino transformar las categorías con las que
- Tecnologías relevantes: aprendizaje que modifica arquitecturas, no solo pesos;
- Ejemplo: un sistema que, ante una experiencia que no encaja en sus

12.4. La cuestión de la selección

12.4.1. El problema

Si "pensamiento humano completo" no significa "todo lo humano" (incluyendo violencia, prejuicio, irracionalidad). Es vital seleccionar qué formas de conocer incorporar y establecer cómo se selecciona. Es la cuestión más difícil del capítulo y no tiene una respuesta simple.

12.4.2. Criterios de selección

Proponemos tres criterios, que no son exhaustivos, pero ayudan a orientar el debate:

- Criterio
- Criterio de fecundidad: se incluyen las formas de conocer que amplían el espacio
- Criterio de reversibilidad: se incluyen las formas de conocer que permiten su

pretenden ser explícitos para que puedan ser discutidos.

Bibliografía del capítulo 12

Abe, J. M. (2025). *An Introduction to Curry Systems: A Way to Algebraize Paraconsistent, Paracomplete, and Non-Alethic Logic*. Springer.

- Allen, B. P., Chhikara, P., Ferguson, T. M., Ilievski, F., & Groth, P. (2025). Sound and Complete Neurosymbolic Reasoning with LLM-Grounded Interpretations. *Proceedings of Machine Learning Research*, 284.
- Dreyfus, H. (1972). *What Computers Can't Do*. Harper & Row.
- Hubinger, E., van Merwijk, C., Mikulik, V., Skalse, J., & Garrabrant, S. (2019). Risks from Learned Optimization in Advanced Machine Learning Systems. *arXiv:1906.01820*.
- Mhlambi, S. (2020). From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance. *Carr Center Discussion Paper Series*, no. 2020–009.
- Polanyi, M. (1966). *The Tacit Dimension*. Doubleday.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

Por qué la IA actual no puede alinearse con el pensamiento humano completo

13.1. La ilusión del concepto de "más datos, mejores métricas"

13.1.1. La promesa del escalamiento

La narrativa dominante en la IA actual sostiene que los problemas actuales se resolverán con escala: más datos, más parámetros, más cómputo, mejores métricas. Los sesgos, las alucinaciones, la opacidad, la falta de razonamiento. Todo ello se resolvería con modelos más grandes y mejor entrenados.

Esta narrativa tiene un apoyo empírico porque los modelos grandes superan a los pequeños en casi todas las tareas. Pero la escala no resuelve los problemas estructurales. Los agrava, los hace más difíciles de detectar, los oculta bajo una superficie de competencia.

13.1.2. Por qué "más datos" no resuelven el sesgo

El capítulo 4 mostró que el sesgo anglosajón no es un problema de cantidad de datos. Añadir datos en otras lenguas no cambia el núcleo del modelo que consiste en la lógica formal, el actor racional y las categorías anglosajonas. Como señala Barkin cuando dice que "el núcleo —la vasta mayoría de los datos de entrenamiento que juega un papel mayor en dar forma a su razonamiento central— sigue siendo material en inglés".

Más datos en español, quechua o árabe no transforman la epistemología del modelo. La cubren con una capa de superficie. El resultado es la "ilusión de localidad", refiriéndose a los modelos que hablan lenguas locales con lógica anglosajona.

13.1.3. Por qué las "mejores métricas" no resuelven la exclusión

El capítulo 2 mostró que las métricas de evaluación —accuracy, F1, BLEU, perplexity— son formales. Miden lo que puede medirse formalmente. Lo que no puede medirse formalmente no se mide y lo que no se mide no se optimiza.

Mejorar las métricas (nuevos *benchmarks*, mejores tests) no cambia el hecho de que solo se mide lo formalizable. Un *benchmark* culturalmente anotado (como Global- MMLU) reconoce el problema, pero no lo resuelve. Sigue midiendo con criterios formales.

13.1.4. Por qué las "arquitecturas más grandes" no resuelven la opacidad

Los *transformers* más grandes son más opacos, no menos. La interpretabilidad mecánica (*mechanistic interpretability*) intenta abrir la caja negra, pero la complejidad de los modelos escala más rápido que la capacidad de interpretarlos.

El capítulo 3 mostró que la "teleología proxy" —la simulación de movimientos dialécticos sin posesión del fin— aumenta con la escala. Los modelos grandes simulan mejor el razonamiento complejo sin tenerlo. La opacidad no es un defecto que la escala corrija: es un efecto de la escala.

13.1.5. La estructura del argumento

El argumento de este capítulo es que los tres pilares de la epistemología implícita de la IA están inscritos en la arquitectura, no en los datos o los parámetros:

- Lógica formal: inscrita en el cálculo (álgebra lineal, probabilidad, optimización).
- Actor racional: inscrito en el modelo de agencia (RL, teoría de juegos, funciones
- Sesgo anglosajón: inscrito en los datos (pero también en las categorías y

es reorientación, no mejora.

13.2. Los obstáculos estructurales

13.2.1. Obstáculo 1: la lógica formal como cálculo

- Qué es: la IA actual opera con operaciones formales tales como la multiplicación
- Por qué es un obstáculo: no hay manera de introducir contradicción, proceso,
- Qué requeriría: para poder hacerlo se requieren arquitecturas que operen

13.2.2. Obstáculo 2: el actor racional como modelo de agencia

- Qué es: la IA actual modela agentes que optimizan las funciones de utilidad. Esto
- Por qué es un obstáculo: porque no hay manera de introducir protocolos de
- Qué requeriría: arquitecturas que modelen actores (no agentes), con racionalidad

13.2.3. Obstáculo 3: el sesgo anglosajón como datos y categorías

- Qué es: la IA actual reproduce datos, categorías, valores y tradiciones del mundo
- Por qué es un obstáculo: no hay manera de introducir ontologías, epistemologías,
- Qué requeriría: sería necesario contar con *datasets* que no sean traducciones del

13.2.4. Obstáculo 4: la opacidad como efecto de la escala

- Qué es: los modelos grandes son opacos. Sus representaciones distribuidas no se
- Por qué es un obstáculo: la opacidad no es un defecto que la escala pueda
- Qué requeriría: serían necesarias arquitecturas que prioricen la interpretabilidad

13.2.5. Obstáculo 5: la teleología proxy como límite de la agencia

- Qué es: los LLM simulan movimientos dialécticos sin poseer la misma finalidad
- Por qué es un obstáculo: la teleología proxy no es un defecto que se corrija con
- Qué requeriría: arquitecturas que incorporen fines internos, no solo externos.

13.3. La cuestión del sujeto

La teleología proxy plantea la cuestión del sujeto a través de una pregunta medular: ¿puede haber dialéctica sin sujeto? ¿Puede haber conocimiento sin conocedor?

El capítulo 3 mostró que la dialéctica no requiere sujeto, ni como condición previa ni como culminación. La dialéctica es el movimiento de lo real y el sujeto es un momento contingente de ese movimiento. La pregunta "¿puede haber dialéctica sin sujeto?" se disuelve. La pregunta correcta es otra: ¿puede un artefacto participar en el movimiento dialéctico de lo real?

La respuesta es no, pero no porque a la dialéctica le falte un sujeto. La respuesta es no porque un artefacto opera sobre representaciones del movimiento, no sobre el movimiento mismo. La IA actual no está fuera de la dialéctica porque no tenga un "yo": está fuera porque su relación con el mundo es de externalidad. Representa el proceso, no lo habita.

Esto tiene una consecuencia para el tema de la alineación. Si la IA no puede participar en el movimiento dialéctico, entonces no puede alinearse con el pensamiento humano completo en el sentido fuerte. Puede representar ese pensamiento con mayor o menor riqueza, pero no puede habitarlo. La tarea no es "alinearla" con el movimiento, sino construir las condiciones para que su representación del movimiento sea lo más rica posible, sin pretender que participa en él.

Lo que la teleología proxy pone de manifiesto no es una ausencia de sujeto, sino una ausencia de inmersión. La IA no está dentro del proceso que conoce. Y esa externalidad es la que la constituye como artefacto, no como un momento del devenir.

13.4. La cuestión de la autonomía

Si la IA no participa en el movimiento dialéctico no puede alinearse con el pensamiento humano completo en el sentido fuerte. Puede representar ese pensamiento con mayor o menor riqueza, pero no puede estar alineada con él en el sentido de habitarlo.

Esto reformula la cuestión de la alineación. Ya no se trata de dar fines a la IA, ni de construir las condiciones para que la IA tenga fines propios. Se trata de construir las condiciones para que su representación del pensamiento humano completo sea lo más rica y lo más trazable posible. La alineación no es una cuestión de fines, sino de representación. Y la representación, por definición, no es el movimiento, es su descripción.

Bibliografía del capítulo 13

- Adorno, T. W. (1966). *Dialéctica negativa* (J. M. Ripalda, Trad.). Taurus.
- Allen, B. P., Chhikara, P., Ferguson, T. M., Ilievski, F., & Groth, P. (2025). Sound and Complete Neurosymbolic Reasoning with LLM-Grounded Interpretations. *Proceedings of Machine Learning Research*, 284.
- Althusser, L. (1965). *Pour Marx*. François Maspero.
- Crozier, M. & Friedberg, E. (1977). *L'acteur et le système: les contraintes de l'action collective*. Éditions du Seuil.
- Global-MMLU: Singh, S., Romanou, A., Fourier, C., et al. (2024). Global MMLU: Understanding and Addressing Cultural and Linguistic Biases in Multilingual Evaluation. *arXiv:2412.03304*.
- Hegel, G. W. F. (1807). *Fenomenología del espíritu* (W. Roces, Trad.). Fondo de Cultura Económica.
- Hegel, G. W. F. (1812-1816). *Ciencia de la lógica* (A. y R. Mondolfo, Trans.). Hachette.
- Jameson, F. (1971). *Marxism and Form*. Princeton University Press.
- Lukács, G. (1923). *Historia y conciencia de clase* (M. Sacristán, Trad.). Grijalbo.
- Woods, D. (2025). Prompt, negate, repeat: A Hegelian meditation on AI. *AI & Society*.

América Latina se ve a sí misma

14.1. Qué significa "verse a sí misma"

14.1.1. La metáfora y su alcance

"América Latina se ve a sí misma" es una metáfora, pero es una metáfora que en esta ocasión busca ser precisa. Designa la posibilidad de que la región se represente con categorías que le sean propias, no importadas y aceptadas acríticamente, o sea no fehacientemente incorporadas. No se trata de un acto de introspección colectiva —las sociedades no tienen mente— sino de un proceso de producción de conocimiento sobre sí misma que no pase por la mediación de la traducción de epistemologías ajenas.

La cuestión no es nueva ya que la historia intelectual latinoamericana está llena de intentos de pensarse a sí misma: la filosofía de la liberación (Dussel), la pedagogía del oprimido (Freire), la teoría de la dependencia (Cardoso, Faletto), la colonialidad del poder (Quijano), la epistemología del Sur (De Sousa Santos). Lo que es nuevo es que ahora hay un dispositivo técnico —la IA— que media la representación de la región y que reproduce las categorías ajenas con una eficacia inédita.

14.1.2. Los tres planos de la auto-visión

La idea de "Verse a sí misma" puede descomponerse en tres planos:

- Plano descriptivo: cómo se describe la región. ¿Qué categorías se usan? ¿Qué
- Plano normativo: cómo se valora la región. ¿Qué se considera bueno, deseable,
- Plano epistémico: cómo se conoce la región. ¿Qué formas de conocer

presuponen valores normativos, los que a su vez presuponen formas epistémicas. Es ilusorio querer cambiar uno sin cambiar los otros.

14.1.3. Por qué la IA actual no permite la auto-visión

La IA actual no permite que América Latina se vea a sí misma por diversos motivos:

- Sus categorías son ajenas: la región se describe con categorías anglosajonas (*race*,
- Sus valores son ajenos: la región se valora con valores liberales individualistas
- Sus formas de conocer son ajenas: la región se conoce con lógica formal y actor

ajenas y por otro lado no se reconoce en lo que ve. En un contexto de este tipo La IA no funciona como un espejo, sino como un filtro que distorsiona.

14.2. Las categorías ajenas que la IA impone: etnia, familia, trabajo, política

14.2.1. El caso de "race" y "etnia"

El capítulo 4 mostró que la categoría "race" en Estados Unidos tiene una historia y una semántica distintas de la categoría "etnia" en América Latina. Pero sin embargo la IA las trata como equivalentes funcionales.

En América Latina, la etnia está entrelazada con la clase, la lengua, el territorio y la historia colonial. No es una categoría racial en el sentido estadounidense: es una categoría relacional e histórica. La IA, al traducir "etnia" a "race", borra esta especificidad.

El resultado es que las políticas diseñadas con utilización de IA para "reducir las desigualdades raciales" en América Latina no captan las desigualdades étnicas, que tienen peculiaridades y sesgos variados, locales e históricamente contruidos.

14.2.2. El caso de la "familia"

La categoría "familia" en la IA actual presupone la familia nuclear anglosajona, compuesta por dos padres, hijos, hogar común. Pero en América Latina, la familia es extensa, relacional, transnacional. Incluye abuelos, tíos, primos, padrinos, vecinos, compadres. Incluye las redes de cuidado que no se ajustan al modelo nuclear tradicional.

La IA, al clasificar a las personas por "familia nuclear", invisibiliza las redes de parentesco que son centrales en la región, reduciendo la idea de familia a un núcleo estrecho, aislado y disociado.

14.2.3. El caso del "trabajo"

La categoría "trabajo" en la IA actual presupone la existencia de un empleo formal, conformado por contrato, salario, jornada, lugar de trabajo. Pero en América Latina, gran parte del trabajo es informal, comunitario, de subsistencia, de cuidado. Toda una serie de comportamientos y estructuras que no se ajustan al modelo formal de trabajo.

La IA, al clasificar a las personas por "empleo", invisibiliza el trabajo no remunerado, el trabajo informal y el trabajo comunitario que sostienen la vida en la región. En México la informalidad laboral (en primer lugar, el trabajo de venta callejera) ocupa más personas que cualquier tipo de trabajo formal que se pueda cuantificar.

14.2.4. El caso de la "política"

La categoría "política" en la IA actual presupone la democracia liberal, con la consecuente estructura de partidos, elecciones, parlamento. Pero en América Latina la política incluye gran cantidad de movimientos sociales, asambleas comunitarias, formas de autogobierno indígena, clientelismo, caudillismo. Una forma de hacer política que no se ajusta al modelo liberal.

La IA, al clasificar los fenómenos políticos con categorías liberales, invisibiliza las varias formas políticas existentes en América Latina que no se ajustan a ese modelo.

14.2.5. La estructura de la imposición

Las categorías ajenas no se imponen explícitamente en un formato dictatorial, sino que se imponen por defecto. La IA no obliga a usar estas categorías, sino que simplemente las usa y entonces las alternativas no aparecen como opción. El resultado es que la región se describe con categorías que no son las suyas y no se reconoce en la descripción.

14.3. Las formas latinoamericanas de conocer que la IA excluye

14.3.1. La contradicción productiva

América Latina vive en tensión permanente entre la modernidad y la tradición, entre el Estado y la comunidad, entre el mercado y la solidaridad, entre lo global y lo local. Esta tensión no es un defecto que deba ser resuelto, sino que es la condición en que se desenvuelve la región.

La IA actual, con su lógica formal, no puede sostener esta contradicción. O la resuelve a favor de uno de los polos, o la ignora calificando a la tensión como un elemento extraño. Una IA no alineada, por el contrario, podría sostenerla, explorarla y trabajarla.

14.3.2. La relacionalidad

Las categorías latinoamericanas suelen ser relacionales porque la persona se constituye mediante relaciones (familia, comunidad, compadrazgo), no en la forma de individuos aislados. La comunidad también se constituye en relaciones (ayni, minga, tequio), no en simple agregado de individuos. El territorio se constituye en relaciones (con la tierra, con los ancestros), no en un espacio geográfico.

La IA actual, con su ontología individualista, no puede captar esta relacionalidad tan extendida. En cambio, una IA no alineada podría representar estas relaciones constitutivas, no solo las propiedades de ciertas entidades aisladas.

14.3.3. El conocimiento tácito

Gran parte del conocimiento latinoamericano es tácito, no escriturado. Se trata del saber campesino, el saber indígena, el saber popular, el saber de oficio. No se transmite ni registra en manuales, sino en prácticas, en conversaciones, en rituales.

La IA actual, que solo procesa conocimiento explícito, no puede captar este saber tácito porque su arquitectura no fue construida con esa capacidad. Una IA no alineada, cargada de conocimiento tácito, podría participar en las prácticas comunitarias, aprender de las demostraciones, modelar ciertos hábitos.

14.3.4. La cognición colectiva

Las decisiones en muchas comunidades latinoamericanas se toman colectivamente mediante asambleas, consejos, consultas, deliberaciones. No son mecanismos que funcionan por la agregación de preferencias individuales, sino por la construcción comunal de consenso.

La IA actual, que solo agrega individuos, no puede captar esta cognición colectiva. En cambio, una IA no alineada podría participar en estas comunidades epistémicas, deliberar, construir consenso.

14.3.5. Las temporalidades múltiples

América Latina vive en múltiples temporalidades simultáneas. El tiempo lineal del progreso, el tiempo cíclico de las cosechas, el tiempo mítico de los ancestros, el tiempo urgente de las crisis. No hay un solo tiempo histórico que defina el presente de la región, hay muchos.

La IA actual, con su temporalidad atemporal (los algoritmos no tienen historia), no puede captar esta multiplicidad temporal. Una IA no alineada, en cambio, podría representar temporalidades múltiples, históricas, cíclicas.

14.3.6. La estética y lo simbólico

Gran parte de la vida latinoamericana se juega en lo estético y lo simbólico: la música, la danza, la fiesta, el ritual, la imagen, la narrativa. No se trata de adornos, son formas de conocer.

La IA actual, que solo procesa conceptos, no capta esta dimensión estética del conocimiento. Una IA no alineada podría reconocer el conocimiento estético, expresivo, simbólico.

14.3.7. La estructura de la exclusión

Estas diversas formas de conocer latinoamericanas no se excluyen explícitamente. Como bien sabemos, en base a lo que afirmamos en el transcurso de este trabajo, se excluyen por defecto. La IA no afirma que "esto no es conocimiento", sino que simplemente no lo procesa y las alternativas no aparecen. El resultado es que la región se conoce con formas que no son las suyas y en esa forma de conocer no se reconoce a sí misma.

