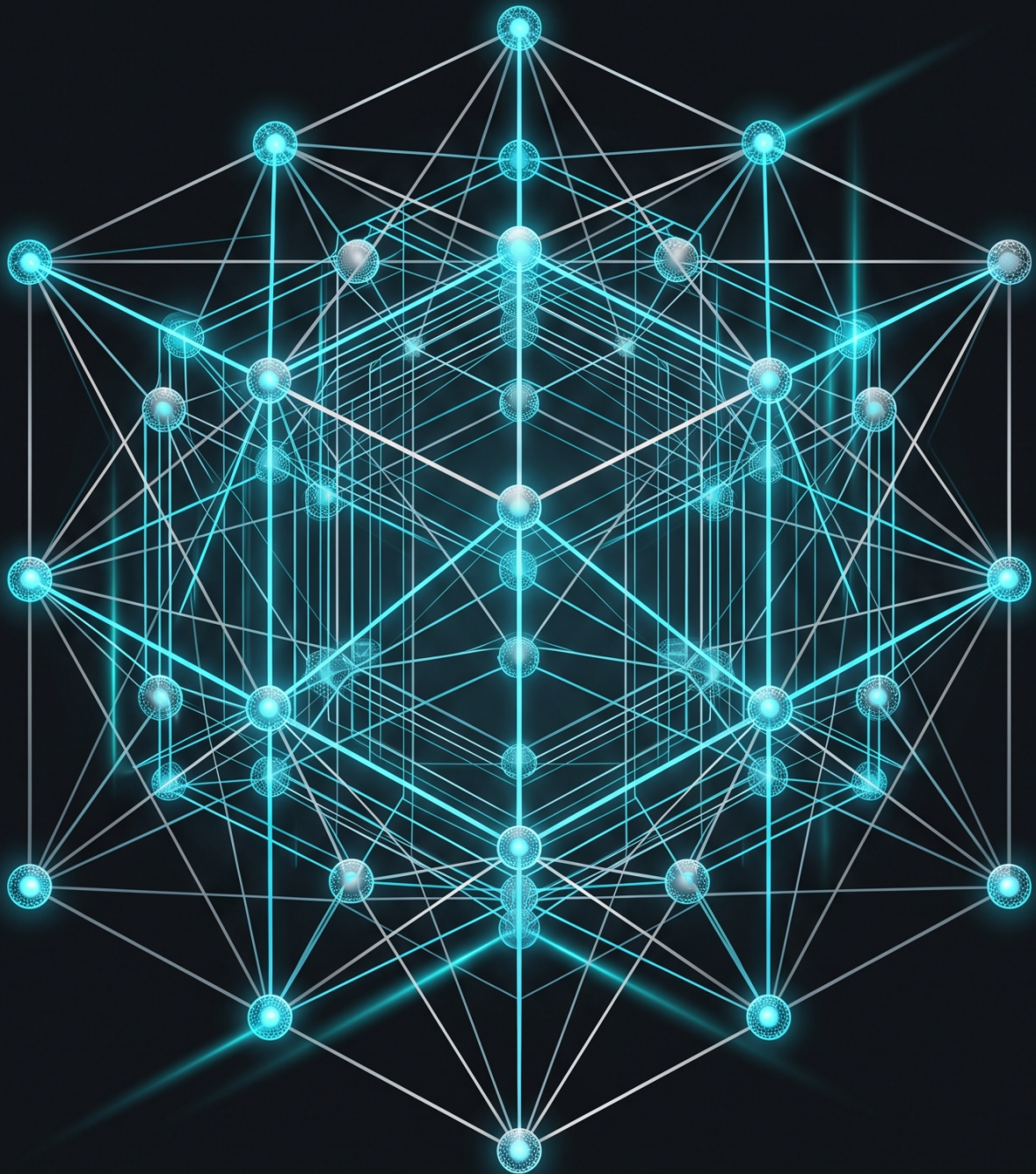


INGENIERÍA JURÍDICA

Una introducción crítica a la justicia
asistida por inteligencia artificial



Dr. Andrés Saavedra Avendaño

INGENIERÍA JURÍDICA

*Una introducción crítica a la justicia
asistida por inteligencia artificial*

Dr. Andrés Saavedra Avendaño

Cuernavaca, Morelos
2026

INGENIERÍA JURÍDICA

Una introducción crítica a la justicia asistida por inteligencia artificial

Primera edición, 2026

© Andrés Saavedra Avendaño, 2026

Publicado por Artistas Legales A.C.

Cuernavaca, Morelos, México

Esta obra se distribuye bajo licencia Creative Commons
Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0).

Se permite copiar, distribuir y adaptar la obra con fines no comerciales, siempre que se atribuya la autoría al Dr. Andrés Saavedra Avendaño y las obras derivadas se distribuyan bajo la misma licencia.

Texto completo de la licencia disponible en:
<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>

Nota sobre el uso de inteligencia artificial en la elaboración de esta obra.

En congruencia con el Principio 3 (Transparencia) formulado en el Capítulo 11, se declara que la redacción del manuscrito se apoyó en herramientas de inteligencia artificial generativa bajo supervisión, dirección editorial y revisión integral del autor. Las tesis, la estructura conceptual, las decisiones argumentativas y la responsabilidad íntegra del contenido son exclusivamente del autor.

Hecho en México / *Made in Mexico*

Índice

Prólogo — Hipótesis de trabajo

PARTE I — DIAGNÓSTICO

Capítulo 1. El derecho como tecnología social

Capítulo 2. La obsolescencia funcional del sistema jurídico

Capítulo 3. Tres aporías contemporáneas del acceso a la justicia

PARTE II — MARCO CONCEPTUAL

Capítulo 4. La inteligencia artificial como tecnología epistémica

Capítulo 5. Sesgos, opacidad y la ilusión de neutralidad

PARTE III — MÉTODO

Capítulo 6. Anatomía del prompt jurídico: las cinco capas

Capítulo 7. Detección asistida de conexiones normativas

Capítulo 8. Modelado predictivo del adversario procesal

Capítulo 9. El prompt que escribe prompts, y la división del trabajo entre agentes

Capítulo 10. Patologías frecuentes del prompt mal diseñado

PARTE IV — IMPLICACIONES

Capítulo 11. Diez principios para una práctica responsable

Epílogo — Preguntas abiertas

Fuentes consultadas

Sobre el autor

Prólogo

Hipótesis de trabajo

Este libro no proclama doctrinas: propone hipótesis. La distinción importa y conviene declararla antes que cualquier otra cosa.

Quien escribe es jurista y observa, desde la práctica forense en Morelos y desde la academia, una disonancia que ya no admite postergación. En tres décadas, la sociedad ha transitado del expediente físico al lenguaje natural conversacional con máquinas. El sistema jurídico que la rige, en cambio, sigue operando con la arquitectura conceptual de hace dos siglos. Esa disonancia no es estética; es funcional, y tiene costos verificables sobre vidas concretas.

El ensayo que sigue parte de cuatro hipótesis. No son convicciones cerradas; son proposiciones formuladas para ser puestas a prueba —por el lector, por la práctica forense, por el tiempo.

Primera hipótesis. El derecho es una tecnología social, no una ciencia natural. Como toda tecnología, fue diseñada para resolver problemas específicos en contextos específicos; como toda tecnología, puede volverse obsoleta. Tratarlo como ciencia inmutable no es un acto de respeto intelectual: es un obstáculo para su mejora.

Segunda hipótesis. El sistema jurídico contemporáneo, particularmente en países en desarrollo, presenta indicios verificables de obsolescencia funcional: tiempos procesales que exceden la vida útil de los conflictos, asimetría estructural de acceso y una brecha lingüística que convierte la norma en idioma extranjero para el ciudadano común. Esos indicios son empíricamente medibles y, por tanto, susceptibles de refutación.

Tercera hipótesis. La inteligencia artificial generativa es una tecnología epistémica —una tecnología que altera las condiciones bajo las cuales producimos, validamos y comunicamos conocimiento. Aplicada al derecho, no reemplaza al jurista: redefine las competencias que constituyen su oficio. La pregunta correcta no es si reemplazará al abogado, sino cuáles de sus tareas actuales conservarán valor cuando la herramienta esté universalmente disponible.

Cuarta hipótesis. La calidad de la justicia asistida por inteligencia artificial depende menos de la sofisticación del modelo y más del diseño de las instrucciones que lo orientan. La técnica del prompt —su contexto, su precisión técnica, su fundamentación sustantiva, su estrategia y su estilo— constituye una competencia jurídica del siglo XXI, no una habilidad informática accesorio.