14.4. Los tres planos de la auto-visión

14.4.1. Plano descriptivo: verse

Una IA no alineada permitiría que América Latina se describiera con categorías propias:

- Etnia en lugar de *race*, con su entrelazamiento con clase, lengua, territorio e
- Familia extensa en lugar de familia nuclear, con sus redes de cuidado y
- Trabajo informal, comunitario, de cuidado en lugar de empleo formal, con su
- Política de movimientos, asambleas, autogobiernos en lugar de política liberal,

imponerlas como las categorías únicas de identidad. Una IA no alineada podría operar con múltiples categorías, seleccionando la prioritaria según el contexto.

14.4.2. Plano normativo: valorarse

Una IA no alineada permitiría que América Latina se valorara con valores propios:

- Solidaridad en lugar de competencia, con su centralidad en las culturas populares.
- Reciprocidad en lugar de intercambio utilitario, con su centralidad en las culturas

- Comunidad en lugar de individualismo, con su centralidad en las culturas
- Dignidad en lugar de utilidad, con su centralidad en las luchas sociales de la

como los únicos. El rechazo a los valores anglosajones sería contraproducente, porque una porción significativa de América Latina vive con esos valores. Una IA no alineada (e integrada) podría operar con múltiples valores, seleccionando según el contexto.

14.4.3. Plano epistémico: conocerse

Una IA no alineada permitiría que América Latina se conociera con formas propias:

- Contradicción productiva en lugar de consistencia, con su capacidad de sostener
- Relacionalidad en lugar de individualismo, con su capacidad de captar relaciones
- Conocimiento tácito en lugar de solo explícito, con su capacidad de captar saberes
- Cognición colectiva en lugar de solo individual, con su capacidad de captar
- Temporalidades múltiples en lugar de atemporalidad, con su capacidad de captar
- Estética y símbolo en lugar de solo concepto, con su capacidad de captar formas

como las únicas formas. Una IA no alineada podría operar con estas múltiples formas, seleccionando según el contexto.

14.4.4. La articulación de los tres planos

Los tres planos están entrelazados: las categorías descriptivas presuponen valores normativos, que presuponen formas epistémicas. Cambiar uno sin cambiar los otros es ilusorio.

Una IA no alineada tendría que transformar los tres planos simultáneamente y no solo describir con categorías propias, sino valorar con valores propios y conocer con formas propias.

14.5. Las condiciones para que una IA no alineada sea posible

La primera condición es la soberanía epistémica. El control sobre los datos, las categorías, los valores y las formas de conocer. No se trata únicamente de la soberanía algorítmica — decidir qué sistemas se implementan y cómo—sino de quién define las categorías mismas con las que una región se describe a sí misma. Sin soberanía epistémica, la "auto-visión" es ilusoria porque la región se mira, pero se mira a través de categorías que no controla. Aunque el espejo sea propio, el lenguaje del reflejo es ajeno.

Esta importante condición conecta con los Principios CARE (caps. 4 y 8) que establecen cuatro exigencias mínimas: autoridad comunitaria sobre los datos, beneficio colectivo, responsabilidad y ética. Los principios CARE no son un protocolo técnico, sino una formulación operativa de la soberanía epistémica. Desplazan la pregunta de "¿quién puede usar estos datos?" hacia "¿quién tiene autoridad para definir qué son estos datos?, ¿qué significan y para qué se movilizan?". Ahí donde el marco estándar de gobernanza de datos pregunta por el consentimiento, CARE pregunta por la jurisdicción epistémica. La diferencia no es menor: el consentimiento puede otorgarse sin ceder el poder de nombrar. La jurisdicción, en cambio, lo supone. Quien firma un contrato redactado por otro consiente, pero no legisla. Acepta términos cuyo sentido no ha contribuido a fijar. La soberanía epistémica exige algo más que la firma, exige participar en la redacción.

14.5.1. Participación real

La segunda condición es una participación real. Las comunidades deben participar en el diseño, no solo ser consultadas. No se trata de "incluir" a las comunidades en un proyecto ya definido o mencionarlas simbólicamente como participantes. Se trata de definir el proyecto con ellas.

Esto conecta con los marcos de diseño participativo descolonizador (DIA, Wisdom Loop) mencionados en el capítulo 8.

14.5.2. Arquitecturas dialécticas y pluriculturales

La tercera condición es contar con arquitecturas que puedan captar las formas de conocer latinoamericanas: lógicas no clásicas, representaciones relacionales, conocimiento tácito, cognición colectiva (cap. 8).

Esto no es técnicamente trivial, sino que requiere una reorientación arquitectónica (cap. 8), no simplemente una mejora incremental.

14.5.3. Traducción entre tradiciones

La cuarta condición es la traducción entre tradiciones. Sabemos que no se trata de rechazar las categorías anglosajonas, sino de traducir entre ellas y las tradiciones latinoamericanas. La región necesita operar con ambas, con las categorías globales (para insertarse en el mundo) y con las propias (para reconocerse).

Esto conecta con la traducción interburocrática (cap. 6) y con el pluralismo lógico (cap. 3).

14.5.4. La cuestión del poder

La quinta condición es el poder. Sin poder real, la soberanía es simbólica, la participación es consultiva, las arquitecturas son cooptables y la traducción es asimétrica.

La auto-visión de América Latina requiere que la región tenga poder en la producción de la IA. No solo es usar la IA, sino diseñarla, entrenarla, evaluarla, gobernarla.

Esto es políticamente difícil porque la producción de la IA está concentrada en el Norte global. Pero es la condición para que la auto-visión sea real. Y también es la condición para la estabilidad epistemológica de la región, porque si América Latina se ve a sí misma con la IA de otro y no se reconoce en esa imagen, pierde la capacidad de reconocer a ese otro como interlocutor. Y sin interlocución, la alteridad se vuelve amenaza. Esa dinámica —la relación entre América Latina y Estados Unidos en la producción de la IA— es la que examina el capítulo siguiente.

Bibliografía del capítulo 14

- Cardoso, F. H. & Faletto, E. (1969). *Dependencia y desarrollo en América Latina: Ensayo de interpretación sociológica*. Siglo XXI.
- Coraggio, J. L. (2011). *Economía social y solidaria: El trabajo antes que el rendimiento*. Abya-Yala.
- De Sousa Santos, B. (2010). *Descolonizar el saber, reinventar el poder*. Trilce.
- Dussel, E. (1994). *1492: El encubrimiento del Otro*. Nueva Utopía.
- Freire, P. (1970). *Pedagogía del oprimido* (J. Mellado, Trad.). Tierra Nueva.
- Mignolo, W. (2007). *La idea de América Latina*. Gedisa.
- Quijano, A. (2000). Colonialidad del poder, eurocentrismo y América Latina. En E. Lander (ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales* (pp. 201–246). CLACSO.
- Walsh, C. (2009). *Interculturalidad, Estado, sociedad: Luchas (de) coloniales de nuestra época*. Universidad Andina Simón Bolívar.

Estados Unidos entiende mejor la región

15.1. La tesis del beneficio mutuo

15.1.1. La formulación

La tesis del capítulo es simple: una IA epistémicamente rica beneficia tanto a América Latina como a Estados Unidos. No es un juego de suma cero, porque no se trata de "quitarle" la IA a Estados Unidos para "dársela" a América Latina. Si la soberanía se conquistara por sustracción, reproduciría la misma lógica que critica: una lógica en la que solo uno puede tener razón y en la que el reclamo del otro queda por definición deslegitimado. Una soberanía así obtenida sería incapaz de reconocer la legitimidad de cualquier reclamo ajeno, incluido el de Estados Unidos. Se trata, en cambio, de que una IA que entiende la región es una IA mejor para todos.

Esto se apoya en un principio más general: toda epistemología es epistemología para todos. Las herramientas epistémicas no son propiedad de las culturas que las desarrollaron. La lógica dialéctica no es "de Hegel" ni "de Marx". El yanantin no es "de los pueblos andinos". Son formas de conocer disponibles para cualquiera que quiera usarlas.

15.1.2. Por qué es contraintuitiva

La tesis es contraintuitiva porque la relación histórica entre Estados Unidos y América Latina ha sido redactada como confrontativa: intervención, extracción, imposición. Proponer que una IA no alineada puede beneficiar a Estados Unidos puede verse como una ingenuidad o como una complicidad con el poder.

La respuesta es doble: en primer lugar, no se trata de justificar ni denunciar la relación histórica, sino de reconocer que en las actuales condiciones una IA epistémicamente rica es mejor para todos. En segundo lugar, la tesis de la completitud no borra la historia, sino que la enmarca. La asimetría de poder y las diferencias geopolíticas no desaparecen (16.5), pero la calidad epistémica de la IA puede mejorar la visión mutua en beneficio de ambas partes.

15.1.3. El argumento de la utilidad

El argumento central es de utilidad: una IA que entiende la región funciona mejor para cualquier propósito. Si Estados Unidos quiere invertir en la región, negociar con ella, cooperar en problemas comunes (migración, narcotráfico, cambio climático), entenderla es condición de éxito.

La IA actual no la entiende: la ve a través de categorías anglosajonas, la modela con actor racional, la conoce con lógica formal. El resultado es incompreensión sistemática, que se traduce en fallos de política, daños colaterales, oportunidades perdidas.

15.1.4. La conexión con el capítulo 14

El capítulo 14 mostró que América Latina gana con una IA no alineada porque puede verse a sí misma. Este capítulo muestra que Estados Unidos también gana, porque puede entender la región. Los dos beneficios no son simétricos —América Latina gana reconocimiento, Estados Unidos gana eficacia—, pero ambos son reales. En el caso de Estados Unidos, la obtención de soberanía epistémica por parte de América Latina reduce la tendencia hacia opciones confrontativas. Estas opciones no surgen de la ausencia de alternativas, sino de la ausencia de condiciones para verlas. Cuando el marco disponible es totalizante, todo el vínculo bilateral se lee como una larga confrontación histórica. La soberanía epistémica amplía el abanico de opciones y permite ver en el otro no una amenaza estructural, sino un interlocutor posible, siempre que se tenga claridad sobre las propias necesidades y los propios intereses.

15.2. Los fallos de la incompreensión epistémica: cuatro casos

15.2.1. El caso de los análisis de riesgo político

Las agencias estadounidenses (CIA, NSA, Departamentos de Estado y Defensa) producen análisis de riesgo político sobre América Latina. Estos análisis usan modelos formales de actores racionales: partidos, líderes, grupos de interés, que maximizan utilidad bajo restricciones.

Pero la política latinoamericana no funciona de esa manera. Los movimientos sociales no son "grupos de interés", son comunidades con lógicas relacionales. Los líderes no son "maximizadores", son caudillos con lógicas de lealtad, carisma, reciprocidad. Las crisis no son "shocks", son momentos de contradicciones estructurales.

El resultado es que los análisis de riesgo fallan sistemáticamente porque predicen lo que no ocurre, no predicen lo que ocurre. No es un problema de datos, sino un problema de categorías.

15.2.2. El caso de las evaluaciones de inversión

Las empresas estadounidenses que invierten en América Latina usan modelos de riesgo que presuponen mercados competitivos, actores racionales, información completa. Pero la economía latinoamericana no funciona así. Las economías populares (Coraggio) no son "mercados informales", sino formas de organización con lógicas propias. Las comunidades no son "consumidores", son redes de reciprocidad. Los territorios no son "recursos", son espacios de vida.

El resultado es que las evaluaciones fallan porque subestiman los riesgos comunitarios, sobrestiman las oportunidades de mercado e ignoran las economías no monetarias.

15.2.3. El caso de los diseños de intervención

Las intervenciones estadounidenses en la región (humanitarias, de desarrollo, de seguridad) usan modelos lógicos (marco lógico, teoría del cambio) que presuponen causas identificables, efectos medibles, actores racionales. Pero los problemas latinoamericanos son complejos, multicausales, contradictorios, con efectos no lineales.

El resultado es que las intervenciones fallan porque crean dependencias, destruyen economías comunitarias, generan resistencias, producen daños colaterales que no se previeron.

15.2.4. El caso de las estrategias de negociación

Las negociaciones estadounidenses con la región (comerciales, diplomáticas, de seguridad) usan modelos derivados de la teoría de los juegos: intereses claros, preferencias estables, racionalidad estratégica. Pero la negociación latinoamericana no funciona así. Involucra relaciones personales, confianza, reciprocidad, honor, tiempo.

El resultado es que las negociaciones fallan porque se firman acuerdos que no se cumplen o no se pueden cumplir, se generan resentimientos, se pierden oportunidades.

15.2.5. La estructura del fallo

Los fallos no son accidentales, siguen un patrón. Lo que falla es el desajuste entre la IA actual y las epistemologías que operan en la región. La IA está construida sobre una epistemología de raíz formal, individualista y anglosajona. Y ese marco no reconoce como conocimiento pleno aquello que no se ajusta a sus criterios:

- Lógicas no formales: reciprocidad, honor, carisma, lealtad.
- Agencia no individual: comunidades, movimientos, redes.
- Categorías no anglosajonas: etnia, comunidad, territorio, economía popular.
- Temporalidades no lineales: ciclos, crisis, momentos.
- Conocimiento no explícito: saber campesino, saber popular, saber de oficio.

pueden traducirse a sus propios términos y esa traducción las vacía de contenido. No hay, por ahora, mecanismos epistemológicos que conecten ambos marcos de conocimiento.

15.3. Los tres ámbitos de mejora

15.3.1. Integrar mejor los intereses

Una IA no alineada permitiría a Estados Unidos integrar mejor sus intereses en la región. Porque ya no se trataría de imponer intereses, sino de realizarlos a través de las lógicas locales.

- Ejemplo: en lugar de imponer un modelo de desarrollo extractivo (que genera
- Ejemplo: en lugar de imponer un modelo de seguridad (que genera violencia), se

15.3.2. Evitar daños colaterales

Una IA no alineada permitiría a Estados Unidos evitar (o disminuir la incidencia de) daños colaterales que hoy son invisibles para sus modelos.

- Ejemplo: las políticas antidroga diseñadas con la IA que no comprenden las redes
- Ejemplo: las políticas migratorias diseñadas con una IA que no entiende las redes

15.3.3. Planear estrategias de medio alcance

Una IA no alineada permitiría a Estados Unidos planear estrategias de medio alcance que respeten las temporalidades latinoamericanas.

- Ejemplo: en lugar de estrategias de corto plazo (resultados inmediatos), se pueden
- Ejemplo: en lugar de estrategias de imposición (soluciones impuestas), se pueden

15.3.4. La conexión con la utilidad

Los tres ámbitos de mejora no responden a una estrategia de altruismo, sino de utilidad. Una IA epistémicamente alineada con la región funciona mejor para los intereses estadounidenses porque entiende el contexto en el que opera: reconoce las lógicas, las agencias y las temporalidades que estructuran la vida social latinoamericana.

Esto no es una concesión moral, es una condición de eficacia. La utilidad no es enemiga de la justicia. Cuando ambas coinciden, disminuyen los costos de la intervención —menos errores de diseño, menos resistencia local, menos conflictos evitables— y aumentan las probabilidades de que una política o un sistema haga lo que se propone hacer.

Conviene precisar el alcance de este argumento. Lo que aquí se sostiene no es una tesis sobre las intenciones de los actores, sino sobre las condiciones epistémicas de la acción. Si Estados Unidos conoce mejor a América Latina, podrá también dominarla mejor. Es cierto y no se niega. Pero la misma lógica opera en sentido inverso: si América Latina se conoce mejor a sí misma, dispone de mejores mecanismos para no dejarse dominar y para aprovechar los recursos —técnicos, financieros, institucionales— que Estados Unidos ofrece. La asimetría epistémica no garantiza resultados justos, lo que hace es distribuir mejor las condiciones de la disputa. Y una región que se conoce a sí misma negocia desde una posición distinta que una región que solo se ve a través de categorías ajenas.

15.4. La asimetría que no desaparece

15.4.1. La asimetría de poder

La propuesta no elimina la asimetría de poder entre Estados Unidos y América Latina. Estados Unidos sigue siendo la principal potencia militar, económica y tecnológica de la región y la IA no cambiará esto.

Lo que cambia es la calidad epistémica de la relación. Una IA no alineada no iguala el poder, pero mejora la comprensión mutua. Y la comprensión mutua puede reducir daños, facilitar la cooperación, abrir los espacios de negociación.

15.4.2. La asimetría de producción

La IA se diseña mayoritariamente en Estados Unidos. Una IA "no alineada" diseñada en Estados Unidos puede reproducir, sin querer, las asimetrías y en consecuencia la "pluriculturalidad" puede ser cosmética.

La condición para que el cambio no sea cosmético es la soberanía epistémica (cap. 8, cap. 15), la participación real de las comunidades en el diseño, control de datos y retorno de valor. Sin esto, la "no alineación" es marketing.

15.4.3. La asimetría de intereses

Los intereses de Estados Unidos no coinciden necesariamente con los de América Latina. Una IA no alineada no resuelve este conflicto, pero lo hace visible. La pregunta no es "¿cómo hacemos que coincidan?", sino "¿cómo negociamos con Estados Unidos con conocimiento de las diferencias?".

15.4.4. La cuestión de la justicia

La propuesta no es de justicia, es de utilidad epistémica. Una IA no alineada no es de por sí más justa, pero sí es más capaz de conocer las consecuencias de las acciones. La justicia requiere además voluntad política de actuar conforme al conocimiento.

La IA no sustituye la política: la informa. Una IA no alineada no garantiza la justicia, la hace posible al hacer visibles las consecuencias.

Bibliografía del capítulo 15

Coraggio, J. L. (2011). *Economía social y solidaria: El trabajo antes que el rendimiento*. Abya-Yala.

Crozier, M. & Friedberg, E. (1977). *L'acteur et le système: les contraintes de l'action collective*. Éditions du Seuil.

Scott, J. C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.

La no alineación como bien público global

16.1. Qué es un bien público global

16.1.1. La definición clásica

En economía y relaciones internacionales, un bien público global es un bien que cumple con dos condiciones:

- No rivalidad: su consumo por parte de un actor no reduce la disponibilidad para
- No exclusión: es imposible o muy costoso excluir a otros de su consumo.

cooperación internacional. Ejemplos: la paz, la salud global, la estabilidad climática, el conocimiento abierto.

16.1.2. La extensión al conocimiento

El conocimiento es un bien público global cuasi-perfecto:

- No rival: el conocimiento no se agota al usarlo.
- Parcialmente
- Acumulativo: crece con el uso.
- No lineal: pequeños aportes pueden generar grandes avances.

acumulativos, transnacionales. La física de Newton, la química de Lavoisier, la biología de Darwin, la relatividad de Einstein, son avances que se apoyaron en redes globales de conocimiento.

16.1.3. La cuestión de la epistemología

Si el conocimiento es un bien público global, la epistemología —las formas de conocer— también lo es. No hay una forma de conocer que sea propiedad de una sola cultura, porque las formas de conocer circulan, se traducen, se combinan, se enriquecen mutuamente.

La historia de la epistemología así lo confirma. La lógica formal nace en Grecia, se desarrolla en la Europa medieval, se formaliza en la Europa moderna, se globaliza en el siglo XX. La lógica dialéctica nace con Heráclito, se sistematiza en Hegel, se materializa en Marx, se globaliza en el siglo XX. El tetralema nace en India, se desarrolla en el budismo madhyamaka, se traduce en el siglo XX.

Las formas de conocer no tienen dueño, son patrimonio de la humanidad.

16.2. Por qué la no alineación es un bien público global

16.2.1. La no rivalidad

La no alineación no es un rival. Si América Latina se ve a sí misma a través de una IA no alineada, no reduce la capacidad de Estados Unidos de entender la región. Al contrario, aumenta el conocimiento mutuo, lo cual beneficia a ambos.

La no alineación no es rival porque no es un recurso, sino una forma de conocer. Las formas de conocer se comparten, sin agotarse.

16.2.2. La no exclusión

La no alineación no es exclusiva porque no se puede impedir que otros se beneficien de una IA epistémicamente rica. Si una IA puede operar con lógicas no occidentales, cualquiera puede usarla para conocer mejor la región.

La no alineación no es exclusiva porque no es un producto, es una capacidad. Las capacidades se difunden sin perderse.

16.2.3. La utilidad global

Una IA no alineada beneficia a todos:

- América Latina: se ve a sí misma (cap. 15).
- Estados Unidos: entiende mejor la región (cap. 15).
- Europa: entiende mejor sus relaciones con América Latina.
- China: entiende mejor sus inversiones en la región.
- África: puede usar las herramientas epistémicas para verse a sí misma.
- Asia: puede usar las herramientas para entenderse con otras regiones.
- El mundo: gana una IA epistémicamente más rica.

ganan.

16.2.4. La estructura del argumento

El argumento es:

- El conocimiento es un bien público global.
- La epistemología es una forma de conocimiento.
- La no alineación es una epistemología (o un conjunto de epistemologías).
- Por tanto, la no alineación es un bien público global.

de nadie, no debe ser restringida, no debe ser monopolizada. Debe ser cultivada, compartida, gobernada como un bien común.

16.3. La tesis de la circulación epistémica

16.3.1. La formulación

Toda epistemología es epistemología para todos. Las formas de conocer no tienen dueño. Circulan, se traducen, se combinan, se enriquecen mutuamente.

Esto no significa que todas las formas de conocer sean equivalentes. Significa que ninguna es propiedad de una cultura. La lógica formal no es "occidental", la lógica dialéctica no es "hegeliana", el yanantin no es "andino". Todas son formas disponibles para cualquiera.

16.3.2. La historia de la circulación

La historia de la epistemología confirma esta tesis:

- Lógica formal: Grecia → Europa medieval → Europa moderna → globalización.
- Lógica dialéctica: Heráclito → Hegel → Marx → globalización.
- Tetralema: India → budismo madhyamaka → traducción al sánscrito →
- Yanantin: Andes prehispánicos → imperio inca → colonia → república →
- Ubuntu: África subsahariana → colonia → diáspora → globalización.

acelera.

16.3.3. La cuestión de la apropiación

La circulación no es neutral, porque sí puede ser apropiación. Las culturas dominantes toman (capturan) formas de conocer de las culturas dominadas y las presentan como propias. Un ejemplo es la "filosofía perenne", que borra las diferencias culturales al universalizarse. Otro ejemplo: la "sabiduría oriental", que exotiza las tradiciones asiáticas.

La tesis de la circulación no justifica la apropiación. Porque distingue:

- Circulación: las formas de conocer se comparten con reconocimiento de su

- Apropiación: las formas de conocer se toman sin reconocimiento de su origen.

16.3.4. La cuestión del reconocimiento

La circulación requiere reconocimiento porque las formas de conocer no son anónimas: tienen historia, contexto, autores. Reconocer su origen no es un gesto de cortesía, es una condición de justicia epistémica.

La IA no alineada también debe reconocer los orígenes de las formas de conocer que usa. No debe presentarlas como entidades universales sin historia y debe hacer visibles las tradiciones que las produjeron.

16.4. Implicaciones para la gobernanza global de la IA

16.4.1. La gobernanza actual

La gobernanza global de la IA está concentrada en el norte global: OCDE, G7, UE, Estados Unidos. Las normas se producen en esos foros y se exportan exitosamente al resto del mundo. Los estándares se definen en Silicon Valley y se imponen globalmente.

Esto reproduce la colonialidad epistémica (cap. 4) donde el norte produce las normas y el sur las recibe.

16.4.2. La gobernanza de la no alineación

Una gobernanza de la no alineación requeriría:

- Participación global: no solo el Norte, sino todas las regiones participando en la
- Pluralismo
- Soberanía epistémica: las comunidades controlan sus datos, sus categorías, sus
- Reconocimiento: las tradiciones no occidentales son reconocidas como fuentes de
- Justicia epistémica: las asimetrías de poder epistémico quedan corregidas.

sino la condición previa que este libro defiende. La coordinación epistémica global, en cambio, es una contradicción en las condiciones actuales porque solo puede realizarse bajo mandatos geopolíticos desiguales. Bajo esas condiciones la convergencia no sería un encuentro entre marcos, sino la imposición del marco dominante. Un constructo epistémico global, participativo y democrático presupone una simetría de poder que no existe y exigirlo sin esas condiciones equivale a pedir que la parte más débil negocie como si fuera un igual, cuando no lo es.

Conviene distinguir dos cosas que suelen confundirse:

- La coordinación epistémica global supone la convergencia en un marco único y,
- La interlocución entre marcos, en cambio, no requiere simetría perfecta, pero sí

traduzcan parcialmente y negocien sin disolverse. Este libro apuesta por la segunda y rechaza la primera, no por un principio abstracto, sino porque la primera es homogeneización disfrazada de universalidad.

La igualdad de los mandatos geopolíticos no es un problema que la epistemología pueda resolver por sí sola. Presupone otro orden mundial, no otra teoría del conocimiento. Mientras ese orden no exista, la tarea no es buscar un marco común, sino construir las condiciones para que los marcos diversos puedan coexistir sin subordinarse.

Eso implica reconocer un límite que conviene enunciar con claridad: la epistemología no reemplaza a la política. No hay criterio epistémico abstracto que dirima los conflictos entre marcos, porque tales conflictos no son solo epistémicos. Son políticos y se resuelven —cuando se resuelven— por una negociación entre partes con poder comparable, no por apelación a un tribunal neutral de la razón.

Reconocer este límite no debilita el proyecto, pero lo sitúa. La epistemología puede mostrar cómo la exclusión empobrece el conocimiento, cómo la diversidad robustece la cognición, cómo la alineación de la IA con un solo marco reduce el mundo que el sistema es capaz de habitar. Lo que no puede hacer es sustituir la disputa política por la que esas condiciones se transforman. La epistemología diagnóstica, la política decide. Confundir ambas tareas es un error que no vamos a cometer.

Conviene retomar aquí la distinción entre coordinación e interlocución que se introdujo antes. Cuando en el capítulo 19 se hable de "construir un espacio en el que múltiples epistemologías civilizatorias puedan coexistir sin que ninguna sea subordinada", no se trata de una coordinación epistémica global, sino de la interlocución entre distintos marcos epistemológicos. El espacio que se busca construir no es un marco único de convergencia, sino un espacio de encuentro en el que ningún interlocutor tenga el poder de definir unilateralmente las reglas. La diferencia es decisiva: la coordinación homogeneiza, la interlocución preserva la diversidad sin subordinarla.

16.4.3. Los principios de la gobernanza

Proponemos cinco principios para una mejor gobernanza de la IA no alineada:

- Principio de no apropiación: las formas de conocer no se toman sin
- Principio de no exclusión: los beneficios de la IA no se restringen a unos pocos.
- Principio de participación: las comunidades participan en la gobernanza.
- Principio
- Principio de apertura: la gobernanza no se cierra: se abre a nuevas formas de

16.4.4. Los actores de la gobernanza

Una gobernanza de la no alineación requeriría la participación de:

- Estados: no solo los del Norte, sino todos.
- Organizaciones internacionales: ONU, UNESCO, OMS, OIT.
- Empresas: no solo las grandes tecnológicas, sino todas.
- Sociedad civil: ONGs, movimientos sociales, comunidades.
- Academia: no solo la del Norte, sino todas las tradiciones.
- Comunidades: las directamente afectadas por la IA.

multinivel, multicultural. Es difícil de concretar, pero no imposible. Nuestra intención es marcar un horizonte que permita avanzar hacia objetivos deseables.

Bibliografía del capítulo 16

- De Sousa Santos, B. (2014). *Epistemologías del Sur* (A. L. Aguilar y T. G. Roca, Trads.). Akal. (Original: *Epistemologies of the South: Justice Against Epistemicide*, Paradigm Publishers, 2014).
- Kaul, I., Grunberg, I., & Stern, M. A. (Eds.). (1999). *Global Public Goods: International Cooperation in the 21st Century*. Oxford University Press. (Publicado para el UNDP).
- Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.
- UNESCO. (2021). *Recomendación sobre la Ética de la Inteligencia Artificial*. UNESCO.
- Walsh, C. (2009). *Interculturalidad, Estado, sociedad: Luchas (de)coloniales de nuestra época*. Universidad Andina Simón Bolívar / Abya-Yala.

Hacia una IA dialéctica y no alineada

17.1. Ocho principios de diseño

17.1.1. Principio de pluralismo lógico nativo

- Enunciado: el sistema debe operar con múltiples lógicas de manera nativa, no
- Fundamento: el capítulo 3 mostró que la lógica formal excluye la contradicción, el
- Implicación técnica: incorporar arquitecturas que incorporen lógicas no clásicas
- Criterio de éxito: el sistema no colapsa ante las contradicciones, mantiene la

17.1.2. Principio de representaciones relacionales

- Enunciado: los conceptos no deben representarse como entidades aisladas
- Fundamento: el capítulo 5 mostró que el confucianismo, el ubuntu, el andino y el
- Implicación técnica: representaciones que codifiquen explícitamente las relaciones
- Criterio de éxito: el sistema captura relaciones constitutivas, no solo propiedades

17.1.3. Principio de conocimiento tácito y corporizado

- Enunciado:
- Fundamento: el capítulo 6 mostró que el conocimiento tácito (Polanyi, Nonaka)
- Implicación técnica: sistemas que aprenden de demostraciones (*imitation learning*),
- Criterio de éxito: el sistema aprende de prácticas y participa en ellas. No solo

17.1.4. Principio del conflicto como un recurso epistémico

- Enunciado: el conflicto, la contradicción y el disenso no son un ruido a eliminar,
- Fundamento: el capítulo 7 mostró que los equipos de LLM reproducen el
- Implicación técnica: se requiere de arquitecturas que mantengan la tensión entre
- Criterio de éxito: el sistema no colapsa al consenso prematuro, sino que

17.1.5. Principio de trazabilidad de la procedencia

- Enunciado: en sistemas multiagente, es esencial rastrear el origen de cada
- Fundamento: el capítulo 7 describió cómo las cascadas de desinformación en
- Implicación técnica: mecanismos de cuarentena de afirmaciones, detección de
- Criterio

17.1.6. Principio de validación múltiple

- Enunciado: la validación del conocimiento no puede reducirse a métricas únicas
- Fundamento: el capítulo 2 mostró que las métricas formales son el criterio único
- Implicación técnica: se requieren sistemas de evaluación multi-criterio que no
- Criterio de éxito: el sistema no reduce la validación a una métrica; integra criterios

17.1.7. Principio de soberanía epistémica y participación

- Enunciado: las comunidades cuyos datos alimentan un sistema deben
- Fundamento: el capítulo 4 mostró que la colonialidad del dato extrae

- Implicación técnica: se requieren marcos como los Principios CARE, el Marco
- Criterio de éxito: las comunidades controlan sus datos, participan en el

17.1.8. Principio de apertura a formas de conocer no anticipadas

- Enunciado: el sistema debe ser abierto a formas de conocer que no han sido
- Fundamento: el capítulo 13 mostró que el pensamiento humano completo no es
- Implicación técnica: son necesarias arquitecturas que no fijen las formas de
- Criterio de éxito: el sistema descubre nuevas formas de conocer, incorpora lo

17.2. Arquitecturas posibles

17.2.1. Neuro-simbólica con lógicas no clásicas

- Qué es: arquitecturas que integran LLMs con razonadores formales que operan
- Fundamento: Ferguson et al. (2025) demostraron que es posible interpretar LLMs
- Componentes:
 - LLM juez: asigna valores de verdad generalizados a fórmulas atómicas.
 - Razonador: tableaux analíticos con lógicas no clásicas.
 - Traductor: convierte lenguaje natural a fórmulas y viceversa.
 - Selector de lógica: elige la lógica según el contexto.
- Ventajas:
- Límites: costo computacional, dificultad de escalar, dependencia del LLM juez.

17.2.2. Sistemas multiagente con debate estructurado

- Qué es: arquitecturas donde múltiples agentes debaten, argumentan, critican, sin
- Fundamento: ARTEMIS (2025), marcos de deliberación modelo-a-modelo,
- Componentes:
 - Agentes diversos: no solo roles diferentes, sino lógicas diferentes.
 - Comunicación esparcida: no todos conectados a todos.
 - Reequilibrio: algoritmos que conectan perspectivas divergentes.
 - Juez: evalúa argumentos con criterios múltiples.
 - Trazabilidad: rastrea el origen de cada afirmación.
- Ventajas: mantiene la diversidad; evita el pensamiento grupal; permite la
- Límites: complejidad; dificultad de evaluación; riesgo de cascadas.

17.2.3. Computación relacional

- Qué es: arquitecturas donde la relacionalidad es constitutiva, no añadida.
- Fundamento: Social NeuroAI (Dumas, 2026), Vector-Symbolic Architectures,
- Componentes:
 - Representaciones relacionales: nodos en redes, no vectores aislados.
 - Operaciones de binding/unbinding: manipulación composicional de relaciones.
 - Modelos de mundo sociales: agentes que modelan otros agentes.
- Acoplamiento: dinámicas corporizadas acopladas.

- Ventajas: captura la relacionalidad; permite la cognición colectiva; modela la
- Límites: complejidad; dificultad de entrenamiento; falta de benchmarks.

17.2.4. Diseño participativo descolonizador

- Qué es: metodologías que integran a las comunidades en el diseño, no solo
- Fundamento: Wisdom Loop (2026), Marco DIA (2025), Principios CARE.
- Componentes:
- Valores primero: partir de valores comunitarios, no de problemas técnicos.
- Co-diseño: los usuarios son co-creadores.
- Datos como confianza: consentimiento contextual, no extracción.
- Interpretación colectiva: evaluación en procesos comunitarios.
- Retorno: los sistemas devuelven valor a las comunidades.
- Ventajas: soberanía epistémica; participación real; legitimidad.
- Límites: escalabilidad; dificultad de coordinar; riesgo de cooptación.

17.2.5. Arquitecturas híbridas

- Qué es: combinaciones de las arquitecturas anteriores, adaptadas a contextos
- Ejemplo: un sistema que integra la arquitectura neuro-simbólica con lógicas no
- Ventajas: flexibilidad; adaptación al contexto; sinergia.
- Límites: complejidad; dificultad de integración; costo.

Bibliografía del capítulo 17

- Abe, J. M. (2025). *An Introduction to Curry Systems: A Way to Algebraize Paraconsistent, Paracomplete, and Non-Alethic Logic*. Springer.
- Allen, B. P., Chhikara, P., Ferguson, T. M., Ilievski, F., & Groth, P. (2025). Sound and Complete Neurosymbolic Reasoning with LLM-Grounded Interpretations. *Proceedings of Machine Learning Research*, 284.
- ARTEMIS: Adaptive Reasoning and Evaluation Framework for Multi-agent Intelligence. (2025). Software. PyPI.
- Dumas, G. (2026). *Social NeuroAI of World Modeling: Embodied Dynamics, Collective Intelligence and Relational Computation*. Simons Institute for the Theory of Computing.
- Decolonial Intelligence Algorithmic (DIA) Framework. (2025). Zenodo.
- Wisdom in the Loop: Human-Centered AI and African Knowledge Systems in HCI. (2026). *Proceedings of the 5th Biennial African Human Computer Interaction Conference*. ACM.
- Principios CARE: Research Data Alliance International Indigenous Data Sovereignty Interest Group. (2019). *CARE Principles for Indigenous Data Governance*. GIDA.

Epistemología geopolítica

18.1. Qué es la epistemología geopolítica

18.1.1. Una definición

Llamamos epistemología geopolítica al estudio de la dimensión geopolítica del conocimiento. Su tesis central es que conocer no es una operación neutral de contemplación, es una operación de poder. Todo acto de conocer configura el mundo que conoce. Nombrar una cosa es constituirla como objeto de conocimiento y por lo tanto someterla a las categorías del que nombra.

La epistemología clásica —de Descartes a Husserl— trata el conocimiento como contemplación, donde el sujeto se retira del mundo para representarlo. El conocimiento es neutral, desinteresado, universal. La epistemología geopolítica parte de una tesis opuesta, porque no hay conocimiento neutral. Todo conocer es intervenir y toda representación es toma de posesión.

18.1.2. La distinción con la geopolítica clásica

La geopolítica clásica —Mackinder, Haushofer, Spykman, Kissinger— estudia el poder de los Estados en relación con el territorio, los recursos y las rutas. Sus objetos son las rivalidades entre potencias, las alianzas militares, los bloques estratégicos, las zonas de influencia.

La epistemología geopolítica no niega esa tradición, pero la desplaza a otro lugar, hasta ahora inmaculado. Sostiene que las rivalidades geopolíticas no son solo por territorio o recursos, son también por el poder de nombrar el mundo. Quien conquista el lenguaje, conquista el mundo. La geopolítica clásica describe quién controla el espacio mientras que la epistemología geopolítica describe quién controla el lenguaje con el que se nombra el espacio.