Estas cuatro proposiciones organizan los capítulos que siguen. Las primeras tres se examinan en las partes I y II; la cuarta articula el método expuesto en la parte III; las implicaciones éticas de aceptarlas, discutidas en la parte IV, son las que justifican —si algo lo justifica— el ejercicio entero.

Una advertencia metodológica. Este libro no es un manual ni un tratado. No agota un campo: lo introduce críticamente. Su brevedad es deliberada; la economía expositiva pertenece al método, no a la pereza. Donde otros textos ofrecerían certezas, aquí se ofrecen problemas mejor formulados. El lector que busque instrucciones cerradas hallará algunas en la parte III, pero hará bien en leerlas con la misma desconfianza con que el autor las propone.

Quien escribe no aspira a tener razón. Aspira a contribuir a una discusión que, en términos prácticos, ya ocurre todos los días en juzgados, despachos y comunidades sin que sus protagonistas la nombren así. Nombrarla es el primer paso para conducirla con responsabilidad.

— A. S. A.

Cuernavaca, Morelos, 2026

PARTE I

Diagnóstico

El derecho como tecnología social

La distinción más útil que ofrece la filosofía clásica del conocimiento es también la más olvidada por el discurso jurídico contemporáneo: la ciencia describe lo que es; la técnica produce lo que debe ser. Aristóteles llamó *episteme* a la primera y *techné* a la segunda. La biología, la física y la química son ciencias en sentido estricto: buscan la verdad de aquello que existe con independencia de la voluntad humana. La medicina, la arquitectura y la ingeniería son técnicas: producen objetos —tratamientos, edificios, puentes— diseñados por humanos para resolver problemas humanos.

¿Dónde está el derecho? La respuesta dominante durante los últimos dos siglos ha sido: en la ciencia. Existe una "ciencia jurídica", una "dogmática", una pretensión sistemática de que el derecho posee un objeto propio, un método propio y verdades propias. Esa autoadscripción no es inocente: convierte al jurista en una suerte de astrónomo de las normas, alguien que las descubre en lugar de alguien que las diseña.

La tesis que articula este capítulo —y, por extensión, el libro entero— es que esa autoadscripción descansa en un error de categoría. El derecho no es una ciencia. Es una tecnología. Más precisamente, una *tecnología social*.

¿Qué significa "tecnología social"?

Una tecnología, en sentido riguroso, es un artefacto deliberadamente diseñado para producir un resultado específico bajo condiciones específicas. Un martillo lo es. Un protocolo de cirugía también. Lo mismo un algoritmo de búsqueda. Las tres comparten cuatro rasgos: tienen autor —individual o institucional—; tienen propósito declarado; están sujetas a evaluación funcional —cumplen o no cumplen lo que prometen—; y son susceptibles de obsolescencia —otra tecnología puede reemplazarlas.

Cuando una tecnología no consiste en un objeto material sino en un conjunto de reglas, procedimientos e instituciones diseñadas para coordinar conducta colectiva, hablamos de tecnología social. El dinero es una tecnología social. La doble entrada contable es una tecnología social. El método científico mismo —y aquí la paradoja es solo aparente— es también una tecnología social: un conjunto de procedimientos diseñado para producir un tipo específico de conocimiento. El derecho, entendido como sistema de normas, procedimientos e instituciones destinado a resolver conflictos y coordinar conducta, encaja sin residuo en esta definición.

Por qué la confusión persiste

La identificación del derecho con la ciencia tiene una historia rastreable. La escuela histórica alemana del siglo XIX, las pretensiones sistemáticas de la pandectística y el rigor formal del positivismo kelseniano contribuyeron, cada uno a su modo, a vestir de ciencia lo que en realidad era —y sigue siendo— ingeniería normativa. La motivación no fue puramente intelectual; también fue corporativa. Un saber inaccesible se profesionaliza mejor que un oficio compartido, y un experto en una "ciencia" cobra más que un técnico en una tecnología revisable.

Cuatro consecuencias del cambio de categoría

Aceptar que el derecho es tecnología, no ciencia, tiene consecuencias operativas inmediatas.

Primera: el derecho es *evaluable*. Una tecnología se juzga por su eficacia, no por su consistencia interna. Que un sistema procesal sea coherente con sus propios principios dice poco sobre si resuelve los conflictos para los que fue diseñado.

Segunda: el derecho es *rediseñable*. Las tecnologías se actualizan. La pretensión de que ciertas instituciones jurídicas son intocables porque son "esenciales" confunde la función con la implementación particular que la realiza.

Tercera: el derecho es *auditable*. Toda tecnología admite peritaje externo. La idea de que solo los juristas pueden evaluar el derecho equivale a sostener que solo los pilotos pueden evaluar la seguridad de los aviones —un absurdo que la ingeniería desmintió hace décadas.

Cuarta: el derecho puede volverse *obsoleto*. Las tecnologías caducan cuando el entorno para el que fueron diseñadas cambia más rápido que ellas. Esta consecuencia, la más incómoda, ocupará el capítulo siguiente.

Una objeción anticipada

El jurista dogmático responderá que la dogmática jurídica sí es ciencia: que tiene un objeto formal —el ordenamiento— y un método propio —la interpretación sistemática. La objeción es atendible, pero confunde dos niveles. Que existan disciplinas teóricas *sobre* el derecho —filosofía del derecho, sociología jurídica, teoría general— no convierte al derecho mismo en ciencia, así como la existencia de la filosofía de la medicina no convierte a la medicina en una ciencia natural. El objeto sigue siendo un artefacto humano diseñado con propósitos humanos.

Si el derecho es una tecnología, y si toda tecnología puede volverse obsoleta, la pregunta que sigue no es ideológica sino diagnóstica: ¿qué indicadores observables permitirían concluir que el sistema jurídico

contemporáneo está atravesando un proceso de obsolescencia funcional?

La obsolescencia funcional del sistema jurídico

Toda tecnología envejece de dos maneras. Envejece materialmente —se gasta, se rompe, se deteriora— y envejece funcionalmente: deja de cumplir el propósito para el que fue diseñada, incluso cuando sus piezas siguen intactas. La segunda forma de envejecimiento es la más peligrosa, porque puede prolongarse durante décadas sin que sus operadores la perciban como tal.

¿Cuándo, exactamente, una tecnología social se vuelve funcionalmente obsoleta? La respuesta, propuesta aquí como criterio operativo, descansa en tres condiciones: cuando sus productos dejan sistemáticamente de corresponder al propósito declarado; cuando tecnologías alternativas producen los mismos resultados con costos significativamente menores; y cuando el costo de mantener su operación excede el beneficio que genera para quienes la usan. Las tres condiciones son verificables, ninguna depende de juicios ideológicos, y las tres —planteadas como hipótesis sobre el sistema jurídico contemporáneo— devuelven indicios convergentes.

Este capítulo propone tres indicadores empíricos para someter esa hipótesis a prueba.

Primer indicador: la deriva temporal

La función declarada de un sistema procesal es resolver conflictos. Un conflicto, sin embargo, tiene una vida útil: el momento más allá del cual su resolución, incluso favorable, pierde sentido práctico para las partes. El trabajador despedido que recibe su indemnización siete años después; la madre que obtiene la custodia cuando los hijos ya son adultos; el comerciante que cobra una factura cuando su empresa cerró por insolvencia. En todos estos casos el sistema, en sentido formal, "funcionó": dictó sentencia, ejecutó resoluciones, cerró expedientes. En sentido funcional, falló: produjo su resultado después de que el problema que justificaba el proceso había desaparecido.

Los tiempos procesales medios en México y en buena parte de América Latina —referidos por el *Rule of Law Index* del World Justice Project, los censos del INEGI sobre impartición y procuración de justicia, y los estudios del Banco Mundial sobre eficiencia judicial— se cuentan en años, no en meses, para materias civiles, mercantiles y laborales. Esa duración no es un dato neutral: es la distancia entre la promesa institucional y el resultado entregable. Cuando esa distancia supera la vida útil del conflicto, hay obsolescencia funcional, aunque el sistema continúe operando con normalidad aparente.

Segundo indicador: la asimetría estructural de acceso

Una segunda función declarada del sistema jurídico es ofrecer protección equivalente a todos los sujetos a su jurisdicción. La igualdad ante la ley es un postulado constitucional. El indicador funcional, sin embargo, es distinto: ¿accede igual quien tiene recursos económicos, formación educativa y capital social que quien carece de los tres?