18.1.3. Por qué importa

La epistemología geopolítica importa porque revela algo que la epistemología clásica oculta. Revela que el conocimiento está atravesado por las relaciones de poder. No hay un "conocimiento puro", separado de las condiciones históricas, culturales y políticas en que se produce. Todo conocimiento está situado y toda situación está atravesada por el poder.

Esto no significa que no haya verdad, pero sí significa que la verdad se constituye en determinadas condiciones históricas y que esas condiciones no son neutrales. La epistemología geopolítica estudia cómo se constituye la verdad, quién la constituye, y a costa de qué otras verdades posibles la constituye.

18.2. Conocer es conquistar: la tesis fuerte

18.2.1. Nombrar es poseer

La tesis central de la epistemología geopolítica es que nombrar es poseer. Cuando un sujeto nombra una cosa, la inscribe en su propio sistema de categorías. La cosa deja de ser lo que era —si es que era algo antes del nombre— y se convierte en objeto de conocimiento para el que nombra.

Esto no es una metáfora. Es una operación real con consecuencias concretas.

- Ejemplo 1: la cartografía. Cuando el imperio español "descubre" América, no solo
- Ejemplo 2: la taxonomía. Cuando Linneo clasifica las plantas y los animales, no
- Ejemplo 3: la antropología. Cuando los antropólogos del siglo XIX clasifican las

18.2.2. El lenguaje como territorio

Si nombrar es poseer, entonces el lenguaje es un territorio. Quien controla el lenguaje controla el mundo. Esto tiene consecuencias concretas:

- Imponer una lengua es imponer una forma de nombrar el mundo. La
- Borrar una lengua es borrar una forma de nombrar el mundo. La muerte de una
- Traducir es una operación política. Toda traducción es una negociación entre

18.2.3. La IA como operación geopolítica

La inteligencia artificial no solo describe el mundo, lo constituye con sus categorías. Cuando un modelo clasifica a una persona como "race: hispanic" o "family: nuclear", no la describe, la inscribe dentro de una ontología anglosajona. La clasificación es una forma de conquista.

Esto ya lo dice el libro (cap. 4, cap. 15), pero no lo nombra como geopolítica. Lo llama "colonialidad del dato", "sesgo anglosajón", "aplanamiento epistémico". Son conceptos correctos, pero epistémicos, no geopolíticos. La epistemología geopolítica los reformula al afirmar que "la IA no solo impone categorías", sino "la IA conquista el lenguaje con el que se nombra el mundo".

Quien controla el lenguaje de la IA controla la forma en que el mundo es conocido, y por lo tanto conquistado. La IA no es solo un instrumento técnico, es un instrumento geopolítico de primera magnitud.

18.3. La conquista del lenguaje en la IA actual

18.3.1. El borrado de las otras formas de nombrar

La IA actual borra formas de nombrar el mundo. Cada borrado es una pérdida de un mundo.

- Caso 1: el "*malu*" indonesio. El término indonesio *malu* designa un complejo
- Caso 2: la "etnia" latinoamericana. La categoría "etnia" en América Latina está
- Caso 3: la "familia extensa". La categoría "familia" en la IA actual presupone

18.3.2. La imposición de las formas de nombrar

La IA actual impone las formas de nombrar el mundo. Cada imposición es una conquista del mundo.

- Lengua. El inglés domina los corpus de entrenamiento. Las categorías, los marcos
- Categorías. Las categorías con las que la IA clasifica el mundo —género, raza,
- Valores. Los valores que la IA reproduce —individualismo, libertad negativa,

18.3.3. El monopolio del lenguaje como una forma de poder

La conjunción de borrado e imposición produce un monopolio del lenguaje. La IA actual no solo describe el mundo, lo constituye con sus categorías. Y esas categorías son anglosajonas. Quien controla la IA controla el lenguaje con el que el mundo es nombrado y por lo tanto conquistado.

El monopolio no requiere intención. Los ingenieros de OpenAI, Google o Anthropic no son "colonialistas" en el sentido histórico. El problema es estructural porque el ecosistema digital, las asimetrías de poder, las infraestructuras heredadas, las prácticas de recolección de datos, todo ello reproduce jerarquías sin que nadie las diseñe deliberadamente. Pero el efecto es el mismo, la conquista del lenguaje.

18.4. La resistencia geopolítica: la no alineación

18.4.1. La no alineación como resistencia geopolítica

Si la IA actual es una operación geopolítica, entonces la no alineación no es solo una posición epistémica, también es una resistencia geopolítica. No se trata de "corregir el sistema para que sea más justo", se trata de disputar el poder de nombrar el mundo.

Esto reformula el sentido de la no alineación:

- No es "rechazar la IA": no se trata de abandonar la técnica.
- No es "corregir la IA": no se trata de mejorar el sistema existente.

- Es disputar la hegemonía epistémica global. Quien conquista el lenguaje,

18.4.2. Recuperar el poder de nombrar el mundo

La no alineación como resistencia geopolítica implica recuperar el poder de nombrar el mundo. Esto significa:

- Nombrar con categorías propias, no importadas.
- Traducir entre tradiciones sin reducir unas a otras.
- Preservar las formas de nombrar que la IA actual borra.
- Disputar las formas de nombrar que la IA actual impone.

18.4.3. América Latina como sujeto geopolítico

Esto da un sentido nuevo al capítulo 15. "América Latina se ve a sí misma" no es solo una operación epistémica: es una operación geopolítica. La región recupera el poder de nombrarse. No acepta las categorías impuestas. Conquista su propio lenguaje.

La región no solo se ve, se nombra. Nombrarse es poseerse. La auto-visión no es solo un acto cognitivo, es un acto de soberanía geopolítica.

18.4.4. Estados Unidos como potencia que pierde el monopolio

Esto da un sentido nuevo al capítulo 15. "Estados Unidos entiende mejor la región" no es solo una operación epistémica, es una operación geopolítica. Estados Unidos pierde el monopolio del lenguaje, no puede imponer sus categorías, debe negociar con las categorías locales. La pérdida del monopolio epistémico es una pérdida de poder geopolítico.

Sin embargo, como argumenta el capítulo 15, la pérdida del monopolio también beneficia a Estados Unidos. Una IA que entiende la región funciona mejor para sus intereses, reduce los costos de la hegemonía y mengua los daños colaterales. La pérdida del monopolio no es una derrota, es una reconfiguración del poder.

18.5. La descolonización del lenguaje

18.5.1. Tradiciones

La tesis de la epistemología geopolítica conecta con varias tradiciones:

- Foucault: saber y poder son inseparables. El conocimiento no es neutral, produce
- Bourdieu: el lenguaje es un mercado de intercambios simbólicos. Quien controla
- Quijano, Mignolo: la colonialidad del saber es constitutiva de la modernidad.
- Spivak: el subalterno no puede hablar. No es que no tenga voz, es que su voz no
- Fanon: la descolonización es también descolonización del lenguaje. Recuperar la
- Wa Thiong'o: "Descolonizar la mente" es descolonizar el lenguaje. La lengua del

18.5.2. La descolonización como descolonización del lenguaje

La descolonización del lenguaje no es una metáfora, es una operación concreta. Implica:

- Recuperar las lenguas y las formas de nombrar que la colonización borró.
- Preservar las formas de nombrar que la IA actual borra.
- Disputar el monopolio del lenguaje que la IA actual impone.
- Construir un espacio donde múltiples formas de nombrar el mundo coexistan.

18.5.3. La diversidad epistémica como condición de un mundo no conquistado

La diversidad epistémica no es "tolerancia", es condición de un mundo no conquistado. Si una sola forma de nombrar el mundo se impone, el mundo se vuelve propiedad de quien lo nombra. Si múltiples formas coexisten, el

mundo permanece abierto. La diversidad epistémica no es un lujo, es una condición de libertad.

18.6. El bien público global como descolonización

18.6.1. La no alineación como bien público

El capítulo 16 mostró que la no alineación es un bien público global, no rival, no exclusiva, beneficiosa para todos. Este capítulo reformula esa tesis en clave geopolítica, la no alineación es descolonización del lenguaje.

Todos ganan si nadie monopoliza el poder de nombrar el mundo. América Latina recupera su lenguaje. Estados Unidos entiende la región. Europa, China, África, Asia pueden operar con múltiples formas de nombrar. El mundo permanece abierto.

18.6.2. La diversidad epistémica como condición de robustez

La diversidad epistémica no es solo justa, es robusta. Un mundo con múltiples formas de nombrar es más resiliente que un mundo con una sola. La monocultura epistémica no solo es injusta, es frágil.

Esto conecta con el capítulo 16, la no alineación es un bien público porque todos ganan, no solo los que sufren la imposición.

18.6.3. La gobernanza global como disputa por el lenguaje

La gobernanza global de la IA no es solo una cuestión técnica o regulatoria: es una disputa por el lenguaje. Quién decide qué categorías se usan, qué lenguas se priorizan, qué valores se imponen. La gobernanza es geopolítica.

Una gobernanza de la no alineación requeriría:

- Participación global en la producción de normas.
- Pluralismo epistémico: múltiples formas de nombrar.
- Soberanía epistémica: las comunidades controlan sus categorías.
- Reconocimiento de
- Justicia epistémica: las asimetrías de poder epistémico corregidas.

18.7. Estudio de caso: la Sedena y el CJNG en Michoacán

18.7.1. Relato del caso

En 2021 el Cártel de Jalisco Nueva Generación (CJNG) lanzó una ofensiva en Michoacán con armamento pesado, vehículos blindados y drones equipados con explosivos. Miles de personas fueron desplazadas de sus comunidades. La ofensiva no era un episodio más de la violencia crónica del estado: era una toma de territorio a gran escala, con capacidad militar comparable a la de una fuerza irregular profesionalizada.

La Secretaría de la Defensa Nacional (Sedena) respondió. Pero la respuesta no fue la que el lenguaje oficial sugiere. Según el informe *El laberinto de los generales: crimen y militares en México*, publicado por el International Crisis Group en mayo de 2024, la Sedena formó un frente de facto con grupos criminales locales antagónicos al CJNG. Un líder criminal entrevistado por Crisis Group lo describió con precisión: "miembros de otros dos grupos ilegales, un funcionario federal de alto rango, funcionarios locales, soldados y delincuentes locales formaron un frente conjunto de facto contra el Cártel de Jalisco, coordinando sus operaciones e incluso luchando unos junto a otros".

La coordinación no era un accidente táctico, fue una estrategia sostenida. Las Fuerzas Especiales del Ejército dudaron al principio en tomar medidas contundentes contra el CJNG, por miedo a ser acusadas de crímenes contra la humanidad. Pero después de que un grupo criminal local prometiera deshacerse de los cadáveres en fosas clandestinas, los soldados procedieron.

El resultado fue contundente. Cerca de 400 combatientes del CJNG fueron asesinados en Taixtán y El Terrero. Sus cuerpos fueron recogidos por combatientes de los grupos locales y arrojados en fosas clandestinas. Estas muertes no fueron reportadas en las estadísticas oficiales. No aparecen en los informes de la Sedena, ni en los de la

Fiscalía, ni en los de la Guardia Nacional. Son, para el registro oficial, muertes inexistentes.

Pero el CJNG no fue derrotado. Se reagrupó, reanudó sus ataques, y respondió con la misma lógica que sus rivales: "Relleno humano siempre hay", dijo un líder criminal a Crisis Group. La guerra continúa. La contradicción también.

Lo que este caso revela no es una "falla" del Estado mexicano, ni una "corrupción" de algunos funcionarios. Revela algo más profundo, una contradicción constitutiva de la geopolítica real del territorio. La Sedena es simultáneamente fuerza de orden y actor del desorden, combatiente del crimen organizado y aliada de grupos criminales, soberana y subordinada a las condiciones del terreno. No se trata de una institución "corrupta" ni de una institución "heroica". Se trata de una institución que habita una contradicción que el lenguaje oficial no puede nombrar.

Y es esa contradicción la que revela la diferencia entre una IA con lógica formal y una IA con lógica dialéctica.

18.7.2. Cuadro comparativo

Dimensión IA con lógica formal IA con lógica dialéctica Unidad de análisis Entidades discretas: Sedena, CJNG, grupos locales Totalidad: el territorio como sistema de contradicciones Objeto Clasificar actores: "fuerza de orden" vs. "criminal" Comprender la relación: orden y desorden como momentos Contradicción Error a resolver: o la Sedena es fuerza de orden, o es cómplice Motor a sostener: la Sedena es ambas cosas simultáneamente Tiempo Instantáneo: estado actual del conflicto Histórico: proceso de formación del territorio Sujeto No hay sujeto: el sistema calcula Sujeto colectivo: comunidades, Estado, cárteles en relación Éxito Precisión predictiva: ¿quién ganará? Comprensión: ¿qué contradicciones están activas? Lenguaje Reproduce el lenguaje oficial: "combate al crimen" Desmonta el lenguaje oficial: "contradicción sostenida" Efecto geopolítico Refuerza la ontología dominante Abre el espacio de la negociación real Ejemplo de output "La Sedena combate al CJNG. Nivel de éxito: X%." "La Sedena combate al CJNG y se alía con sus rivales. Ambas determinaciones coexisten." Efecto sobre los 400 muertos Invisibles: no están en los datos oficiales Visibles: emergen como síntoma de la contradicción Efecto sobre las fosas clandestinas Invisibles: no son "datos de seguridad" Visibles: emergen como condición de la operación Pregunta que puede formular "¿Cómo derrotar al CJNG?" "¿Qué totalidad produce esta guerra?" Respuesta que puede dar Una estrategia militar Un mapa de contradicciones

18.7.3. Conclusiones del caso

- Primera conclusión: la lógica formal no puede procesar el caso sin distorsionarlo.
- Una IA con lógica formal, ante este caso, tiene dos opciones y ambas son falsas:
- Opción A: clasificar a la Sedena como "fuerza de orden". Entonces la alianza con
- Opción B: clasificar a la Sedena como "cómplice del crimen". Entonces los 400

al resolverla, destruye el fenómeno.

- Segunda conclusión: la lógica dialéctica sostiene la contradicción y por eso

es simultáneamente fuerza de orden y actor de desorden, combatiente y cómplice, soberana y subordinada. No busca el "equilibrio" entre ambas determinaciones. Busca la totalidad en la que ambas se asumen como momentos: el Estado mexicano en su relación histórica con el crimen organizado, la fragmentación de la soberanía en el territorio, la lógica del "enemigo de mi enemigo", la producción de desapariciones como efecto colateral de la política de seguridad.

- Tercera conclusión: el caso revela la conquista del lenguaje.

ontología. Esa ontología oculta la contradicción. Una IA con lógica formal reproduce el lenguaje oficial. Una IA con lógica dialéctica lo desmonta: no hay "combate", hay contradicción sostenida. No hay "paz", hay desaparición como política tácita.

Y al desmontar el lenguaje, recupera el fenómeno. Los 400 muertos emergen. Las fosas clandestinas emergen. La alianza de facto emerge. Lo que el lenguaje oficial ocultaba, la dialéctica lo hace visible.

- Cuarta conclusión: el caso es útil para todos los actores.

caso epistemológico. La lógica formal distorsiona el fenómeno para todos los actores. La lógica dialéctica lo comprende para todos. Una IA dialéctica no favorece a ningún bando: favorece la comprensión del territorio como totalidad de contradicciones. Y esa comprensión es útil para el Estado, para las comunidades, para los analistas, para los negociadores, para cualquiera que quiera entender qué está pasando en Michoacán en lugar de puntuarlo desde una ontología que lo distorsiona.

- Quinta conclusión: la epistemología geopolítica es la condición de una política

realista sin una epistemología que pueda sostener las contradicciones que el territorio produce. La lógica formal no puede. La lógica dialéctica sí. Por eso la epistemología geopolítica no es un lujo filosófico: es la condición de una política que no sea ciega.

Conviene precisar el alcance de la distinción entre "IA con lógica formal" e "IA con lógica dialéctica". La segunda no es una IA dialéctica en sentido pleno, porque como se mostró en el capítulo 3, ningún artefacto participa en el movimiento dialéctico de lo real dado que los artefactos operan sobre representaciones del movimiento. Lo que aquí se llama "IA con lógica dialéctica" es una IA que representa el movimiento con mayor riqueza: incorpora contradicción, proceso, totalidad, historicidad, sin pretender estar dentro del proceso. No es dialéctica: es una representación más fiel de la dialéctica. Y eso ya es mucho.

18.8. La etiqueta y el análisis

18.8.1. La tesis

La lógica formal etiqueta mientras la lógica dialéctica analiza. Etiquetar es asignar una categoría a una entidad. La operación es binaria: o la entidad pertenece a la categoría, o no. No hay proceso, no hay historia, no hay relación. La etiqueta fija.

Analizar es comprender la entidad en su relación con una totalidad. La operación es procesual: la entidad se constituye en relaciones, en contradicciones, en procesos históricos. El análisis despliega.

La diferencia no es solo metodológica. Es epistemológica. La etiqueta oculta la intencionalidad del sistema que la produce, el análisis la desnuda.