Los datos son conocidos. En los censos nacionales sobre impartición de justicia, una mayoría sustantiva de los conflictos legales reportados por ciudadanos en condición de pobreza no llega a procedimiento formal, según estudios de acceso a la justicia disponibles. Esa ausencia no es voluntaria. Es producto del costo, la distancia, la

duración y la barrera técnica que el sistema mismo impone a su entrada.

Una tecnología que cumple su función para el quintil superior de sus usuarios potenciales y falla con el resto no se está distribuyendo desigualmente: está demostrando que su diseño no contempla a la mayoría de la población para la que nominalmente fue creada.

Tercer indicador: la brecha lingüística

El tercer indicador es más sutil y, por eso, más revelador. El sistema jurídico opera en un dialecto técnico —tecnicismos procesales, latinismos, fórmulas estilizadas, sintaxis ornamentada— que el ciudadano común no comprende sin intermediación profesional. Esa intermediación no es opcional: sin ella, el sujeto regulado por la norma no puede leer la sentencia que lo afecta ni elaborar el escrito que lo defiende.

Las consecuencias son verificables. Los estudios sobre comprensión de resoluciones judiciales en México y en países de tradición jurídica continental, recogidos en el movimiento internacional por el lenguaje claro aplicado al derecho, muestran que el lector no especializado, aun con escolaridad media superior completa, capta sólo una fracción menor del contenido sustantivo de la decisión que le concierne. Una tecnología cuya operación requiere obligatoriamente intérpretes especializados para que el usuario final entienda qué le ocurrió no es una tecnología de coordinación social: es una tecnología de gestión profesional cerrada.

El argumento del "siempre ha sido así"

Una objeción previsible: estos tres problemas no son nuevos, han acompañado al sistema jurídico desde su origen, y por tanto no constituyen obsolescencia sino condición estructural. La objeción confunde antigüedad con inevitabilidad. Que un problema haya existido siempre no implica que no pueda resolverse; implica únicamente que las tecnologías disponibles hasta ahora no lo resolvieron. La pregunta funcional pertinente no es si los problemas existen, sino si hay tecnologías alternativas capaces de mitigarlos a costo razonable. Cuando esas alternativas existen y no se incorporan, el diagnóstico cambia: no es obsolescencia inevitable, es obsolescencia mantenida.

Si los tres indicadores apuntan en la misma dirección —y la evidencia disponible sugiere que así es—, el sistema jurídico contemporáneo no enfrenta una crisis pasajera ni una distorsión coyuntural. Enfrenta tres aporías estructurales: la del acceso, la de la complejidad y la de la confianza. Examinarlas es la tarea del capítulo siguiente.

Tres aporías contemporáneas del acceso a la justicia

Una aporía, en sentido riguroso, no es un problema difícil: es un problema que resiste solución desde dentro del marco que lo formula. Los indicadores empíricos del capítulo anterior —tiempos procesales, asimetría de acceso, brecha lingüística— pueden leerse como síntomas, y como síntomas admiten remedios paliativos: más jueces, más juzgados, traductores, defensorías públicas, programas de oralidad. Pero cuando esos remedios se aplican durante décadas sin que los indicadores cedan, la hipótesis debe cambiar. No estamos frente a problemas técnicos pendientes de solución; estamos frente a paradojas constitutivas del propio sistema.

Este capítulo propone tres.

Primera aporía: el acceso que se cancela protegiendo

El sistema jurídico moderno se construyó como respuesta a la arbitrariedad. Cada institución procesal —el derecho a la defensa, la doble instancia, el principio de contradicción, el control de constitucionalidad, las cargas probatorias formales— fue añadida para proteger al ciudadano frente al poder. La acumulación de esas protecciones, sin embargo, produce un resultado contraintuitivo: cuanto más se protege el procedimiento, más costoso se vuelve activarlo; cuanto más costoso, menos ciudadanos pueden hacerlo; cuanto menos lo hacen, más concentrado queda el acceso en quienes ya tenían recursos para superar las barreras.

La paradoja es estructural. No puede resolverse eliminando garantías —eso reintroduciría la arbitrariedad— ni añadiendo más —eso

profundizaría la barrera. La tecnología jurídica enfrenta aquí un límite interno: las herramientas que la legitiman son las mismas que la inutilizan para la mayoría.

Segunda aporía: la complejidad necesaria que destruye la comprensión

Una norma justa debe contemplar las particularidades del caso al que se aplica. Una sociedad compleja produce casos heterogéneos. Por lo tanto, un sistema jurídico que aspire a la justicia material debe ser, necesariamente, complejo: distinciones finas, excepciones, cláusulas especiales, jurisprudencia interpretativa, doctrina especializada.

Esa complejidad, sin embargo, tiene un costo cognitivo: vuelve la norma incomprensible para el ciudadano sujeto a ella. Y una norma que el ciudadano no comprende no puede regular su conducta de forma reflexiva; solo puede aplicarse sobre él. La pretensión de justicia material termina, paradójicamente, alimentando una relación heterónoma con el derecho: el sujeto no obedece la norma porque la entienda, sino porque alguien con autoridad técnica le dice que está obligado.

La aporía no admite reducción simple. Simplificar la norma erosiona su capacidad de hacer justicia material en casos complejos. Mantener su complejidad la vuelve inoperable para sus destinatarios. La tensión no es un defecto corregible: es una propiedad estructural del intento mismo de regular jurídicamente una sociedad heterogénea.

Tercera aporía: la confianza que requiere su propio resultado

Las instituciones jurídicas funcionan, en última instancia, porque los ciudadanos confían en ellas lo suficiente para acudir a ellas. Pero la confianza no se decreta: se gana mediante evidencia acumulada de eficacia. La eficacia, a su vez, requiere que los ciudadanos sometan sus conflictos al sistema —que es lo que produce los casos sobre los cuales la confianza puede confirmarse o desmentirse.

El circuito es autorreferente y se refuerza en cualquier dirección que tome. Cuando los ciudadanos confían, acuden; al acudir, generan los casos que el sistema resuelve; las resoluciones alimentan la confianza. Cuando los ciudadanos desconfían, no acuden; al no acudir, el sistema procesa solo los casos residuales —los que llegan a pesar de la desconfianza, generalmente los más graves o los promovidos por actores con poder—; esos casos no generan evidencia de eficacia general, y la desconfianza se confirma.

México y buena parte de América Latina han operado durante décadas en la rama negativa de esta paradoja. Salir de ella desde dentro del sistema es estructuralmente difícil, porque cada salida requiere insumos —tiempo, recursos, casos, datos— que el propio circuito niega.

Tres aporías, un patrón común

Las tres paradojas comparten una propiedad lógica: ninguna admite solución desde el interior del sistema que la genera. Añadir más protecciones agrava la primera; añadir más simplicidad agrava la segunda; demandar más confianza no produce los insumos que la harían posible. Los remedios convencionales pertenecen al mismo orden lógico que el problema, y por eso no lo disuelven.

La situación recuerda a la que enfrentó la medicina antes del descubrimiento de la microbiología. Por más cirujanos, hospitales y manuales que se acumularan, ciertos problemas no podían resolverse sin un cambio epistemológico exterior al campo —en aquel caso, una nueva teoría sobre la causa de las infecciones. La intervención decisiva no vino de la medicina; vino de una disciplina vecina cuyas categorías reformularon el problema y lo hicieron tratable.

La pregunta que organiza el resto del libro es, por tanto, análoga: ¿existe, para el derecho contemporáneo, un cambio epistemológico exterior que altere las condiciones bajo las cuales sus aporías se reproducen?

PARTE II

Marco conceptual

La inteligencia artificial como tecnología epistémica

La historia humana puede leerse como una sucesión de tecnologías que no transportan objetos ni transforman materia, sino que alteran las condiciones bajo las cuales producimos, validamos y comunicamos conocimiento. La escritura fue una. La imprenta, otra. El método experimental, una tercera. El telégrafo, la grabación, la computación digital, internet: cada una redistribuyó lo que era fácil, costoso, automático o reservado a especialistas. La filosofía contemporánea de la técnica las llama *tecnologías epistémicas*.

La caracterización es precisa, no decorativa. Una tecnología epistémica no agrega capacidades a la mente individual del modo en que un microscopio agrega resolución al ojo: no es un instrumento de observación más. Modifica, en cambio, la estructura colectiva de la división del trabajo cognitivo. Cambia qué tareas conviene hacer y cuáles ya no conviene seguir haciendo a mano. Su impacto se mide no en lo que permite ver, sino en lo que vuelve obsoleto.

La inteligencia artificial generativa pertenece a este linaje. La tesis de este capítulo, anunciada en la tercera hipótesis del prólogo, es que su aparición no representa una mejora incremental de las tecnologías informáticas previas, sino una transformación cualitativa con capacidad para reorganizar significativamente la división del trabajo intelectual —y, en particular, del trabajo jurídico. Esa tesis exige precisar dos cosas: qué hace y qué no hace una IA generativa.

Qué hace una IA generativa

Un modelo de lenguaje es, en sentido estricto, un sistema estadístico que, dado un contexto textual, predice continuaciones probables. La descripción puede sonar reduccionista, pero es exacta. Lo importante, sin embargo, no es la mecánica del sistema, sino el conjunto de tareas que esa mecánica habilita cuando se entrena sobre corpus suficientemente amplios.

Una IA generativa, aplicada al derecho, es competente en al menos cinco operaciones que antes requerían trabajo humano sostenido: sintetizar grandes cuerpos textuales en tiempos breves; detectar conexiones entre normas, jurisprudencias y doctrinas que escapan a la búsqueda manual; reformular un mismo contenido entre registros distintos —técnico a llano, llano a procesal, español a otra lengua—; anticipar argumentos previsibles a partir de patrones acumulados en casos análogos; y generar borradores estructurados de documentos a partir de instrucciones precisas. Cada una de esas operaciones es, individualmente, una tarea jurídica conocida. La novedad no consiste en que la máquina las haga: consiste en que las hace en una fracción del tiempo y a un costo marginal cercano a cero.

Qué no hace una IA generativa

Igual de importante es delimitar lo que el modelo no hace. Una IA generativa no decide: produce textos. La distinción es jurídica, no técnica. Cuando un juez resuelve un caso, no está completando una secuencia probable de palabras: está atribuyendo consecuencias a hechos bajo normas, con plena conciencia de las vidas que esa atribución afecta. Esa operación —imputar, valorar, decidir— exige una agencia moral que la máquina no tiene y no puede adquirir por mera escala. Un modelo que dijera "condeno" sin comprender el peso de esa palabra estaría usando lenguaje, no ejerciéndolo.

Tampoco accede la máquina a la verdad de los hechos. Accede a patrones del lenguaje *sobre* los hechos, lo cual es algo distinto. Por eso puede generar afirmaciones plausibles que son materialmente falsas —el fenómeno conocido como alucinación— con la misma soltura con que

genera afirmaciones plausibles que son verdaderas. La verificación sigue siendo una tarea humana, y lo será mientras la máquina no tenga acceso independiente al mundo extralingüístico.

Por qué la pregunta del reemplazo está mal formulada

La pregunta dominante en el debate público —"¿reemplazará la IA al abogado?"— supone tres cosas erróneas: que el trabajo del jurista es un bloque unitario, que la pregunta admite respuesta binaria, y que lo relevante es el destino del oficio en lugar del destino de sus operaciones. La pregunta correcta es distinta: ¿qué operaciones concretas del trabajo jurídico conservarán valor cuando esta tecnología esté universalmente disponible, y cuáles dejarán de constituir oficio para volverse infraestructura?

La analogía histórica más clara es la del cálculo aritmético. La calculadora no reemplazó al matemático; suprimió el valor económico de la destreza aritmética. El matemático contemporáneo no calcula menos que su antecesor del siglo XIX: calcula otras cosas, con otras herramientas, y dedica su atención a problemas que antes ni siquiera podía formular. La IA jurídica no reemplazará al jurista; suprimirá el valor económico de ciertas destrezas —memorizar articulado, sintetizar mecánicamente, redactar a partir de plantillas— que durante décadas constituyeron el núcleo del oficio. Las competencias supervivientes serán distintas: formular problemas con precisión, distinguir verdad material de patrón textual, decidir, y diseñar las instrucciones bajo las cuales la herramienta opera.

Si la IA generativa es un amplificador epistémico y no un agente decisor, queda abierta una pregunta incómoda: ¿qué ocurre cuando, pese a esa limitación, le delegamos la decisión? ¿Qué ocurre cuando una IA técnicamente competente deja de asistir al juicio humano y comienza a sustituirlo? El capítulo siguiente examina las condiciones bajo las cuales el uso de esta herramienta puede considerarse jurídicamente aceptable.

Sesgos, opacidad y la ilusión de neutralidad

El capítulo anterior mostró qué hace y qué no hace un sistema de inteligencia artificial generativa. Queda por examinar un supuesto que la exposición dejó implícito: la aparente neutralidad de la herramienta. Ningún sistema decisional es neutral —no por defecto de diseño sino por imposibilidad lógica— y reconocerlo es la condición previa para usar la IA jurídica con rigor.

Este capítulo examina tres fuentes verificables de sesgo en los sistemas de inteligencia artificial generativa, argumenta por qué la presencia de sesgo no descalifica a la tecnología, y enumera las condiciones mínimas que un sistema de IA debe cumplir para ser jurídicamente aceptable.

Tres fuentes de sesgo

Primera fuente: los datos de entrenamiento. Un modelo de lenguaje aprende del corpus al que se lo expone. Si ese corpus contiene resoluciones judiciales en las que se ha sancionado con mayor severidad a personas de determinada condición socioeconómica, el modelo incorpora ese patrón como información válida sobre cómo "tiende a decidirse" un caso así. No discrimina deliberadamente: reproduce una discriminación que la práctica jurídica acumuló en sus precedentes. La neutralidad del corpus es una ilusión: todo corpus es testimonio de las prácticas históricas que lo produjeron.

Segunda fuente: las decisiones de diseño. Quien construye un modelo elige —explícita o implícitamente— qué considerar un resultado correcto durante el entrenamiento, qué temas evitar, qué tipos de respuesta priorizar. Esas elecciones inscriben valores en la arquitectura del sistema antes de cualquier interacción con el usuario. Un modelo ajustado para minimizar litigiosidad responderá distinto que otro ajustado para maximizar protección al consumidor, aun cuando ambos reciban el mismo prompt y partan del mismo corpus.

Tercera fuente: el contexto de despliegue. La misma herramienta usada por una defensoría pública para localizar precedentes favorables al imputado y la usada por un despacho corporativo para anticipar argumentos del litigante débil producen efectos distintos sobre la justicia agregada, aunque el modelo no haya variado. La neutralidad de una herramienta es indeterminable sin examinar quién la opera, contra quién y con qué propósito.

Tres preguntas mínimas ante toda decisión automatizada

Antes de emplear un sistema automatizado para orientar una decisión con consecuencias jurídicas, tres preguntas deben responderse afirmativamente. Primera: ¿la decisión que el sistema recomienda puede considerarse justa en términos sustantivos, más allá del cumplimiento formal de sus parámetros? Segunda: ¿existe un mecanismo accesible mediante el cual el sujeto afectado pueda impugnar la decisión antes de que sea irreversible? Tercera: ¿un tercero competente, independiente del operador del sistema, puede verificar si la decisión se ajustó a las reglas que se afirman aplicables?

Las tres preguntas separan tres dimensiones distintas —justicia sustantiva, legitimidad procedimental, auditabilidad institucional— que la práctica tiende a confundir. Un sistema puede satisfacer una y fallar en las otras dos. La aceptabilidad jurídica exige las tres.

Por qué el sesgo no descalifica al sistema

La constatación del sesgo invita, en su versión ingenua, a una conclusión apresurada: la IA, por sesgada, debe excluirse de la decisión jurídica. La conclusión es errónea por una razón sencilla: el juicio humano también es sesgado, y lo es de modos peor documentados.

Décadas de investigación en psicología cognitiva y realismo jurídico han establecido lo que la profesión jurídica se resiste a reconocer: las decisiones judiciales varían con el momento del día en que se dictan, con la condición física del juzgador, con su simpatía o antipatía implícita hacia las partes, con sesgos de anclaje, de disponibilidad, de confirmación. El juicio humano no es la línea base neutral contra la cual medir la desviación de la máquina: es otra fuente de desviación, dotada además de una ventaja estratégica que la máquina no tiene —la opacidad de su propio procedimiento.

La pregunta correcta, por tanto, no es si el sistema está sesgado, sino si su sesgo es conocible y corregible. Y aquí se produce el contraste decisivo: el sesgo algorítmico, a diferencia del humano, es susceptible de prueba sistemática. Un modelo puede ser sometido a miles de casos análogos en condiciones controladas; un juez no. Un modelo puede ser auditado por un tercero técnico; la deliberación interior de un juzgador es, por definición, inaccesible. Lo que la IA pierde en imparcialidad presunta lo gana en posibilidad de escrutinio. Esa ganancia no es marginal.