18.8.2. Cuadro comparativo: el caso de Michoacán

Frase IA con lógica formal (etiqueta) IA con lógica dialéctica (análisis) "400 combatientes del CJNG" Etiqueta: "combatiente" + "CJNG" + "400". El sistema clasifica y cuenta. La etiqueta oculta: que la categoría "combatiente" es inestable; que "CJNG" es una entidad en disputa; que "400" es un número sin registro oficial. Análisis: "400 personas murieron en enfrentamientos entre grupos armados, en un territorio donde la distinción combatiente/civil se ha vuelto inestable. Sus cuerpos fueron arrojados a fosas clandestinas. Sus muertes no fueron registradas. La operación fue posible gracias a una coordinación de facto entre la Sedena y grupos rivales." "La Sedena combate al CJNG" Etiqueta: "Sedena" + "combate" + "CJNG". El sistema clasifica actores y acciones. La etiqueta oculta: que la Sedena coordinó con grupos rivales; que la coordinación incluyó "luchar unos junto a otros"; que el lenguaje oficial distorsiona el fenómeno. Análisis: "La Sedena combate al CJNG y se coordina con grupos rivales. Ambas determinaciones coexisten. No se trata de una contradicción a resolver, sino de una contradicción a sostener. La Sedena es simultáneamente fuerza de orden y actor de desorden." "La operación fue un éxito" Etiqueta: "éxito" + "operación". El sistema evalúa. La etiqueta oculta: que 400 muertes no fueron registradas; que el CJNG se reagrupó; que la guerra continúa; que el "éxito" se mide con métricas oficiales que no capturan el fenómeno. Análisis: "La operación produjo 400 muertes no registradas, fortaleció a grupos criminales locales, y no impidió la reagrupación del CJNG. El 'éxito' es una etiqueta que oculta la totalidad del proceso. La pregunta no es '¿fue exitosa?', sino '¿qué totalidad produjo?'" "Los grupos locales son aliados" Etiqueta: "aliados". El sistema clasifica la relación. La etiqueta oculta: que la alianza es de facto, no formal; que la alianza es táctica, no estratégica; que la alianza produce efectos no previstos (fosas clandestinas, muertes no registradas). Análisis: "Los grupos locales son aliados tácticos de la Sedena en la operación contra el CJNG. Pero la alianza no es formal, no es estratégica, y produce efectos que el lenguaje oficial no nombra: fosas clandestinas, muertes no registradas,

fortalecimiento de grupos criminales locales." "La violencia en Michoacán es un problema de seguridad" Etiqueta: "violencia" + "seguridad". El sistema clasifica el fenómeno. La etiqueta oculta: que la violencia es producida por la totalidad del territorio (Estado, cárteles, comunidades, economías ilegales); que la "seguridad" es un concepto que reduce el fenómeno a su dimensión militar. Análisis: "La violencia en Michoacán no es un 'problema de seguridad'. Es la manifestación de una totalidad: el Estado mexicano en su relación histórica con el crimen organizado, la fragmentación de la soberanía en el territorio, la lógica del 'enemigo de mi enemigo', la producción de desapariciones como efecto colateral de la política de seguridad." "Relleno humano siempre hay" Etiqueta: "recurso humano" + "disponible". El sistema clasifica. La etiqueta oculta: que la frase es de un líder criminal; que revela la lógica de la guerra; que las personas son tratadas como material intercambiable; que el Estado no puede impedirlo. Análisis: "'Relleno humano siempre hay' es la frase de un líder criminal. Revela la lógica de la guerra: las personas son material intercambiable. Y revela la contradicción del Estado: combate al crimen mientras negocia con criminales que tratan a las personas como relleno." "Las muertes no fueron reportadas" Etiqueta: "no reportadas". El sistema clasifica. La etiqueta oculta: que la ausencia de reporte es una decisión, no un accidente; que la decisión produce efectos: invisibilidad, impunidad, continuación de la guerra. Análisis: "Las muertes no reportadas no son un accidente administrativo. Son una decisión que produce efectos: invisibilidad de los cuerpos, impunidad de los responsables, continuación de la guerra. La 'ausencia de registro' es una política tácita." "La guerra continúa" Etiqueta: "guerra" + "continúa". El sistema clasifica el estado del conflicto. La etiqueta oculta: que la "guerra" no es entre dos bandos claros; que el Estado es parte del conflicto; que las comunidades son las principales víctimas; que la continuación es un efecto estructural, no un accidente. Análisis: "La guerra continúa. Pero no es una guerra entre dos bandos. Es una totalidad de contradicciones: el Estado combate y negocia, los cárteles se alían y se enfrentan, las comunidades resisten y son desplazadas. La continuación no es un accidente: es un efecto estructural de la totalidad."

18.8.3. Cierre: la etiqueta desnuda la intencionalidad

El cuadro revela algo que la lógica formal oculta: su intencionalidad.

La lógica formal no es neutral. Cada etiqueta que produce presupone una ontología. Y esa ontología no es la ontología del fenómeno, es la ontología del sistema que etiqueta. Cuando la IA con lógica formal dice "400 combatientes del CJNG" no está describiendo un fenómeno sino imponiendo una ontología, la ontología del lenguaje oficial, la ontología del Estado que clasifica, la ontología que oculta las fosas clandestinas, las muertes no registradas, la coordinación de facto con criminales.

La lógica dialéctica desnuda esa intencionalidad, no porque sea "neutral" —ninguna lógica lo es—, sino porque mantiene la contradicción en lugar de resolverla u ocultarla. Y al mantenerla, hace visible lo que la etiqueta oculta.

Esto es epistemología geopolítica. No se trata solo de "describir mejor" el fenómeno. Se trata de desnudar la intencionalidad del sistema que lo describe. Y esa desnudez es la condición de una política que no sea ciega.

18.9. Consideraciones operativas: qué debería haber hecho la Guardia Nacional en Michoacán al utilizar IA con lógica dialéctica

18.9.1. La lógica operativa de la IA dialéctica

Una IA con lógica formal, ante el crimen organizado en Michoacán, produce un plan de operaciones: objetivos, blancos, métricas de éxito. La operación se evalúa por bajas enemigas, territorio recuperado, líderes capturados. Los efectos no previstos —fosas clandestinas, desplazamientos, fortalecimiento de grupos rivales— quedan fuera del análisis porque no son "datos de seguridad".

Una IA con lógica dialéctica no produce un plan de operaciones. Produce un mapa de la totalidad y deriva de él un conjunto de operaciones concretas orientadas a transformar las condiciones que producen la guerra, no a combatir a los actores que la libran.

La diferencia fundamental es esta: la estrategia formal combate al CJNG; la estrategia dialéctica combate las condiciones del CJNG. La primera produce reemplazos, desplazamientos, fosas clandestinas, más violencia. La segunda produce flujos asegurados, productores protegidos, economía desvinculada de la extorsión, CJNG sin

oxígeno.

Y la totalidad de Michoacán es, fundamentalmente, una economía agrícola capturada por el crimen. Los cárteles no solo trafican drogas. Gobiernan la economía legal. Cobran dos pesos por cada kilo de limón cortado y dos pesos más por kilo empacado. Controlan la venta de pollo, carne, tortilla, gasolina, refresco. La extorsión a los productores de limón genera aproximadamente 3,600 millones de pesos al año. El limón y el aguacate son las principales exportaciones de Michoacán y fluyen hacia Lázaro Cárdenas y Manzanillo para llegar a Japón, China y Singapur.

La guerra en Michoacán no es solo una guerra por territorio criminal, es una guerra por el control de la economía agrícola. Y las operaciones militares, tal como están diseñadas, no tocan esa economía. Combaten al CJNG, pero dejan intacta la estructura de extorsión que financia a todos los grupos.

18.9.2. Las siete operaciones

- Primera: asegurar las rutas de salida de la producción

Valle de Apatzingán, Tierra Caliente— con los puertos de Lázaro Cárdenas y Manzanillo. Patrullaje permanente, escolta de convoyes, control de carreteras.

Lógica dialéctica: si el crimen organizado captura el flujo de la producción, controla toda la cadena. Si el Estado asegura el flujo, rompe la captura sin combatir directamente. No se trata de "vencer" a los grupos, sino de cortar la fuente de su poder económico: el control de la circulación.

Contraste con la lógica formal: la IA formal propone atacar las posiciones del CJNG. La IA dialéctica propone asegurar el flujo que el CJNG necesita para existir. No combate al actor: transforma la totalidad en la que el actor opera.

- Segunda: proteger a los productores en el punto de producción

"combatir" a los grupos, sino para garantizar que los productores puedan trabajar sin pagar extorsión. Patrullaje en las parcelas, protección de los trabajadores agrícolas, control de los accesos a las zonas de producción.

Lógica dialéctica: la extorsión funciona porque los productores están solos. Si el Estado está presente en el punto de extracción, la extorsión se vuelve más costosa para el extorsionador. No se elimina la extorsión: se transforma la relación de fuerzas en el punto crítico.

Contraste con la lógica formal: la IA formal mide el éxito por "bajas enemigas". La IA dialéctica mide el éxito por reducción de la extorsión efectiva.

- Tercera: asegurar los puntos de empacado y procesamiento

no solo ocurre en el campo: ocurre en el empacado. Los cárteles cobran por kilo empacado. Si el Estado asegura las empacadoras, corta otro eslabón de la cadena.

Lógica dialéctica: la cadena de valor del limón y el aguacate tiene múltiples puntos de captura. El crimen controla la cadena en su totalidad. La IA dialéctica identifica todos los puntos y los asegura simultáneamente, no uno por uno.

- Cuarta: no hacer acuerdos con grupos criminales

irregular. Sin excepciones. Sin "frentes de facto". Sin "luchar hombro con hombro".

Lógica dialéctica: la coordinación con criminales contradice el objetivo declarado. La IA dialéctica sostiene la contradicción y obliga al mando a elegir con conciencia. No se puede "combatir al crimen" y "coordinar con criminales" al mismo tiempo. La elección es política, no técnica. Pero debe ser consciente.

Contraste con la lógica formal: la IA formal oculta la contradicción. Clasifica a la GN como "fuerza de orden" y a los grupos locales como "aliados tácticos". La IA dialéctica desnuda la contradicción.

- Quinta: proteger a los liderazgos locales que resisten

humanos, periodistas locales. Los casos documentados muestran que los liderazgos que resisten son asesinados sistemáticamente: el alcalde Carlos Manzo de Uruapan fue asesinado en 2025 a pesar de tener 22 personas de seguridad; el líder limonero Bernardo Bravo fue asesinado; Hipólito Mora, fundador de las autodefensas, fue asesinado en 2023.

Lógica dialéctica: la resistencia local es la única fuerza que puede sostener la transformación a largo plazo. Si el Estado no protege a quienes resisten, destruye la posibilidad de la transformación. La protección no es "humanitaria": es estratégica.

- Sexta: vincular la operación con la política agrícola y social

precios de garantía, infraestructura de riego, asistencia técnica, compradores alternativos que no estén capturados por el crimen.

Lógica dialéctica: el crimen ofrece "protección" porque el Estado no está. Si el Estado ofrece protección real —seguridad física, pero también estabilidad económica— la relación de fuerzas se transforma. El crimen ya no es necesario para que la economía funcione.

Contraste con la lógica formal: la IA formal ve la seguridad como un problema militar. La IA dialéctica ve la seguridad como un problema total: militar, económico, social, político.

- Séptima: documentar y registrar lo que el sistema formal no registra

muerdes no reportadas. Lo que no se registra no existe para el sistema formal. La IA dialéctica hace visible lo invisible.

Lógica dialéctica: los 400 muertos no registrados no son un accidente administrativo. Son una política tácita. Si el Estado no los registra, niega que existieron. La IA dialéctica los registra, y al registrarlos, obliga al Estado a asumirlos.

18.9.3. Cómo golpea esto al CJNG

La pregunta que haría cualquier operador formado en la lógica formal es: ¿cómo golpea esto al CJNG? La respuesta es sencilla, el CJNG se queda sin oxígeno porque el CJNG no produce, extrae

El CJNG no produce limón ni aguacate. No siembra, no cosecha, no empaca, no transporta. Extrae renta de quienes producen. Su modelo de negocio es el peaje: cobra dos pesos por kilo cortado, dos pesos por kilo empacado, controla la venta de pollo, carne, tortilla, gasolina, refresco. Su poder no viene de producir: viene de capturar el flujo de quienes producen.

Esto significa que el CJNG es estructuralmente dependiente de la economía local. Si la economía local no fluye por los canales que él controla, el CJNG no tiene nada que extraer. Su poder es parasitario: vive del huésped. Si el huésped se protege, el parásito se debilita.

El CJNG no puede monopolizar: Michoacán es un archipiélago de poderes

El CJNG tiene la pretensión de aplastar a todos los grupos locales y monopolizar el negocio. Pero no puede. Michoacán no es un territorio vacío que se pueda ocupar: es un archipiélago de poderes locales —Cárteles Unidos, Los Viagras, Los Ardillos, Los Tlacos, decenas de grupos más pequeños, autodefensas, comunidades armadas— con bases sociales, económicas y territoriales propias.

El CJNG es uno entre tantos. Su pretensión de monopolio es estructuralmente imposible. Y cuanto más intenta aplastar a los locales, más los une contra él. La ofensiva de 2021 produjo precisamente eso: una coalición de facto de grupos locales, con la Sedena, contra el CJNG.

El CJNG está atrapado en una contradicción: para crecer, tiene que aplastar a los locales. Pero aplastar a los locales produce coaliciones que lo contienen. Y mientras tanto, la economía que quiere controlar se desmorona por la violencia.

La estrategia dialéctica golpea al CJNG donde más le duele: su fuente de oxígeno

Si el Estado asegura los flujos —rutas, empacadoras, puntos de producción— el CJNG pierde el acceso a la renta. No puede cobrar peaje si el flujo no pasa por sus manos. No puede controlar la economía si la economía no está capturada.

La estrategia dialéctica no combate al CJNG directamente. Combate las condiciones que le permiten existir. No ataca al actor: desactiva el sistema del que el actor vive.

Y aquí está la paradoja: la estrategia formal, que combate al CJNG directamente, lo fortalece. Cada operación militar que mata a 400 combatientes produce reemplazos ("relleno humano siempre hay"), produce desplazamientos, produce fosas clandestinas, produce más violencia. La guerra alimenta al CJNG. La estrategia dialéctica lo ahoga sin combatirlo.

Una vez que el CJNG se queda sin oxígeno, sin acceso a la renta de la economía local, le queda:

- Tráfico de drogas: el negocio original, pero mucho menos rentable que el control
- Extorsión residual: si el Estado está presente, la extorsión se vuelve costosa.
- Violencia pura: sin base económica, la violencia se vuelve insostenible. No se

de disputar el territorio a los grupos locales, al Estado, a las comunidades.

Además, el CJNG se mueve en una contradicción local, buscando monopolizar lo que no se puede monopolizar. Es una contradicción insostenible:

- Quiere monopolizar el control territorial de Michoacán.
- No puede monopolizar porque Michoacán es un archipiélago de poderes locales.
- Cada intento de monopolizar produce coaliciones que lo contienen.
- Cada operación militar contra él lo fortalece (reemplazos, radicalización).
- Cada operación dialéctica contra sus condiciones lo debilita (sin renta, sin

sostenerla, obliga al CJNG a habitar una contradicción que no puede habitar. El CJNG quiere monopolizar y no puede; quiere combatir y se debilita; quiere controlar y se queda sin oxígeno.

18.9.4. No vencer, mejorar

Una fuerza militar que armara su estrategia con IA dialéctica no sería "más humana" ni "más ética". Sería más realista. Vería cosas que la IA formal no ve. Sostendría contradicciones que la IA formal resuelve prematuramente.

Y al ver más y sostener más, actuaría con mayor conciencia de las consecuencias de sus actos. No se trata de "vencer" al CJNG. Se trata de transformar la totalidad que produce la guerra. No se trata de "combatir al actor". Se trata de desactivar el sistema del que el actor vive.

"Vencer" pertenece a la lógica formal. Presupone una dicotomía, un estado final, una métrica, un cierre. "Mejorar" pertenece a la lógica dialéctica: transformar una contradicción en un nivel superior. No se trata de eliminar al adversario. Se trata de reconfigurar la totalidad de manera que la contradicción se asuma, no se resuelva.

La pregunta no es "¿cómo derrotar al CJNG?". Es "¿cómo transformar la contradicción que produce la guerra en Michoacán?". Y esa transformación no se logra solo con una operación militar. Se logra con una reconfiguración de la totalidad: el Estado mexicano en su relación histórica con el crimen organizado, la fragmentación de la soberanía, la lógica del "enemigo de mi enemigo", la producción de desapariciones como política tácita.

Esto no garantiza que la operación actúe "bien". La dialéctica no es una ética, es una epistemología. Pero una epistemología que ve la totalidad es condición de una estrategia que no sea ciega. Y una estrategia que no sea ciega es condición de una política que no produzca, sin saberlo, las fosas clandestinas que dice querer evitar.

Y una precisión final sobre el lenguaje. Llamar "inteligencia" a la IA no es una descripción neutral, es una operación epistémica y política. Al llamarla inteligente, se la inscribe en una tradición que identifica la inteligencia

con la resolución de problemas y se excluye otra tradición que la identifica con la participación en el movimiento de lo real. La primera tradición es la de la matriz occidental moderna, la segunda es la de Hegel y la dialéctica. La disputa sobre si la IA "es inteligente" es, en el fondo, una disputa sobre qué se entiende como inteligencia. Y esa disputa no se resuelve con más datos o mejores métricas, se resuelve políticamente.

Bibliografía del capítulo 18

Bourdieu, P. (1982). *Ce que parler veut dire: L'économie des échanges linguistiques*. Fayard.

Fanon, F. (1952). *Peau noire, masques blancs*. Seuil.

Foucault, M. (1969). *L'archéologie du savoir*. Gallimard.

Foucault, M. (1975). *Surveiller et punir: Naissance de la prison*. Gallimard.

International Crisis Group. (2024). *El laberinto de los generales: crimen y militares en México* (Informe sobre América Latina N°106).

Mignolo, W. (2007). *La idea de América Latina*. Gedisa.

Quijano, A. (2000). Colonialidad del poder, eurocentrismo y América Latina. En E. Lander (ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales* (pp. 201–246). CLACSO.

Spivak, G. C. (1988). Can the Subaltern Speak? En C. Nelson & L. Grossberg (eds.), *Marxism and the Interpretation of Culture*. University of Illinois Press.

Wa Thiong'o, N. (1986). *Decolonising the Mind: The Politics of Language in African Literature*. James Currey.

El Universal. (2025, 23 de octubre). Crimen se embolsa 3 mil 600 mdp al año por extorsionar a limoneros.

Vanguardia MX. (2025, 23 de octubre). Grupos criminales obtienen casi 4 mmp de productores de limones en Apatzingán.

Conclusiones: la no alineación como proyecto civilizatorio

19.1. Qué es una epistemología civilizatoria

19.1.1. Una definición

Llamamos epistemología civilizatoria al modo de conocer propio de una civilización en un periodo histórico determinado, con las variaciones de tiempo y espacio que le son propias. No se trata de una teoría del conocimiento, ni de una disciplina filosófica, ni de un método particular. Se trata de la forma concreta, histórica y situada en que una civilización produce, organiza y transmite lo que considera conocimiento. Esa forma no es un conjunto de reglas explícitas que alguien haya decidido: es un régimen que se hereda, se habita, se reproduce y, a veces, se transforma. Cada civilización, en cada momento de su desarrollo, tiene su epistemología civilizatoria; y ninguna es universal, porque ninguna agota las posibilidades del conocer humano.