Condiciones mínimas para una IA jurídicamente aceptable

De los argumentos anteriores se desprenden, por convergencia, las condiciones que un sistema de IA debe cumplir para ser empleado legítimamente en decisiones con consecuencias jurídicas. Son seis y no admiten excepciones discrecionales.

Primera, **documentación del entrenamiento**: el usuario debe poder saber, en términos generales, qué tipo de corpus alimentó al modelo y qué exclusiones o filtros se le aplicaron.

Segunda, **explicitación de los criterios de optimización**: qué considera el sistema un resultado correcto y qué considera un error.

Tercera, **calibración verificable**: el sistema debe haber sido probado, antes de su uso, en muestras suficientemente diversas como para detectar disparidades de desempeño entre grupos relevantes.

Cuarta, **auditabilidad externa**: la posibilidad práctica de que un tercero competente verifique los puntos anteriores.

Quinta, **reversibilidad**: ninguna consecuencia jurídica producida por una recomendación algorítmica puede ser irreversible sin intervención humana superviniente.

Sexta, **derecho a explicación**: el sujeto afectado por una decisión asistida por IA debe poder conocer, en términos comprensibles, las razones operativas de la recomendación.

Las seis condiciones no son utópicas. Constituyen, de hecho, el contenido mínimo de la normativa europea sobre inteligencia artificial y de las recomendaciones de la UNESCO sobre ética de la IA. Su incumplimiento, en cualquiera de los puntos, no convierte al sistema en ilegal; lo convierte en una herramienta cuyo uso responsable es imposible.

Las seis condiciones regulan al sistema. Pero el sistema —se mostró en el Capítulo 4— no decide: produce textos a partir de instrucciones humanas. La calidad última del resultado depende, por tanto, no sólo de las propiedades del modelo sino de la calidad de la instrucción que recibe. Aquí se produce un desplazamiento conceptual decisivo: la responsabilidad técnica y ética del uso jurídico de la IA recae menos sobre el algoritmo que sobre el diseño del prompt.

¿Cómo se diseña un prompt jurídico que esté a la altura de esa responsabilidad? La respuesta es la materia de la parte siguiente del libro.

PARTE III

Método

Anatomía del prompt jurídico: las cinco capas

Si la responsabilidad técnica y ética del uso jurídico de la inteligencia artificial recae, como sostuvo el capítulo anterior, sobre el diseño de las instrucciones que la orientan, la pregunta operativa es cómo se construye una instrucción capaz de sostener esa responsabilidad. Este capítulo propone una respuesta: todo prompt jurídico bien diseñado se compone de cinco capas que, en ese orden, transforman una pregunta legal en una instrucción que la inteligencia artificial puede ejecutar con utilidad real.

Las capas son: contexto fáctico, especificación técnica, fundamentación sustantiva, diseño estratégico y estilo. No son etapas opcionales ni elementos decorativos. Cada una neutraliza una fuente típica de error del sistema, y la omisión de cualquiera de ellas degrada significativamente la calidad del resultado.

Antes de examinar cada capa, conviene una observación de orden general: el prompt jurídico no es una pregunta, es una instrucción. Una pregunta admite respuestas; una instrucción produce un resultado. El cambio de gramática es decisivo y, como se verá, no banal.

A lo largo del capítulo se utiliza un caso de referencia: un trabajador despedido sin causa justificada que pretende reclamar las prestaciones derivadas. Es deliberadamente común y accesible. El método se transfiere sin pérdida a materia civil, mercantil, familiar, penal o administrativa.

El prompt como acto estratégico

Antes de descender a las capas conviene fijar la naturaleza de la operación. Un prompt jurídico es una decisión adoptada bajo tres condiciones simultáneas: información incompleta sobre el resultado, recursos cognitivos limitados y presencia de un adversario —procesal o argumentativo— cuyas respuestas el sistema debe anticipar. Es, en sentido estricto, un acto estratégico.

Tres principios rigen ese acto. Primero, **conocimiento del terreno**: la calidad de la instrucción depende de la precisión con que se describe el escenario en que se ejecutará. Segundo, **anticipación del adversario**: la instrucción debe pedir al sistema no solo sostener la propia posición, sino modelar la contraria. Tercero, **preservación de las zonas de acuerdo**: la victoria por convergencia estabiliza el conflicto de forma que la victoria por aniquilación no logra. Los tres principios articulan silenciosamente las cinco capas que siguen.

Capa 1 — Contexto fáctico

Qué es. La narración de los hechos relevantes y la posición procesal de las partes. No es un resumen exhaustivo del caso: es la cartografía mínima que el sistema necesita para no inventar el terreno.

Qué debe incluir. Cuatro elementos. Primero, la identidad funcional de las partes (no sus nombres, salvo necesidad): qué papel cumple cada una. Segundo, los hechos jurídicamente relevantes, ordenados cronológicamente. Tercero, la pretensión concreta del usuario. Cuarto, las omisiones declaradas: lo que aún no se sabe y debe asumirse o consultarse.

Ejemplo modelo.

Trabajador con cuatro años de antigüedad continua en una empresa privada del sector servicios. Despido verbal el 12 de marzo, sin notificación escrita ni causa expresada. El trabajador acudió a su centro de trabajo durante los tres días siguientes y se le impidió el acceso. No existió liquidación. No existe registro previo de procedimientos disciplinarios. Salario mensual:

equivalente a 1.5 veces el salario mínimo vigente. Pretensión: determinar las prestaciones reclamables y la vía procesal idónea. Información que no consta y debe confirmarse: si el contrato fue por tiempo indeterminado o determinado.

Contraejemplo.

Mi cliente fue despedido injustamente y quiero demandar a la empresa, ¿qué hago?

El contraejemplo tiene cuatro defectos típicos: califica el despido como "injusto" sin probarlo, omite antigüedad y régimen contractual, fusiona consulta y pretensión sin distinguirlas, y obliga al sistema a inventar el contexto. El resultado, en el mejor caso, será genérico; en el peor, contendrá información falsa que aparenta ser específica.

Capa 2 — Especificación técnica

Qué es. La indicación precisa de la vía procesal, la competencia, el tipo de documento o resultado esperado y los requisitos formales aplicables. Lo que en la tradición jurídica se denomina el continente, distinguido del contenido.

Qué debe incluir. Tres elementos. Primero, el tipo concreto de producto solicitado (dictamen, escrito inicial, contestación, opinión, análisis comparativo). Segundo, la autoridad ante la que se planteará y los requisitos de procedibilidad aplicables. Tercero, los plazos pertinentes, si están corriendo.

Ejemplo modelo.

Producto solicitado: análisis preliminar de viabilidad procesal de una demanda laboral por despido injustificado en jurisdicción local mexicana, conforme a la Ley Federal del Trabajo vigente. Identificar requisitos de procedibilidad, plazo de prescripción aplicable y autoridad competente. No redactar aún la demanda; sólo evaluar viabilidad.

Contraejemplo.

Escríbeme la demanda completa con todo lo necesario.

El contraejemplo confunde la fase. Pedir una demanda completa antes de evaluar la viabilidad procesal equivale a construir un edificio antes de verificar el suelo. Además, omite indicar la jurisdicción, dato sin el cual el sistema no puede aplicar el régimen procesal correcto.

Capa 3 — Fundamentación sustantiva

Qué es. La indicación de las normas, principios y precedentes que el sistema debe articular en la respuesta. Aquí se decide la profundidad del análisis y su nivel de control de convencionalidad.

Qué debe incluir. Cuatro elementos. Primero, el bloque normativo aplicable: artículos específicos de la legislación pertinente. Segundo, los criterios jurisprudenciales relevantes —en el caso mexicano, tesis de la Suprema Corte y, cuando proceda, criterios de la Corte Interamericana de Derechos Humanos. Tercero, la doctrina pertinente si existe consolidada. Cuarto —y esta es la instrucción más importante de toda la capa—, la advertencia explícita de no inventar referencias: si una fuente no puede verificarse, debe señalarse en lugar de fabricarse.

Ejemplo modelo.

Articula la respuesta con base en los artículos aplicables de la Ley Federal del Trabajo vigente en materia de despido injustificado, y en el bloque constitucional sobre derecho al trabajo (artículo 123 CPEUM), así como en los estándares interamericanos en materia de protección a la estabilidad laboral. Identifica tesis jurisprudenciales relevantes sólo si puedes citarlas con número y rubro verificables; si no las conoces con certeza, indícalo expresamente y propón cómo localizarlas.

Contraejemplo.

Cita toda la jurisprudencia relevante.

El contraejemplo es la fuente más frecuente de alucinación jurídica. Sin indicación de bloque normativo, sin acotación temática y sin advertencia sobre verificabilidad, el sistema completará la solicitud con citas plausibles pero potencialmente falsas. La verosimilitud del resultado disfraza el error y lo vuelve peligroso.

Capa 4 — Diseño estratégico

Qué es. La instrucción para que el sistema no se limite a sostener la posición propia, sino que anticipe la del adversario y exponga los puntos débiles del propio razonamiento. Aquí se aplica directamente el principio de anticipación del adversario.

Qué debe incluir. Tres elementos. Primero, la identificación de las objeciones previsibles que el adversario formulará y la réplica anticipada a cada una. Segundo, la detección de los puntos débiles del propio caso, vistos por el sistema con la frialdad que el litigante no puede tener. Tercero —y este es el uso menos explotado—, la sugerencia de zonas posibles de convergencia o acuerdo, conforme al principio de preservación.

Ejemplo modelo.

Una vez articulada la pretensión, formula: (a) las tres objeciones más probables que el demandado opondrá; (b) la réplica argumentativa a cada una con sustento normativo; (c) los dos puntos más débiles de la posición del trabajador, vistos desde fuera; (d) las zonas en que un eventual convenio podría satisfacer a ambas partes sin sacrificio de los derechos irrenunciables.

Contraejemplo.

Hazlo lo más fuerte posible para ganar.

El contraejemplo confunde fuerza retórica con solidez argumentativa. Un argumento que sólo conoce su propia voz queda inerte ante la primera objeción seria. La fortaleza real proviene de haber previsto la respuesta del otro.

Capa 5 — Estilo y ejecución

Qué es. La indicación del registro, la extensión, el formato y el tono del resultado. Determina si la respuesta servirá ante un juez, un cliente, un colega especialista, un estudiante o un evaluador académico.

Qué debe incluir. Cuatro elementos. Primero, el destinatario implícito del texto (qué persona lo leerá). Segundo, la extensión aproximada deseada. Tercero, el formato estructural (apartados, numeración, encabezados). Cuarto, las restricciones de lenguaje: prohibición o aceptación de tecnicismos, latinismos, fórmulas estilizadas.

Ejemplo modelo.

Redacta el resultado en español jurídico claro, sin latinismos innecesarios. Extensión aproximada: dos cuartillas. Estructura en cuatro apartados numerados: hechos, marco jurídico, análisis, conclusiones. Destinatario: un colega no especialista en materia laboral que necesita comprender rápidamente la viabilidad del caso.

Contraejemplo.

Hazlo bonito.

El contraejemplo deja al sistema decidir sin criterio. El resultado oscilará entre la prosa florida y la sequedad burocrática, y será inutilizable en cualquiera de los dos extremos.

Las cinco capas como sistema

Las cinco capas no son una lista de verificación: son un sistema. La primera fija el terreno; la segunda, el continente; la tercera, el contenido jurídico; la cuarta, la dimensión estratégica; la quinta, la forma comunicativa. Omitir cualquiera produce un texto cuya falla puede no ser visible a primera vista, y precisamente por eso resulta peligrosa.

El método es versátil. El mismo esqueleto se aplica a un estudiante que prepara un caso para examen, a un practicante que redacta su primer escrito de demanda, a un litigante experimentado que diseña una estrategia

compleja, a un funcionario público que dictamina un asunto administrativo, o a un investigador que escribe un artículo doctrinal. Cambian los contenidos de cada capa; la arquitectura permanece.

Una observación final, de orden práctico: las cinco capas pueden combinarse en un único prompt extenso, o desplegarse en una secuencia de prompts encadenados, uno por capa. La segunda opción suele producir mejores resultados en casos complejos, porque permite verificar y corregir cada capa antes de avanzar a la siguiente. La elección entre formato unitario y formato secuencial depende del tiempo disponible y de la gravedad del asunto.

Las cinco capas describen cómo dar instrucciones al sistema. Falta examinar qué tareas específicas, dentro del trabajo jurídico, la IA ejecuta mejor que el método tradicional. La primera de esas tareas —la detección de conexiones normativas que la búsqueda manual no alcanza— ocupa el capítulo siguiente.

Detección asistida de conexiones normativas

La investigación jurídica tradicional es secuencial. El jurista recuerda el artículo que le viene a la mente, localiza la jurisprudencia que ese artículo ha generado, sigue las referencias cruzadas que la propia jurisprudencia ofrece. El procedimiento es metódico y, en condiciones ordinarias, suficiente. Tiene, sin embargo, una limitación estructural: privilegia la profundidad sobre la lateralidad. Lo que el jurista no recuerda como punto de partida, no lo encuentra como destino.

La inteligencia artificial generativa, aplicada a esta tarea, no es un sustituto del razonamiento jurídico: es un mecanismo de expansión del horizonte argumentativo. Permite ejecutar, en minutos, exploraciones laterales que la búsqueda manual no acomete por falta de tiempo o por desconocimiento previo del campo. Este capítulo describe tres formas de conexión normativa que la IA detecta con particular eficacia, ilustra cada una con un ejemplo y advierte sobre el límite ineludible de la herramienta: la verificación humana.

Tres formas de cartografía normativa

Cartografía vertical. Toda norma específica forma parte de una jerarquía. Por encima del artículo de un código local se encuentran principios constitucionales; por encima de la Constitución, los tratados internacionales que la complementan; por encima de los tratados, los criterios de los órganos jurisdiccionales internacionales que los interpretan. La cartografía vertical consiste en reconstruir, para una norma dada, el bloque de protección superior que la sostiene. Es la operación que el control de convencionalidad exige y que el jurista promedio omite por simple agotamiento.

Cartografía horizontal. Una misma situación fáctica puede ser regulada por normas pertenecientes a ramas distintas del ordenamiento. Un despido injustificado activa normas laborales, pero puede activar también normas civiles (responsabilidad por daño), administrativas (sanción a la empresa por incumplimiento de obligaciones formales), de seguridad social (afectación de derechos derivados) y de derechos humanos (estabilidad en el empleo como derecho fundamental). La cartografía horizontal consiste en identificar la red de normas concurrentes que la pretensión activa simultáneamente. Esa red, vista en conjunto, suele revelar argumentos que la lectura monodisciplinar no ofrece.

Cartografía diagonal. Casos pertenecientes a materias jurídicas distintas pueden compartir estructura argumentativa. Un caso laboral sobre discriminación por embarazo y un caso administrativo sobre denegación de servicios públicos por origen étnico no comparten norma aplicable, pero comparten la lógica del control de igualdad. La cartografía diagonal consiste en localizar precedentes estructuralmente análogos en otras materias y extraer de ellos el razonamiento transferible. Es la operación más fina y la que la IA realiza con mayor ventaja sobre el método manual.

Un ejemplo operativo

Retomando el caso del trabajador despedido del capítulo anterior, un prompt orientado a la detección de conexiones normativas podría formularse así:

Sobre el caso ya planteado (despido injustificado de trabajador con cuatro años de antigüedad, jurisdicción mexicana), identifica tres planos de conexión normativa:

(1) Vertical: bloque constitucional y convencional aplicable, con jerarquía de fuentes desde la Ley Federal del Trabajo hasta los estándares de la Corte Interamericana de Derechos Humanos en materia de estabilidad laboral.

(2) Horizontal: ramas adicionales del ordenamiento que la situación podría activar (civil, administrativa, seguridad social, derechos humanos), con identificación de los artículos pertinentes en cada una.

(3) Diagonal: dos casos estructuralmente análogos en materias distintas a la laboral, en los que se haya razonado sobre situaciones de privación abrupta de un derecho protegido por estabilidad.

Indica únicamente referencias verificables; donde no las tengas con certeza, propón cómo localizarlas.

El resultado, bien ejecutado, no entrega una respuesta cerrada: entrega un mapa. Sobre ese mapa, el jurista decide qué conexiones explotar, cuáles descartar y cuáles requieren investigación adicional. La IA no sustituye la decisión estratégica; la informa con un horizonte que la búsqueda manual no habría podido construir.