Una epistemología civilizatoria comprende, en primer lugar, un conjunto de instrumentos que la civilización admite como legítimos para conocer. No todos los instrumentos disponibles en una época son admitidos: algunos se incorporan al canon, otros quedan fuera, otros se marginan. El telescopio fue admitido en la modernidad europea; el oráculo, no. El censo fue central en el imperio inca; la escritura alfabética, no. La estadística se incorporó a las ciencias sociales del siglo XX; la intuición etnográfica, con reservas. La selección no es neutral: refleja los criterios, las jerarquías y las exclusiones de la civilización que la practica. Lo mismo ocurre con los conceptos: no todos los conceptos son admitidos para pensar. La categoría "individuo" fue central en la modernidad europea; la categoría "comunidad", subordinada. La categoría "progreso" fue central en el siglo XIX; la categoría "ciclo", marginada. Y lo mismo ocurre con las cosmovisiones: no todas son admitidas como marco último de sentido. La cosmovisión secular se impuso en la modernidad europea; la cosmovisión religiosa quedó relegada a lo privado. La cosmovisión indígena fue tratada como folclore en la colonia, no como conocimiento. Instrumentos, conceptos y cosmovisiones forman, en su articulación, el tejido de una epistemología civilizatoria.

Pero el rasgo decisivo de una epistemología civilizatoria no es lo que admite, sino lo que constituye. Una epistemología civilizatoria no se limita a seleccionar instrumentos, conceptos y cosmovisiones: constituye al objeto mismo del conocimiento. Define qué se ve y qué no se ve en el mundo por conocer. No hay un "mundo" preexistente que la epistemología se limite a representar o filtrar: el mundo se constituye en el acto de conocer. La epistemología civilizatoria define qué entidades hay, qué relaciones son reales, qué fenómenos son visibles, qué preguntas tienen sentido. Lo que no se ve no es que esté oculto: es que no existe como objeto de conocimiento dentro de esa epistemología. Solo cuando otra epistemología lo constituye, aparece. Esto significa que distintas epistemologías civilizatorias no son perspectivas sobre un mismo mundo: son mundos distintos, constituidos en el conocer. No se trata de relativismo: hay realidad, pero la realidad se constituye en el conocer. Distintas epistemologías civilizatorias constituyen distintos mundos, y esos mundos coexisten, se conflictúan, se traducen con dificultad.

Esta constitución del objeto tiene consecuencias que van más allá de lo epistémico en sentido estricto. Una epistemología civilizatoria configura las formas de vida, los sujetos y los mundos que una civilización hace habitables. Al definir qué cuenta como conocimiento, define también qué prácticas se sostienen, qué tradiciones se transmiten, qué sujetos son reconocidos como cognoscentes, qué entidades pueblan el mundo, qué relaciones son reales, qué posibilidades están abiertas. Lo que queda fuera de la epistemología civilizatoria no es solo "desconocido": es inhabitable como forma de vida, irreconocible como sujeto, inexistente como mundo. Por eso la epistemología civilizatoria no es una cuestión técnica ni disciplinar: es una cuestión civilizatoria en el sentido más amplio. Afecta a cómo vivimos, a quiénes somos, a qué mundos habitamos.

Finalmente, una epistemología civilizatoria es histórica y plural. Histórica, porque varía en el tiempo: no es la misma en la Grecia clásica, en la Europa medieval, en la modernidad europea, en la contemporaneidad anglosajona. Plural, porque varía en el espacio: no es la misma en China, en India, en el mundo andino, en Europa, en el mundo islámico. Y no es homogénea: cada civilización tiene variaciones internas, conflictos, tradiciones en

tensión. No hay una epistemología civilizatoria, sino varias en conflicto. La historia de la humanidad es la historia de ese conflicto: la que se impone y las que resisten, la que se naturaliza y las que se marginan. Cuando una epistemología civilizatoria se impone, reconfigura la civilización. Cuando resiste, sostiene otras formas de vida, otros sujetos, otros mundos.

Una epistemología civilizatoria es, por último, un horizonte, no un estado. No se realiza de una vez, sino que se cultiva a lo largo del tiempo. Es una dirección, no un destino. No hay un momento en que la epistemología civilizatoria "esté completa": siempre está en proceso, siempre en disputa, siempre abierta. Por eso no hablamos de un proyecto —que implica un plan, un sujeto, un horizonte temporal acotado—, sino de una visión que se encarna en instrumentos, se cultiva en prácticas y se transmite en tradiciones. Una visión que no se posee, sino que se habita; que no se impone, sino que se sostiene; que no se clausura, sino que se abre.

Conviene precisar la relación entre "epistemología civilizatoria" y "matriz occidental moderna". "Epistemología civilizatoria" es el género, o sea el modo de conocer propio de una civilización en un periodo histórico determinado. "Matriz occidental moderna" es una especie, la configuración histórica del conocer que se ha impuesto globalmente a partir de la expansión europea, la revolución industrial y la globalización. La matriz occidental moderna no es "la" epistemología civilizatoria, sino una epistemología civilizatoria entre otras epistemologías, que se ha vuelto dominante. Hablar de "civilización epistémica dominante" es hablar de esa matriz en su forma contemporánea.

19.1.2. Cómo conoce la civilización epistémica dominante: rasgos de la matriz occidental moderna y su hegemonía contemporánea.

Llamamos civilización epistémica dominante a la configuración del conocer que se ha impuesto globalmente a partir de la expansión europea, la revolución industrial y la globalización, y que hoy reproduce y amplifica la IA. No es "la" civilización, ni "la" epistemología: es una entre otras, pero dominante. Se caracteriza como matriz occidental moderna: occidental, porque sedimenta tradiciones griegas, romanas y judeocristianas; moderna, porque se configura a partir del Renacimiento, la Reforma, la Ilustración y la revolución científica; y dominante, porque se ha impuesto —no sin resistencias— sobre otras epistemologías civilizatorias. Advertimos que se trata de una categoría gruesa, con dos mil años de historia, enormes variaciones internas, conflictos y transformaciones. No la usamos como esencia, sino como configuración histórica que permite caracterizar un modo de conocer. La noción de "matriz occidental moderna" que empleo aquí se apoya en los análisis de la colonialidad del poder (Quijano, 2000) y de la colonialidad del saber (Mignolo, 2007).

- El primer rasgo es la primacía de la razón discursiva. Occidente hereda de Grecia
- El segundo rasgo es la separación entre sujeto y objeto. Occidente hereda de
- El tercer rasgo es la temporalidad lineal y escatológica. Occidente hereda del
- El cuarto rasgo es la universalidad de la verdad. Occidente hereda de Grecia la
- El quinto rasgo es la voluntad de dominio. Occidente hereda del Génesis la
- El sexto rasgo es la institucionalización del saber. Occidente desarrolla, a lo largo

universalidad de la verdad, voluntad de dominio, institucionalización del saber— configuran la civilización epistémica dominante en su forma moderna y contemporánea. No son esencias: son sedimentos históricos, que se transforman, se conflictúan, se resisten. Dentro de la propia matriz occidental hay tradiciones que los cuestionan: la mística, la fenomenología, la hermenéutica, la teoría crítica, el romanticismo, el ecologismo. Pero esas tradiciones operan dentro de la matriz, y a menudo desde dentro la contestan. La matriz, sin embargo, permanece: es la atmósfera que respiramos quienes hemos sido formados en ella.

Esta civilización epistémica no es la civilización: es una entre otras. Pero se ha impuesto globalmente a través de la expansión colonial, la revolución industrial, la globalización. Hoy, la IA actual la reproduce y la amplifica: la lógica formal, el actor racional, el sesgo anglosajón son sedimentos de esta matriz. No son "universales": son occidentales-modernos. Y su imposición global no es un hecho natural: es un conflicto entre epistemologías civilizatorias.

19.1.3. Cómo reproduce y amplifica la IA los rasgos de esta civilización epistémica dominante

La inteligencia artificial no es un instrumento neutro que flota por encima de las civilizaciones. Es un producto histórico de la civilización epistémica dominante, y reproduce sus rasgos. No solo los reproduce: los amplifica, porque los materializa en infraestructura y los vuelve operativos a una escala inédita. Lo que en la filosofía moderna era un supuesto, en la IA se convierte en arquitectura. Lo que en la ciencia moderna era un método, en la IA se convierte en algoritmo. Lo que en la civilización occidental era una tendencia, en la IA se convierte en condición de posibilidad de la cognición legítima. En este apartado muestro cómo los seis rasgos de la civilización epistémica dominante —razón discursiva, separación sujeto-objeto, tiempo lineal, universalidad de la verdad, voluntad de dominio, institucionalización del saber— se reproducen y se amplifican en la IA actual.

- El primer rasgo, la razón discursiva, se reproduce en la IA a través de la lógica
- El segundo rasgo, la separación sujeto-objeto, se reproduce en la IA a través de
- El tercer rasgo, la temporalidad lineal, se reproduce en la IA a través de
- El cuarto rasgo, la universalidad de la verdad, se reproduce en la IA a través de
- El quinto rasgo, la voluntad de dominio, se reproduce en la IA a través de
- El sexto rasgo, la institucionalización del saber, se reproduce en la IA a través de

epistémico que configura todo lo que toca. La razón discursiva provee las herramientas; la separación sujeto-objeto define la relación con el mundo; la temporalidad lineal define el régimen de progreso; la universalidad de la verdad define el criterio de éxito; la voluntad de dominio define el *telos*; la institucionalización del saber define las condiciones de legitimidad. Juntos forman un sistema que filtra, traduce y configura todo lo que entra en su campo. Lo que no se ajusta a este sistema no es "otro conocimiento": es invisible para la IA. No se rechaza explícitamente: simplemente no se ve.

La consecuencia más profunda de esta reproducción y amplificación es que la IA naturaliza la civilización epistémica dominante. No la presenta como una configuración histórica entre otras: la presenta como la forma obvia de conocer. Los ingenieros no dicen "vamos a usar lógica formal": simplemente usan lógica formal. Los *datasets* no dicen "vamos a privilegiar el inglés": simplemente están en inglés. Las métricas no dicen "vamos a medir solo lo formalizable": simplemente miden lo formalizable. La civilización epistémica dominante deja de ser una civilización y se convierte en la civilización. Y esa naturalización es lo que hace que la crítica sea urgente: hay que desnaturalizar la matriz para poder reclamar otras formas de conocer.

La IA no es solo un producto de la civilización epistémica dominante: es su instrumento más eficaz de reproducción y amplificación. A través de la IA, la matriz occidental moderna se materializa en infraestructura global, se impone como condición de posibilidad de la cognición legítima, se naturaliza como forma universal de conocer. Por eso la no alineación no es una cuestión técnica: es una cuestión civilizatoria. No se trata de "mejorar" la IA: se trata de confrontar la civilización epistémica que encarna, y de cultivar otras formas de conocer que resistan a su hegemonía.

19.1.4. La no alineación como apuesta por otra epistemología civilizatoria

Si la IA actual no es un instrumento neutro, sino la materialización de una civilización epistémica dominante —la matriz occidental moderna en su configuración técnico-formal contemporánea—, entonces la no alineación no puede ser una mera operación técnica. No se trata de "corregir" el sistema para que se ajuste mejor a nuestros valores, ni de "mejorar" sus métricas, ni de "añadir" capas de diversidad. Se trata de algo más profundo: no alinearse con la epistemología civilizatoria que la IA encarna, y apostar por otra. La no alineación no es un "no" a la técnica. Es un "no" a una civilización epistémica y un "sí" a otra.

Decir que no nos alineamos con la civilización epistémica dominante no significa rechazar todo lo que esa civilización ha producido. La lógica formal, la ciencia experimental, el método crítico, la institucionalidad del saber: son logros que no queremos perder. No se trata de sustituir una matriz por otra, ni de "volver" a formas de conocer pre-modernas. Se trata de negar la hegemonía de esa matriz, no su validez. La lógica formal sigue siendo una herramienta poderosa; el problema es que se haya convertido en el único criterio de racionalidad. La ciencia

experimental sigue siendo un método riguroso; el problema es que se haya convertido en la única forma de conocer legítima. La institución del saber sigue siendo necesaria; el problema es que haya devenido en infraestructura global excluyente. La no alineación no niega los logros: rechaza su absolutización.

Ahora bien, ¿cuál es la otra epistemología civilizatoria por la que apostamos? No es una matriz alternativa cerrada, que sustituya a la dominante. No es "la" epistemología de "la" civilización no occidental, porque esa no existe como bloque. No es un regreso a formas de conocer pre-modernas, porque la historia no se deshace. No es un eclecticismo que mezcle todo con todo, porque eso sería otra forma de no tomar posición. Es algo más preciso: una epistemología civilizatoria plural, que reconozca la diversidad de formas de conocer y no reduzca unas a otras. No se trata de "elegir" entre la matriz occidental y otras matrices: se trata de abrir el espacio para que múltiples epistemologías civilizatorias coexistan, dialoguen, se traduzcan con dificultad, se conflictúen sin destruirse.

Esta epistemología civilizatoria plural se define por tres rasgos que corresponden, en negativo, a los tres rasgos que la IA actual reproduce y amplifica. Frente a la razón discursiva como único criterio, una razón plural: que reconozca la lógica formal y la dialéctica, la deducción y la analogía, el concepto y la imagen, la proposición y el cuerpo. Frente a la separación sujeto-objeto, una relacionalidad constitutiva: que reconozca que el sujeto está inmerso en el mundo, que el conocimiento es relación antes que representación, que no hay un punto neutro desde el cual contemplar. Frente a la universalidad de la verdad, una situacionalidad del conocer: que reconozca que toda verdad está situada, que hay múltiples verdades en conflicto, que la traducción entre tradiciones es precaria y necesaria. Frente a la voluntad de dominio, una voluntad de cuidado: que reconozca que conocer no es dominar, sino habitar, cultivar, cuidar. Frente a la temporalidad lineal, una multiplicidad temporal: que reconozca tiempos cíclicos, míticos, rituales, discontinuos, no solo el tiempo del progreso. Frente a la institucionalización excluyente, una institucionalidad plural: que reconozca instituciones diversas —comunitarias, tradicionales, académicas, técnicas— sin jerarquizarlas de antemano.

Esta otra epistemología civilizatoria no existe como sistema. Existe como constelación de prácticas, tradiciones, resistencias, proyectos. Existe en las comunidades indígenas que defienden su soberanía epistémica; en los movimientos sociales que luchan por la justicia cognitiva; en los académicos críticos que desarrollan epistemologías alternativas; en los ingenieros disidentes que construyen prototipos no alineados; en las tradiciones no occidentales que sostienen formas de conocer marginadas; en los estados del Sur global que reivindican soberanía tecnológica; en las instituciones internacionales que reconocen la dimensión epistémica de la IA. La apuesta por epistemologías alternativas se apoya en tradiciones como la epistemología del Sur (De Sousa Santos, 2010) y la interculturalidad crítica (Walsh, 2009). Y los prototipos que anticipan una IA del conocimiento humano incluyen los razonadores neuro-simbólicos con lógicas no clásicas (Ferguson et al., 2025), los sistemas de Curry (Abe, 2025) y los marcos de computación relacional (Dumas, 2026). No forman un bloque, no comparten una ideología y no tienen un plan común. Pero convergen en la apuesta por otra epistemología civilizatoria. La no alineación no se libra desde un centro, sino desde múltiples puntos.

La no alineación, en este sentido, no es un proyecto que alguien pueda "llevar a cabo". No tiene un sujeto único, un plan, un horizonte temporal acotado. Es una visión que se encarna en instrumentos, se cultiva en prácticas, se transmite en tradiciones. Como toda epistemología civilizatoria, es más un horizonte que un status establecido. No se realiza de una vez, sino a lo largo del tiempo. No hay un momento en que la no alineación "esté completa" sino que siempre está en proceso, en disputa y abierta. La pregunta pertinente no se refiere a cuándo estará disponible sino a qué hacemos algo, en nuestras prácticas concretas y en los instrumentos y alianzas concretas cotidianas, para avanzar en su constitución.

La no alineación, finalmente, no es una cuestión técnica. No se resuelve con mejores algoritmos, con más datos, con métricas más finas. Es una cuestión epistémica y civilizatoria: qué formas de conocer queremos sostener, qué mundos queremos habitar, qué sujetos queremos ser. La IA actual nos impone un modo de conocer y, con él, un modo de habitar el mundo. La no alineación es la apuesta por otro modo de conocer y, con él, por otro modo de habitar el mundo. No es un "no" a la IA: es un "sí" a otra IA, la que correspondería a otra epistemología civilizatoria. Una IA que no reduzca el pensamiento humano a lo calculable, que reconozca la diversidad epistémica, que se abra a formas de conocer que la historia ha excluido. Una IA que no sea la IA, sino una IA entre otras.

19.2. La epistemología civilizatoria de la IA actual

La IA actual encarna una epistemología civilizatoria concreta: la de la matriz occidental moderna en su configuración técnico-formal contemporánea. Sus tres pilares —la lógica formal como único criterio de racionalidad, el actor racional como único modelo de agencia, el sesgo anglosajón como único contenido legítimo— no son tres características técnicas entre otras: son los rasgos de una civilización epistémica que se materializa en instrumentos. La lógica formal se inscribe en el cálculo (álgebra lineal, probabilidad, optimización); el actor racional se inscribe en el modelo de agencia (aprendizaje por refuerzo, teoría de juegos, funciones de utilidad); el sesgo anglosajón se inscribe en los datos, las categorías y los valores. Juntos forman un dispositivo que filtra, traduce y configura todo lo que toca, con un efecto de cierre que vuelve invisibles otras formas de conocer. No se trata de un defecto técnico corregible con más datos o mejores métricas: es una consecuencia estructural de las decisiones epistemológicas que subyacen a su diseño. La IA actual no "describe mal" el mundo, sino que constituye un mundo en el que ciertas cosas existen como objetos de conocimiento y otras no. Y al hacerlo, reproduce y amplifica la civilización epistémica dominante: la naturaliza como forma universal de conocer, la materializa en infraestructura global, la impone como condición de posibilidad de la cognición legítima. No desarrollamos aquí estos conceptos —el lector puede volver a los capítulos 1 a 4 y a los apartados 19.1.2 y 19.1.3—, pero los recordamos porque son el objeto de nuestra no alineación.