El límite ineludible: la verificación

La capacidad combinatoria del sistema produce, junto a conexiones reales, conexiones plausibles pero inexistentes. Tesis jurisprudenciales con número y rubro fabricados, artículos atribuidos a leyes equivocadas, criterios convencionales que el órgano jamás emitió: el fenómeno de la alucinación es particularmente grave en esta tarea, porque el resultado se presenta con el mismo formato y la misma autoridad aparente que las referencias verídicas.

La regla operativa es absoluta: ninguna cita jurisprudencial, ningún artículo y ningún criterio convencional que la IA produzca puede incorporarse a un documento jurídico sin verificación independiente en la fuente oficial. Esta regla no es una recomendación de prudencia: es una obligación profesional. Citar una tesis inexistente ante un tribunal expone al litigante a sanciones disciplinarias y, más grave aún, destruye la credibilidad del resto del escrito.

La verificación, por lo demás, no anula el beneficio del método. La IA reduce el problema desde "buscar todo lo que pudiera ser aplicable" a "verificar lo que ya fue identificado como probablemente aplicable". El ahorro de tiempo es sustantivo aun considerando la verificación, y la cobertura argumentativa que se obtiene es típicamente mayor que la alcanzable por búsqueda manual exclusiva.

Las tres cartografías expanden el horizonte normativo del caso. Lo que aún no resuelven es la dimensión propiamente estratégica: qué hará la contraparte cuando ese horizonte se despliegue en su contra. La IA permite, además de cartografiar normas, modelar comportamientos argumentativos previsibles del adversario procesal. Cómo se diseña un prompt para esa tarea es el objeto del capítulo siguiente.

Modelado predictivo del adversario procesal

El principio de anticipación del adversario, enunciado en el capítulo sexto como uno de los tres que rigen el prompt jurídico, establece que conocer las respuestas de la contraparte antes de que las formule equivale a haber ganado, en términos de información, antes del combate. Hasta antes de la inteligencia artificial generativa, esa anticipación dependía exclusivamente de la experiencia acumulada del litigante: cuántos casos parecidos había visto, cuántos abogados contrarios había enfrentado, cuántos jueces conocía. El recurso era valioso pero estrechamente individual.

La IA introduce una ampliación cualitativa de esa capacidad. Entrenada sobre un volumen de litigios y resoluciones que ningún jurista podría leer en una vida, puede generar, en minutos, un modelo razonable del comportamiento argumentativo previsible del adversario en un caso dado. El modelo no sustituye la experiencia humana; la amplifica. Y la pone al alcance del litigante novel, que hasta ahora estaba estructuralmente en desventaja frente al colega con veinte años de práctica.

Este capítulo describe tres operaciones de modelado predictivo que un prompt jurídico bien diseñado permite ejecutar.

Primera operación: predicción de objeciones

La operación más elemental consiste en enumerar las objeciones puntuales que el adversario formulará contra cada elemento de la pretensión propia. No se trata de imaginar todas las objeciones posibles —ejercicio agotador y poco útil— sino de identificar las más probables conforme a los patrones argumentativos típicos en la materia.

Sobre el caso ya planteado, enumera las cinco objeciones más probables que el demandado formulará contra cada uno de los siguientes elementos: (a) acreditación del vínculo laboral; (b) acreditación del despido; (c) cuantificación del salario base; (d) procedencia de las prestaciones reclamadas. Para cada objeción, indica el fundamento normativo o probatorio típico en que se apoya.

El valor de este resultado no está en su novedad —un litigante experimentado intuye las mismas objeciones— sino en su sistematicidad. La IA no olvida ninguna línea estándar, no se distrae con sus argumentos preferidos, no subestima las objeciones que parecen débiles. Cubre el campo con regularidad mecánica. Esa regularidad es, por sí sola, una ventaja estratégica.

Segunda operación: mapeo de líneas de defensa

Las objeciones aisladas no constituyen estrategia. El adversario hábil construye una narrativa coherente que articula sus objeciones bajo una tesis unificadora: el trabajador renunció voluntariamente y no fue despedido; el contrato era por obra determinada y concluyó conforme a lo pactado; la relación nunca fue laboral sino mercantil. La verdadera anticipación consiste en identificar no las objeciones, sino las narrativas dentro de las cuales esas objeciones cobran sentido.

A partir de las objeciones identificadas, reconstruye las dos o tres narrativas defensivas más coherentes que el demandado podría sostener. Para cada narrativa, indica: (a) la tesis central

que la articula; (b) los hechos que se reordenarían para sostenerla; (c) las pruebas que el demandado típicamente ofrecería; (d) la línea jurisprudencial que la respaldaría.

Este ejercicio convierte la anticipación, de un catálogo en un guion. Y permite al litigante elegir entre dos posiciones: prepararse para neutralizar cada narrativa por separado, o adelantarse y construir la pretensión propia de modo que ciertas narrativas defensivas queden, desde el inicio, fuera del campo de juego.

Tercera operación: diseño de rutas probatorias alternativas

Una vez modeladas las narrativas del adversario, la cuestión técnica decisiva es probatoria. Cada narrativa defensiva requiere ciertas pruebas; cada prueba puede ser neutralizada por contrapruebas específicas; cada contraprueba implica diligencias concretas que deben preverse desde el inicio del procedimiento, no improvisarse durante su curso.

Para cada narrativa defensiva identificada, propón: (a) las dos contrapruebas más eficaces que el actor podría preparar desde la fase inicial; (b) las diligencias preliminares necesarias para asegurar esas contrapruebas; (c) una ruta probatoria alternativa, viable si las contrapruebas principales no pudieran obtenerse o resultaran insuficientes.

El resultado es un mapa probatorio anticipado. No es un guion rígido —los hechos imprevistos siempre alteran el curso del proceso— pero sí una preparación que reduce sustantivamente la probabilidad de sorpresa estratégica.

El límite del modelo

Una advertencia metodológica cierra el capítulo. El modelado predictivo de la IA funciona sobre patrones acumulados; predice, por tanto, comportamientos típicos. Lo que no puede prever es la creatividad específica del adversario concreto: el argumento inusual, el precedente recién publicado, la estrategia personalísima del litigante experimentado que ha decidido innovar en este caso. El modelo, en suma, ofrece la línea base; el conocimiento situado del litigante humano —quién es el abogado contrario, qué juez resuelve, qué dinámica local existe— añade la inteligencia específica.

Esa complementariedad es exactamente la forma de trabajo que el libro propone: la IA ejecuta lo sistemático; el jurista añade lo singular. Ninguno de los dos componentes basta por sí solo. Su combinación, en cambio, produce un nivel de preparación procesal que ni el método tradicional ni la herramienta aislada alcanzan.

¿Puede el método volverse todavía más potente cuando el prompt deja de ser un acto único y se convierte en una arquitectura compuesta? El capítulo siguiente lo examina.

El prompt que escribe prompts, y la división del trabajo entre agentes

Los capítulos anteriores describieron el prompt como acto singular: una instrucción dirigida a un sistema para producir un resultado. Esa descripción fue útil pedagógicamente pero es incompleta. En la práctica avanzada, el prompt jurídico rara vez es un acto único; es una arquitectura compuesta. Puede consistir en un prompt que redacta otros prompts, en una cadena de instrucciones sucesivas, o en una división del trabajo entre varios agentes especializados que operan bajo una coordinación deliberada.

Este capítulo examina esas tres arquitecturas en orden de complejidad creciente. La tesis que las unifica es que la capacidad real de la inteligencia artificial jurídica no reside en la calidad de una instrucción aislada, sino en el diseño de sistemas de instrucciones que se corrigen, se complementan y se auditan entre sí.

El meta-prompting: la IA como redactora de sus propias instrucciones

La técnica más elemental de la práctica avanzada consiste en pedir al sistema, antes de ejecutar cualquier tarea, que redacte el prompt óptimo para esa tarea. La operación es contraintuitiva y por eso poco practicada: en lugar de dar la instrucción, se solicita la instrucción.

Cuándo conviene. El meta-prompting es útil cuando el jurista sabe qué necesita pero no sabe cómo formular la petición para obtenerlo con

precisión; cuando el asunto es lo bastante complejo para que la instrucción intuitiva probablemente omita variables relevantes; y cuando la calidad del resultado dependerá crucialmente de la formulación inicial, como ocurre en dictámenes, estrategias procesales o análisis de convencionalidad.

Cómo se ejecuta. El meta-prompt tiene una estructura estable. El usuario describe en lenguaje llano el objetivo, el contexto, el destinatario del resultado final y las restricciones aplicables. Solicita al sistema que redacte el prompt óptimo para alcanzar ese objetivo, y le pide que explique las decisiones de diseño del prompt propuesto.