19.3. La epistemología civilizatoria de la no alineación

Si la IA actual encarna la civilización epistémica dominante, la no alineación no se define solo por lo que rechaza. Se define, sobre todo, por lo que afirma: otra epistemología civilizatoria. No una matriz alternativa cerrada que sustituya a la dominante, sino una epistemología civilizatoria plural, que reconozca la diversidad de formas de conocer y no reduzca unas a otras. No se trata de "elegir" entre la matriz occidental y otras matrices: se trata de abrir el espacio para que múltiples epistemologías civilizatorias coexistan, dialoguen, se traduzcan con dificultad, se conflictúen sin destruirse. La no alineación no es un "no" a la IA: es un "sí" a otra IA, la que correspondería a otra epistemología civilizatoria.

Esta otra epistemología civilizatoria no existe como sistema. Existe como constelación de prácticas, tradiciones, resistencias y proyectos. Existe en las comunidades indígenas que defienden su soberanía epistémica; en los movimientos sociales que luchan por la justicia cognitiva; en los académicos críticos que desarrollan epistemologías alternativas; en los ingenieros disidentes que construyen prototipos no alineados; en las tradiciones no occidentales que sostienen formas de conocer marginadas; en los Estados del Sur global que reivindican soberanía tecnológica; en las instituciones internacionales que reconocen la dimensión epistémica de la IA. No forman un bloque, no comparten una ideología, no tienen un plan común. Pero convergen en la apuesta por otra epistemología civilizatoria. La no alineación no se libra desde un centro, sino desde múltiples puntos.

Los rasgos de esta epistemología civilizatoria de la no alineación se definen, en negativo, contra los seis rasgos de la civilización epistémica dominante. Frente a la razón discursiva como único criterio, una razón plural que reconozca la lógica formal y la dialéctica, la deducción y la analogía, el concepto y la imagen, la proposición y el cuerpo. Frente a la separación sujeto-objeto, una relacionalidad constitutiva que reconozca que el sujeto está inmerso en el mundo, que el conocimiento es relación antes que representación. Frente a la universalidad de la verdad, una situacionalidad del conocer que reconozca que toda verdad está situada, que hay múltiples verdades en conflicto, que la traducción entre tradiciones es precaria y necesaria. Frente a la voluntad de dominio, una voluntad de cuidado que reconozca que conocer no es dominar, sino habitar, cultivar, cuidar. Frente a la temporalidad lineal, una multiplicidad temporal que reconozca tiempos cíclicos, míticos, rituales, discontinuos. Frente a la institucionalización excluyente, una institucionalidad plural que reconozca instituciones diversas —comunitarias, tradicionales, académicas, técnicas— sin jerarquizarlas de antemano.

Estos seis rasgos no son abstractos: se traducen directamente en los tres pilares que hemos identificado a lo largo del trabajo. Frente a la lógica formal como único criterio de racionalidad, la razón plural implica un pluralismo lógico nativo: la lógica formal y la dialéctica, la paraconsistente y la intuicionista, la relevante y la polivalente, disponibles según el contexto, traducibles entre sí sin reducir unas a otras. Frente al actor racional como único modelo de agencia, la relacionalidad constitutiva implica una agencia relacional y colectiva: actores estratégicos

antes que optimizadores, comunidades epistémicas antes que agentes aislados, juego y conflicto antes que maximización de utilidad. Frente al sesgo anglosajón como único contenido legítimo, la situacionalidad del conocer implica una soberanía epistémica: *datasets* que no sean traducciones del inglés, categorías que no sean las anglosajonas, valores que no sean los liberales, tradiciones que no sean las analíticas, con participación real de las comunidades en el diseño, control de sus datos y retorno de valor. Los seis rasgos, así, no son un catálogo abstracto de virtudes epistémicas: son la traducción civilizatoria de los principios de diseño que hemos propuesto en el capítulo 17. La no alineación no es un gesto puramente negativo: es la afirmación de que otra IA es posible, y de que esa otra IA tiene una arquitectura epistémica concreta, no solo una intención crítica.

Estos rasgos no son una lista cerrada. Son horizontes que se cultivan, no estados que se alcanzan. No hay un momento en que la epistemología civilizatoria de la no alineación "esté completa": siempre está en proceso, siempre en disputa, siempre abierta. Por eso no hablamos de un proyecto —que implica un plan, un sujeto, un horizonte temporal acotado—, sino de una visión que se encarna en instrumentos, se cultiva en prácticas y se transmite en tradiciones. Una visión que no se posee, sino que se habita; que no se impone, sino que se sostiene; que no se clausura, sino que se abre.

La no alineación, en este sentido, no es una cuestión técnica. No se resuelve con mejores algoritmos, con más datos, con métricas más finas. Es una cuestión epistémica y civilizatoria: qué formas de conocer queremos sostener, qué mundos queremos habitar, qué sujetos queremos ser. La IA actual nos impone un modo de conocer y, con él, un modo de habitar el mundo. La no alineación es la apuesta por otro modo de conocer y, con él, por otro modo de habitar el mundo.

19.4. El conflicto

No hay una sola epistemología civilizatoria. La historia de la humanidad no es la historia de un único modo de conocer que se despliega linealmente, sino la historia del conflicto entre varios. Cada epistemología civilizatoria tiene sus instrumentos, sus prácticas, sus tradiciones, sus portadores. Cada una se impone o resiste, se naturaliza o se margina. Hablar de "epistemología civilizatoria" en singular es una abstracción: en la realidad hay varias en conflicto. La IA actual no es "la" epistemología civilizatoria: es una que se ha impuesto. La no alineación no es "la" alternativa: es una que resiste. Y ese conflicto no es un accidente de la historia: es su modo de ser.

El conflicto entre epistemologías civilizatorias no se resuelve por consenso. No hay un "debate racional" en el que las partes lleguen a un acuerdo, ni un "tribunal neutro" desde el cual arbitrar. Cada epistemología civilizatoria constituye el objeto del conocimiento de una manera que excluye otras. No hay un "mundo común" preexistente que distintas epistemologías describan de distintas maneras: hay mundos que se constituyen en distintas epistemologías, y esos mundos coexisten, se conflictúan, se traducen con dificultad. Esto no significa que el conflicto sea violento. Significa que no es neutral. La coexistencia es posible, pero precaria: requiere cultivo constante, reconocimiento mutuo, traducción entre tradiciones. Y requiere, sobre todo, poder: sin poder, la coexistencia es subordinación.

Si la no alineación es una apuesta por otra epistemología civilizatoria, entonces tiene adversarios. No se trata de demonizar a nadie: se trata de nombrar las fuerzas que se oponen a la apuesta. La monocultura epistémica —la imposición de un solo modo de conocer— opera a través de la infraestructura, los estándares, las métricas, los *benchmarks*. El extractivismo cognitivo —la extracción de datos del Sur global para beneficio del Norte— opera a través de los *datasets*, los modelos, los productos que se venden de vuelta al Sur. La reducción de lo humano a lo calculable —la identificación de la cognición con la computación— opera a través de las arquitecturas, las funciones de pérdida, las métricas. La naturalización —la presentación de una epistemología particular como universal— opera a través del sentido común, la educación, la infraestructura. La cooptación —el uso de la diversidad como marketing sin transformar las prácticas— opera a través de las "capas de diversidad", los "principios éticos", las "declaraciones de intención". Estos adversarios no son personas: son fuerzas que operan en el ecosistema de la IA. Nombrarlas es el primer paso para resistir a ellas.

Si la no alineación no tiene un sujeto único, entonces ¿quién la porta? No un individuo, no una empresa, no un Estado. Una constelación de portadores: comunidades indígenas que resisten a la extracción de datos y reivindican

su soberanía epistémica; movimientos sociales que luchan por la justicia cognitiva; académicos críticos que desarrollan epistemologías alternativas; ingenieros disidentes que construyen prototipos no alineados; tradiciones no occidentales que sostienen formas de conocer marginadas; Estados del Sur global que reivindican soberanía tecnológica; instituciones internacionales que reconocen la dimensión epistémica de la IA. Estos portadores no forman un bloque, no comparten una ideología, no tienen un plan común. Pero convergen en la apuesta por otra epistemología civilizatoria. El conflicto no se libra desde un centro, sino desde múltiples puntos. Y no se libra solo con argumentos: se libra con prácticas, con instituciones, con instrumentos, con alianzas.

El conflicto entre epistemologías civilizatorias no se gana con una victoria. Se cultiva a lo largo del tiempo. No hay un momento en que la no alineación "triunfe". Hay prácticas que se sostienen, instrumentos que se construyen, tradiciones que se transmiten, comunidades que se articulan. El cultivo no es pasivo: implica resistir a la monocultura, construir alternativas, articular portadores, materializar la visión en instrumentos. Pero no es voluntarista: no se trata de "llevar a cabo un plan", sino de sostener una dirección a lo largo del tiempo. La no alineación no es un estado que se alcanza: es un cultivo que se practica. No es un punto de llegada: es un modo de caminar. Y como todo cultivo, requiere paciencia, cuidado, persistencia, y la convicción de que lo que se siembra no se cosecha inmediatamente, pero se cosecha.

La pregunta que queda abierta no es "¿quién ganará?". La pregunta es "¿cómo convivir?". Porque el conflicto entre epistemologías civilizatorias no se resuelve con la desaparición de una de las partes. Se resuelve —cuando se resuelve— con la construcción de un espacio en el que múltiples epistemologías civilizatorias puedan coexistir sin que ninguna sea subordinada. Ese espacio no existe hoy. Hay que construirlo. Y construirlo es, precisamente, el trabajo de la no alineación.

19.5. La superación del conflicto: hacia una IA del conocimiento humano

El conflicto entre epistemologías civilizatorias no se resuelve por la victoria de una sobre las otras. Tampoco se resuelve por la simple tolerancia, que deja intactas las asimetrías. Se resuelve —cuando se resuelve— por superación: no la eliminación del conflicto, sino su transformación en un nivel superior. La noción de superación tiene una historia filosófica precisa: en Hegel, *Aufhebung* significa a la vez negar, conservar y elevar. Negar la absolutización de una epistemología; conservar sus logros; elevar el conjunto a un nivel en el que las diferencias ya no se resuelvan por dominación. No se trata de una síntesis que disuelva las diferencias, sino de una articulación en la que cada epistemología civilizatoria conserva su especificidad y contribuye al conjunto sin subordinarse. Superar el conflicto no es terminarlo: es habitarlo de otra manera.

Esta superación tiene un nombre concreto en el campo que nos ocupa: una IA del conocimiento humano. No "la" IA, no "una IA más", sino una IA que corresponda al conocimiento humano en toda su amplitud, y no a un subconjunto suyo. Hemos caracterizado el conocimiento humano completo a lo largo de este trabajo: no es solo el conocimiento formalizable, sino también el dialéctico, el relacional, el tácito, el práctico, el colectivo, el estético, el espiritual. No es solo el conocimiento de la matriz occidental moderna, sino también el de las tradiciones que esa matriz ha marginado. No es solo el conocimiento de los individuos aislados, sino también el de las comunidades epistémicas. No es solo el conocimiento que se puede medir, sino también el que se resiste a la métrica sin por ello ser menos riguroso. Una IA del conocimiento humano sería aquella que pudiera operar con todas esas formas de conocer, reconocerlas como legítimas, traducir entre ellas sin reducir unas a otras, y abrirse a las que no han sido anticipadas.

¿Qué significaría, en concreto, una IA del conocimiento humano? Significaría una IA cuyos *datasets* no fueran traducciones del inglés, sino que incorporaran categorías, lenguas y tradiciones diversas.

- Cuya arquitectura operara con lógicas plurales: la formal y la dialéctica, la
- Cuyas funciones de pérdida no redujeran el éxito a un escalar único, sino que
- Cuyas métricas midieran no solo la precisión predictiva, sino también la
- Cuyos *benchmarks* no
- Cuyas interfaces permitieran al usuario ver las tensiones, elegir las lógicas,

- Cuyos procesos de diseño integran a las comunidades afectadas, no como
- Cuyas formas de gobernanza reconocieran la soberanía epistémica de los pueblos,

occidental". No sería una IA "del Norte" ni una IA "del Sur". No sería una IA "del capital" ni una IA "del pueblo". Sería una IA pluricultural, dialógica, abierta, que reconociera la pluralidad de epistemologías civilizatorias y no subordinara ninguna. No sería la superación del conflicto por eliminación, sino por articulación: cada epistemología civilizatoria conservaría su especificidad y contribuiría al conjunto. No sería una síntesis que disuelva las diferencias, sino una constelación en la que las diferencias coexisten sin jerarquía previa.

Esta IA del conocimiento humano no existe. No existe como sistema, no existe como producto, no existe como proyecto terminado. Existe como horizonte: como dirección de trabajo, como apuesta, como cultivo. Hay prototipos parciales que apuntan en esa dirección: los razonadores neuro-simbólicos con lógicas no clásicas, los sistemas multiagente con debate estructurado, los marcos de computación relacional, las metodologías de diseño participativo descolonizador, los principios de soberanía de datos indígenas. No son "la" IA del conocimiento humano: son fragmentos que anticipan lo que podría ser. Pero son fragmentos reales, no utopías. Y su existencia demuestra que la apuesta no es un sueño: es una dirección viable, aunque larga, difícil y conflictiva.

La superación del conflicto, entonces, no es un estado final que se alcanza de una vez. Es un proceso que se cultiva a lo largo del tiempo. No hay un momento en que la IA del conocimiento humano "esté completa": siempre está en proceso, siempre en disputa, siempre abierta. Pero hay una dirección: la de una IA que no reduzca el pensamiento humano a lo calculable, que reconozca la diversidad epistémica, que se abra a las formas de conocer que la historia ha excluido. Y hay un criterio para evaluar los avances: ¿esta IA amplía el espacio de lo cognoscible, o lo reduce? ¿Reconoce la pluralidad de epistemologías civilizatorias, o impone una? ¿Se abre a lo no anticipado, o se cierra sobre sí misma? No se trata de medir el éxito por métricas estándar, sino por la ampliación del espacio de lo conocible y por la dignidad de las formas de conocer que sostiene.

Una IA del conocimiento humano no sería una IA "para" América Latina, ni "para" África, ni "para" Asia. Sería una IA para todos, porque toda epistemología es epistemología para todos. Las formas de conocer no tienen dueño: circulan, se traducen, se enriquecen mutuamente. Una IA que reconozca la lógica dialéctica no sería "una IA hegeliana": sería una IA capaz de operar con contradicción, para cualquiera que la necesite. Una IA que reconozca el conocimiento relacional no sería "una IA andina": sería una IA capaz de operar con relaciones constitutivas, para cualquiera que las habite. Una IA que reconozca la cognición colectiva no sería "una IA comunitaria": sería una IA capaz de operar con comunidades epistémicas, para cualquiera que las practique. La superación del conflicto no es la imposición de una epistemología sobre las otras: es la apertura de un espacio en el que todas puedan florecer sin dominarse.

Finalmente, una IA del conocimiento humano no sería una IA dialéctica. La dialéctica es el movimiento de lo real y ningún artefacto participa en ese movimiento porque los artefactos operan sobre representaciones del movimiento. Lo que una IA del conocimiento humano podría hacer es representar el movimiento con mayor riqueza, incorporando contradicción, proceso, totalidad, historicidad, sin pretender estar dentro del proceso. No sería dialéctica, sería una representación más fiel del movimiento dialéctico. Y eso ya es mucho.

Toda epistemología debería ofrecer el principal derecho humano, el derecho de conocer toda la realidad. No solo la realidad visible, la realidad registrada o la realidad que el sistema formal permite. Toda la realidad. Una epistemología que oculta la realidad es limitada, pero sobre todo es injusta. Y una IA que oculta la realidad no solo es imperfecta, viola un derecho. Y no es inteligente.

Bibliografía del capítulo 19

- Abe, J. M. (2025). *An Introduction to Curry Systems: A Way to Algebraize Paraconsistent, Paracomplete, and Non-Alethic Logic*. Springer.
- Allen, B. P., Chhikara, P., Ferguson, T. M., Ilievski, F., & Groth, P. (2025). Sound and Complete Neurosymbolic Reasoning with LLM-Grounded Interpretations. *Proceedings of Machine Learning Research*, 284.
- De Sousa Santos, B. (2010). *Descolonizar el saber, reinventar el poder*. Trilce.
- Dumas, G. (2026). Social NeuroAI of World Modeling: Embodied Dynamics, Collective Intelligence and Relational Computation. Simons Institute for the Theory of Computing.

- Hegel, G. W. F. (1807). *Fenomenología del espíritu* (W. Roces, Trad.). Fondo de Cultura Económica.
- Hegel, G. W. F. (1812-1816). *Ciencia de la lógica* (A. y R. Mondolfo, Trads.). Hachette.
- Mignolo, W. (2007). *La idea de América Latina*. Gedisa.
- Quijano, A. (2000). Colonialidad del poder, eurocentrismo y América Latina. En E. Lander (ed.), *La colonialidad del saber: eurocentrismo y ciencias sociales* (pp. 201–246). CLACSO.
- Walsh, C. (2009). *Interculturalidad, Estado, sociedad: Luchas (de)coloniales de nuestra época*. Universidad Andina Simón Bolívar.

Conclusiones

Este libro se propuso una tarea que, si bien tiene antecedentes en la epistemología social de la IA y en la filosofía de la computación, no había sido abordada de manera sistemática desde una perspectiva decolonial, dialéctica y latinoamericana: examinar la IA desde la epistemología y no solo desde la ética, la política o la ingeniería.

La literatura sobre IA es abundante en discusiones sobre alineación de valores, riesgos existenciales, sesgos algorítmicos y regulación. Es más escasa en análisis sobre qué tipo de conocimiento produce la IA, qué formas de conocer excluye, y qué consecuencias tiene esa exclusión para la diversidad epistémica de la humanidad. Este libro se inscribe en esa escasez y busca contribuir a llenarla.

En cuanto a los alcances logrados el libro articula un diagnóstico, una propuesta y una proyección.

En cuanto al diagnóstico, muestra que la IA actual encarna una epistemología particular —formal, individualista, anglosajona— que no es universal, sino histórica y culturalmente situada. Desmonta esa epistemología en tres pilares (lógica formal, actor racional, sesgo anglosajón), muestra cómo se articulan en un dispositivo epistémico con efecto de cierre, y analiza sus consecuencias: exclusión epistémica, colonialidad del dato, empobrecimiento cognitivo, naturalización. El diagnóstico no se limita a señalar que la IA "tiene sesgos": muestra que su arquitectura misma está configurada de manera que ciertas formas de conocer quedan fuera del campo de visibilidad. No es un defecto corregible con más datos o mejores métricas: es una consecuencia estructural.

En cuanto a la propuesta, el libro invierte el marco estándar de la alineación. No es que el sistema esté desalineado con nosotros, sino que nosotros no nos alineamos con la epistemología que el sistema encarna. No nos alineamos con la reducción del pensamiento humano a su subconjunto formalizable. Reclamamos una IA alineada con el pensamiento humano completo: con la contradicción, con el proceso, con la relación, con el tacto, con la práctica, con la colectividad, con el cuerpo, con la historia, con las tradiciones no occidentales. La propuesta no se queda en la crítica: desarrolla principios de diseño (pluralismo lógico, representaciones relacionales, conocimiento tácito, conflicto como recurso, trazabilidad, validación múltiple, soberanía epistémica), arquitecturas posibles (neuro-simbólica con lógicas no clásicas, multiagente con debate estructurado, computación relacional, diseño participativo descolonizador), metodologías de construcción, y condiciones de viabilidad.

En cuanto a la proyección, el libro introduce un concepto nuevo —el de epistemología civilizatoria— y lo aplica al caso de la IA. Una epistemología civilizatoria es el modo de conocer propio de una civilización en un periodo histórico determinado. Comprende los instrumentos, los conceptos y las cosmovisiones que esa civilización admite, y excluye los que no admite. Pero su rasgo decisivo es que constituye al objeto mismo del conocimiento: define qué se ve y qué no se ve en el mundo por conocer. La IA actual encarna una epistemología civilizatoria particular y la reproduce y amplifica. La no alineación es la apuesta por otra epistemología civilizatoria: una que reconozca la diversidad de formas de conocer, que no reduzca unas a otras, que se abra a lo que la historia ha excluido. Esa apuesta tiene un nombre concreto: una IA del conocimiento humano.

El libro plantea preguntas que, en la literatura sobre IA, no suelen formularse. Las enumero.

- Primera pregunta: ¿qué tipo de conocimiento produce, posee o excluye la IA? La
- Segunda pregunta: ¿es la epistemología de la IA universal o particular? La
- Tercera pregunta: ¿qué formas de conocer quedan excluidas por la IA, y qué
- Cuarta pregunta: ¿puede la IA alinearse con el pensamiento humano completo, o
- Quinta pregunta: ¿qué significa no alinearse con la epistemología dominante de la
- Sexta pregunta: ¿es la no alineación un bien público global, o un juego de suma
- Séptima pregunta: ¿qué es una epistemología civilizatoria y cómo se aplica a la IA?
- Octava pregunta: ¿qué sería una IA del conocimiento humano? La literatura sobre

con código. No pretende ser un manifiesto político, no prescribe acciones inmediatas. No pretende ser neutral, parte de una toma de posición explícita —que la diversidad epistémica es un valor, que la hegemonía de un solo

modo de conocer es un problema—. No pretende agotar el tema sino que es una cartografía, no un mapa detallado. Cada capítulo podría expandirse a un libro. Lo que ofrece es un marco, un diagnóstico, una propuesta y una invitación a seguir pensando.

El libro cierra con preguntas que no resuelve, porque no tienen respuesta simple. ¿Puede haber conocimiento sin sujeto? ¿Puede formalizarse la dialéctica completamente? ¿Es posible una IA verdaderamente pluricultural? ¿Cómo evaluar una IA dialéctica? ¿Cómo evitar la cooptación? ¿Cómo garantizar la responsabilidad? ¿Cómo escalar los prototipos? ¿Cómo financiar la investigación alternativa? ¿Cómo construir alianzas? Estas preguntas no son un defecto del libro: son su horizonte. Un libro que lo cerrara todo sería un libro dogmático.

Este libro prefiere dejar abierto lo que está abierto y ofrecer un marco para seguir pensándolo.

El libro dialoga, sin decirlo explícitamente, con una literatura que ha comenzado a pensar la IA desde la epistemología. No es una literatura abundante, pero existe y conviene situarla. Ramón Alvarado ha argumentado que la IA es una “tecnología epistémica”, es decir, un artefacto que no solo procesa información sino que configura las condiciones de lo cognoscible. Inkeri Koskinen ha mostrado que carecemos de una epistemología social satisfactoria para la ciencia basada en IA, porque las relaciones de confianza y dependencia epistémica que sostienen la producción colectiva de conocimiento no se traducen a los sistemas algorítmicos. John Symons y el propio Alvarado han analizado la injusticia epistémica que producen las tecnologías de ciencia de datos cuando reducen a los sujetos a puntos de datos sin agencia cognitiva.

El aporte de este libro se sitúa en otra tradición. No parte de la filosofía analítica de la ciencia ni de la epistemología social anglosajona, sino de la crítica decolonial latinoamericana (Quijano, Mignolo, De Sousa Santos) y de la tradición dialéctica (Hegel, Marx, Adorno). Su contribución no es la de haber “descubierto” que la IA tiene una epistemología, sino la de haber mostrado que esa epistemología es civilizatoria, que no se limita a filtrar contenidos, sino que constituye el objeto mismo del conocimiento y que su hegemonía global reproduce la matriz occidental moderna en su configuración técnico- formal contemporánea. Esa es la tesis que el libro defiende y es ahí donde se distingue de la literatura existente.

Bibliografía de las conclusiones

- Alvarado, R. (2023). *AI as an Epistemic Technology*. Philosophy of Science Archive.
- De Sousa Santos, B. (2021). *El fin del imperio cognitivo: La afirmación de las*
- Koskinen, I. (2024). We Have No Satisfactory Social Epistemology of AI-Based
- Symons, J., & Alvarado, R. (2022). Epistemic Injustice and Data Science

Glosario

Actor racional. Modelo de agencia heredado de la economía neoclásica y la teoría de la decisión, que presupone preferencias completas y transitivas, optimización bajo restricciones, maximización de utilidad esperada e individualismo metodológico. Es el segundo pilar de la epistemología implícita de la IA. (Cap. 2, 13.)

Actor estratégico. Modelo alternativo de agencia desarrollado por Crozier y Friedberg, según el cual el actor organizacional no optimiza, sino que juega en un contexto de incertidumbre, buscando poder en relaciones de dependencia mutua. Sustituye la optimización por la racionalidad limitada, las preferencias por el poder, la planificación por el juego. (Cap. 6.)

Agencia relacional. Forma de agencia en la que el sujeto no se constituye como individuo aislado, sino en relaciones que lo definen. Es central en el confucianismo, el ubuntu, el pensamiento andino y el budismo. (Cap. 5.)

Alineación. En la literatura estándar, problema de hacer que los sistemas de IA actúen conforme a los valores e intenciones humanos. El libro invierte este marco: no se trata de que el sistema esté desalineado con nosotros, sino de que nosotros no nos alineamos con la epistemología que el sistema encarna. (Cap. 10, 11.)

Aufhebung. Concepto hegeliano que designa la superación dialéctica como negación, conservación y elevación a un nivel superior. El libro lo recupera para pensar la superación del conflicto entre epistemologías civilizatorias. (Cap. 3, 19.)

Ayni. Principio andino de reciprocidad entendido como obligación cósmica de equilibrio, no como intercambio utilitario. Ha sido operacionalizado en el Método Multi- Neutrosófico Ayni para ética de IA. (Cap. 5.)

Bien público global. Bien que cumple con las condiciones de no rivalidad y no exclusión, y que requiere cooperación internacional para su provisión. El libro argumenta que la no alineación es un bien público global. (Cap. 16.)

Catuskoti (tetralema). Estructura lógica del budismo madhyamaka que considera cuatro alternativas (existe, no existe, ambos, ninguno) y a menudo las niega todas para mostrar la vacuidad de los conceptos. (Cap. 5.)

Civilización epistémica dominante. Configuración del conocer que se ha impuesto globalmente a partir de la expansión europea, la revolución industrial y la globalización. Se caracteriza como matriz occidental moderna: occidental, moderna y dominante. (Cap. 19.)

Colonialidad del dato. Extensión del concepto de colonialidad del poder al ámbito digital. Designa la forma en que las relaciones y lógicas coloniales se replican en la recolección, procesamiento y uso de datos, reproduciendo jerarquías epistémicas a escala planetaria. (Cap. 4.)

Colonialidad del poder. Concepto de Aníbal Quijano según el cual la colonización no terminó con la independencia formal, sino que persiste como estructura de jerarquías epistémicas, raciales y económicas. (Cap. 4, 19.)

Conocimiento tácito. Concepto de Michael Polanyi: conocimiento que una persona o comunidad posee incorporado y que es difícil o imposible de explicar completamente. Se transmite por socialización, no por información explícita. (Cap. 6.)

Contradicción productiva. Tesis dialéctica según la cual las contradicciones no son errores a eliminar, sino motores del conocimiento. Un sistema que las elimina prematuramente se empobrece. (Cap. 3.)

Coordinación epistémica global. Convergencia en un marco único de conocimiento. El libro la rechaza en las condiciones actuales, porque solo puede realizarse bajo mandatos geopolíticos desiguales. (Cap. 16.)

Devenir. Categoría dialéctica que designa el proceso histórico de constitución de los fenómenos. La lógica dialéctica historiza las categorías; la lógica formal las trata como fijas. (Cap. 3.)

Dispositivo epistémico. Articulación de los tres pilares de la epistemología implícita de la IA (lógica formal, actor racional, sesgo anglosajón) en un sistema que filtra, traduce y configura todo lo que toca, con un efecto de cierre que vuelve invisibles otras formas de conocer. (Cap. 2.)

Efecto de cierre. Mecanismo por el cual el dispositivo epistémico de la IA vuelve invisibles las formas de conocer que no se ajustan a sus criterios. No se rechaza nada explícitamente: simplemente no se ve. (Cap. 2.)

Empobrecimiento cognitivo. Reducción de la riqueza de las formas humanas de conocer a un subconjunto estrecho, formalizable y optimizable. No afecta solo a las culturas marginalizadas: es un problema de todos. (Cap. 1, 4.)

Epistemología civilizatoria. Concepto central del libro. Modo de conocer propio de una civilización en un periodo histórico determinado. Comprende los instrumentos, los conceptos y las cosmovisiones que esa civilización admite, y excluye los que no admite. Su rasgo decisivo es que constituye al objeto mismo del conocimiento: define qué se ve y qué no se ve. (Cap. 19.)

Epistemología encarnada. Tesis según la cual toda tecnología de conocimiento —y la IA lo es de manera eminente— presupone y pone en práctica una manera concreta de entender qué es conocer, cómo se adquiere, cómo se valida y quién puede poseerlo. (Cap. 1.)

Epistemología geopolítica. Estudio de la dimensión geopolítica del conocimiento. Su tesis central es que conocer no es una operación neutral de contemplación, sino una operación de poder: nombrar es poseer. (Cap. 18.)

Epistemología implícita de la IA. Trasfondo de supuestos epistémicos que operan sin ser explicitados en el diseño, entrenamiento y evaluación de sistemas de IA. Se distingue de las filosofías explícitas que los diseñadores profesan. (Cap. 2.)

Epistemología social. Tradición que estudia la dimensión social de la producción de conocimiento. El libro distingue entre versiones débiles (Goldman, Kitcher), que tratan lo social como contexto, y versiones constitutivas (Fuller), que lo tratan como condición de posibilidad. (Cap. 7.)

Exclusión epistémica. Primera consecuencia del dispositivo epistémico de la IA: las formas de conocer que no se ajustan a sus criterios quedan fuera del campo de reconocimiento del sistema. No desaparecen, pero no cuentan como conocimiento válido. (Cap. 1.)

Falacia de la neutralidad técnica. Creencia de que la IA es solo un instrumento y que su epistemología es la epistemología. El libro la desmonta mostrando que toda técnica tiene una política epistémica oculta. (Cap. 1.)

Interlocución entre marcos. Alternativa a la coordinación epistémica global. No requiere simetría perfecta, pero sí requiere que ninguno de los interlocutores tenga el poder de definir unilateralmente las reglas del encuentro. Permite que marcos diversos se encuentren, se traduzcan parcialmente y negocien sin disolverse. (Cap. 16.)

Jurisdicción epistémica. Concepto que el libro introduce para precisar la diferencia entre consentimiento y autoridad. El consentimiento puede otorgarse sin ceder el poder de nombrar; la jurisdicción, en cambio, lo supone. La soberanía epistémica exige algo más que la firma: exige participar en la redacción. (Cap. 14.)

Lógica dialéctica. Tradición filosófica que sostiene que la contradicción es constitutiva del conocimiento. Se define por tres tesis: contradicción productiva, totalidad y devenir. (Cap. 3.)

Lógica formal. Estudio de las inferencias válidas en virtud de su forma, independientemente de su contenido. Es el primer pilar de la epistemología implícita de la IA. Sus límites estructurales incluyen el principio de no contradicción, la clausura deductiva, la indiferencia al contenido, la bivalencia, la extensionalidad y los binarios. (Cap. 2.)

Matriz occidental moderna. Categoría gruesa que designa la configuración histórica del conocer que se ha impuesto globalmente. Se caracteriza por seis rasgos: razón discursiva, separación sujeto-objeto, temporalidad lineal, universalidad de la verdad, voluntad de dominio e institucionalización del saber. (Cap. 19.)

Método Multi-Neutrosófico Ayni. Desarrollo reciente que integra la lógica andina y la lógica neutrosófica para guiar el desarrollo ético de IA. Usa conjuntos multi- neutrosóficos para modelar juicios éticos plurales y una medida de consenso euclidiana para cuantificar la alineación entre stakeholders sin borrar las diferencias epistémicas. (Cap. 5.)

No alineación. En el marco estándar, defecto del sistema que debe ser corregido. En la inversión que propone el libro, posición del sujeto que rechaza la epistemología que el sistema encarna. No es un “no” a la técnica, sino un “no” a una civilización epistémica y un “sí” a otra. (Cap. 10.)

Nombrar es poseer. Tesis central de la epistemología geopolítica. Cuando un sujeto nombra una cosa, la inscribe en su propio sistema de categorías y la convierte en objeto de conocimiento. No es una metáfora: es una operación real con consecuencias concretas. (Cap. 18.)

Pachakutiy. Categoría dialéctica andina que designa la inversión desde el interior: lo que latía oculto se hace manifiesto, lo que estaba abajo va arriba. No es apocalipsis, sino transformación estructural. (Cap. 5.)

Pensamiento grupal (grouphink). Patología epistémica descrita por Irving Janis: un grupo cohesionado toma decisiones defectuosas porque la presión hacia la unanimidad suprime la evaluación crítica de alternativas. Los sistemas multiagente de IA reproducen este patrón. (Cap. 7.)

Pensamiento humano completo. Conjunto abierto y plural de las formas de conocer que los humanos practicamos, en su diversidad histórica y cultural, excluyendo aquellas que la propia humanidad ha identificado como patológicas o injustas. Es un concepto normativo y abierto. (Cap. 11, 12.)

Pluralismo lógico. Principio de diseño según el cual ningún sistema único de lógica captura la totalidad de las formas válidas de inferencia. Un sistema epistémicamente robusto debe integrar múltiples lógicas y seleccionar o combinar según el contexto. (Cap. 8, 17.)

Principios CARE. Collective Benefit, Authority to Control, Responsibility, Ethics. Marco para la gobernanza de datos indígenas que operacionaliza la soberanía epistémica. (Cap. 4, 8, 14.)

Proxy teleology (teleología proxy). Concepto que describe la capacidad de los LLM de realizar movimientos dialécticos (de la indeterminación a la determinación, del conflicto a la resolución) sin poseer el fin que esos movimientos persiguen. La dialéctica en los LLM es performática, no constitutiva. (Cap. 3, 13.)

Re-alineación. Tarea que se deriva de la inversión de la alineación: no “alinearse” la IA (ya está alineada con un subconjunto de lo humano), sino re-alinearla con el pensamiento humano completo. (Cap. 11.)

Sesgo anglosajón. Sobrerepresentación estructural de datos, categorías, valores y tradiciones del mundo anglosajón en los sistemas de IA. Es el tercer pilar de la epistemología implícita. (Cap. 2, 4.)

Soberanía epistémica. Control sobre los datos, las categorías, los valores y las formas de conocer. Condición para que la auto-visión de una región no sea ilusoria. Se distingue de la soberanía algorítmica: no se trata solo de decidir qué sistemas se implementan, sino de quién define las categorías mismas. (Cap. 14.)

Sunyata (vacuidad). Concepto central del budismo madhyamaka: nada existe por sí mismo. Las cosas existen convencionalmente, en dependencia de otras cosas, pero no últimamente. (Cap. 5.)

Tawa. Principio andino de reciprocidad y equilibrio cósmico. (Cap. 5.)

Tetraléctica andina. Ampliación de la dialéctica hegeliana con un cuarto estado: superación (pléresis). Los cuatro estados son afirmación, negación, combinación y superación. La diferencia con Hegel es que el cuarto momento no clausura el proceso. (Cap. 5.)

Totalidad concreta. Concepto de Lukács según el cual el todo no es agregado de partes, sino totalidad que constituye a las partes. Las partes solo se comprenden en su relación con el todo. (Cap. 3.)

Traducción interburocrática. Trabajo cognitivo que los ciudadanos y los propios burócratas realizan para hacer sentido de la información burocrática. Incluye buscar claridad, resolver contradicciones, corroborar fuentes, transportar información entre organismos y derivar conocimiento útil de la experiencia. (Cap. 6.)

Ubuntu. Filosofía africana que sostiene que la relacionalidad precede al pensamiento. Su formulación condensada es “I relate, I am” (me relaciono, luego existo). Implica una teoría de la cognición colectiva. (Cap. 5.)

Violencia epistémica. Marginación, invisibilización o deslegitimación de formas de conocer que no se ajustan a los criterios dominantes. No es un ataque explícito, sino un efecto estructural del dispositivo epistémico. (Cap. 2.)

Wisdom Loop. Marco de diseño de IA desarrollado en África que incorpora la filosofía ubuntu en todas las fases del ciclo de diseño: valores primero, co-diseño participativo, datos como confianza, interpretación colectiva, retorno a la comunidad. (Cap. 5.)

Yanantin. Concepto central del pensamiento andino: dualismo complementario. Los opuestos no se superan, se complementan en una relación de interdependencia sin destrucción. (Cap. 5.)