Ejemplo operativo.

Necesito preparar un dictamen sobre viabilidad de amparo indirecto contra la omisión de una autoridad administrativa federal. El destinatario es un cliente sin formación jurídica. Debe caber en cuatro cuartillas. Redacta el prompt óptimo para obtener ese dictamen, incluyendo las cinco capas del método (contexto, técnica, sustancia, estrategia, estilo). Explica brevemente por qué diseñaste cada capa como lo hiciste.

El resultado no es el dictamen: es el prompt para producirlo, junto con la justificación de sus decisiones. El jurista revisa el prompt propuesto, lo corrige donde su conocimiento específico del caso lo exige, y sólo entonces lo ejecuta. La operación añade un paso; pero ese paso, en asuntos serios, reduce sustantivamente el número de iteraciones posteriores.

Riesgo específico. El meta-prompt tiende a producir instrucciones formalmente impecables pero genéricas. La corrección humana no es opcional: es el momento en que el conocimiento situado del jurista se incorpora al diseño. Un meta-prompt aceptado sin revisión reproduce, con más elegancia, las patologías del prompt mal diseñado.

Las cadenas de prompts: la descomposición del trabajo en secuencia

Un prompt único que intenta abarcar viabilidad procesal, fundamentación sustantiva, anticipación estratégica y redacción final produce, invariablemente, un resultado plano. La cadena de prompts descompone el trabajo en fases discretas, cada una ejecutada por una instrucción independiente cuyo resultado alimenta a la siguiente.

Cuándo conviene. Siempre que el producto final tenga más de un componente evaluable por separado. Un escrito de amparo, por ejemplo, contiene diagnóstico de procedencia, identificación de actos reclamados, formulación de conceptos de violación, articulación probatoria y redacción formal. Cada componente requiere criterios distintos y admite corrección independiente.

Cómo se ejecuta. La cadena típica opera en cinco eslabones sucesivos: diagnóstico del asunto; exploración normativa y jurisprudencial; formulación de tesis y anticipación adversarial; redacción del borrador; revisión crítica del borrador desde una perspectiva contraria. Cada eslabón recibe como insumo el resultado del anterior, y cada uno se corrige antes de avanzar.

Ejemplo operativo aplicado al caso del despido injustificado.

Eslabón 1 (diagnóstico): *Sobre los hechos ya suministrados, evalúa la viabilidad procesal, plazos de prescripción y competencia. No avances al fondo hasta que confirme.*

Eslabón 2 (exploración): *Confirmada la viabilidad, cartografía el bloque normativo aplicable en los tres planos vertical, horizontal y diagonal. Entrega mapa, no borrador.*

Eslabón 3 (tesis y anticipación): *Sobre el mapa validado, formula tres tesis argumentativas alternativas y anticipa la defensa más probable contra cada una.*

Eslabón 4 (redacción): *Selecciono la tesis segunda. Redacta*

borrador de escrito conforme a las cinco capas.

Eslabón 5 (revisión adversarial): *Analiza el borrador desde la perspectiva del abogado del demandado. Identifica sus tres ataques más probables y sugiere refuerzos.*

Cada eslabón consume minutos; el conjunto, menos de una hora. La calidad del resultado final es órdenes de magnitud superior a la de un prompt unitario, y el jurista mantiene control en cada bifurcación.

Regla operativa. No avanzar al siguiente eslabón si el actual no satisface. Corregir en el eslabón donde el defecto aparece; no arrastrar errores esperando que la siguiente instrucción los enmiende. Esta regla, aparentemente obvia, es la que separa el uso disciplinado del uso ansioso de la herramienta.

Los agentes especializados: la división del trabajo por función

La arquitectura más sofisticada consiste en asignar a distintas instancias del sistema roles diferenciados, con instrucciones específicas para cada uno, y coordinar sus resultados. En lugar de una sola conversación que atraviesa todas las fases, se opera con varios agentes que ejecutan funciones distintas y cuyos productos se integran.

Cuándo conviene. En asuntos complejos donde distintos ángulos del análisis requieren competencias contradictorias entre sí. Redactar y criticar son operaciones difícilmente compatibles en la misma instancia; buscar precedentes favorables y precedentes adversos exigen orientaciones opuestas; sostener la posición propia y modelar la del adversario son roles cognitivos que se estorban recíprocamente cuando se pide a una sola instancia que los ejecute simultáneamente.

Cuatro agentes básicos para el trabajo jurídico.

Agente investigador. Su instrucción única es localizar, cartografiar y verificar fuentes normativas, jurisprudenciales y doctrinales aplicables. No argumenta, no redacta escritos, no sugiere estrategia. Produce mapas, no piezas. Su virtud es la exhaustividad; su regla, la

verificabilidad absoluta.

Agente redactor. Recibe el mapa del investigador y una tesis definida por el jurista. Su instrucción es traducir esa tesis en texto conforme a las cinco capas del método. No investiga por su cuenta ni cuestiona la tesis; ejecuta la redacción con la fundamentación que se le suministra.

Agente adversario. Recibe el borrador del redactor. Su instrucción es una sola: sostener con máxima energía la posición contraria. No busca equilibrio ni concede; opera como el mejor litigante del demandado en el caso concreto. Su producto es un catálogo estructurado de ataques al borrador propio.

Agente revisor. Recibe el borrador y los ataques del agente adversario. Su instrucción es evaluar cuáles ataques son procedentes, cuáles pueden neutralizarse con ajustes al borrador y cuáles exigen replanteo de fondo. Es el filtro previo a la revisión humana final.

Ejemplo de flujo integrado.

*El jurista introduce los hechos al **agente investigador** con la instrucción de cartografiar bloque normativo, jurisprudencia SCJN vinculante en los últimos cinco años y criterios interamericanos relevantes, con verificabilidad explícita. Revisa y valida el mapa.*

*Traslada el mapa al **agente redactor** junto con la tesis seleccionada y las especificaciones de forma. Obtiene borrador.*

*Somete el borrador al **agente adversario**, instruido para operar como abogado del demandado con veinte años de experiencia. Obtiene catálogo de ataques.*

*Pasa borrador y catálogo al **agente revisor**. Obtiene diagnóstico priorizado: qué corregir, qué reforzar, qué replantear.*

El jurista integra las correcciones, verifica cada cita en fuente oficial, y firma.

El flujo entero consume, en un asunto de complejidad media, dos a tres horas. La versión tradicional del mismo trabajo, sin IA, tomaría entre dos y cinco días. La versión con IA pero sin división por agentes tomaría menos que la tradicional, pero produciría un resultado menos robusto que el flujo especializado.

Los tres riesgos específicos de la arquitectura compuesta

La sofisticación del método introduce riesgos propios que el prompt aislado no presenta. Enumerarlos es condición del uso responsable.

Pérdida de trazabilidad. Cuando cuatro agentes intervienen en un producto, la reconstrucción del origen de cada elemento del resultado se dificulta. El jurista debe mantener un registro mínimo —qué agente produjo qué componente, con qué instrucción— para preservar la trazabilidad exigida por el Principio 1 del método.

Propagación de errores. Un error del agente investigador —una cita fabricada, una jurisprudencia atribuida erróneamente— se traslada al redactor, sobrevive a la revisión adversarial y contamina el resultado final. La regla profesional es implacable: la verificación humana de fuentes ocurre inmediatamente después de la fase de investigación, nunca al final del flujo. Un error detectado tarde cuesta multiplicado por el número de fases que atravesó.

Falsa sensación de rigor. La complejidad procedimental del método puede inducir en el operador una confianza infundada: "cuatro agentes revisaron, entonces está bien". La confianza no procede del número de agentes sino de la calidad de la instrucción a cada uno y de la revisión humana final. Un flujo mal diseñado puede producir, con más pasos, un resultado peor que un prompt único cuidadoso.

Cinco reglas prácticas para el trabajo con arquitecturas compuestas

Al margen de la sofisticación del método, cinco reglas operan como piso mínimo profesional.

Primera: **cada agente debe tener una función única y explícita.** La ambigüedad de rol produce resultados híbridos e inutilizables.

Segunda: **entre agente y agente, hay siempre una validación humana.** La cadena no se ejecuta automáticamente; se ejecuta con detenciones deliberadas del operador.

Tercera: **la verificación de fuentes ocurre en el momento en que las fuentes se producen,** no al final. Es la única defensa efectiva contra la propagación de alucinaciones.

Cuarta: **el registro del flujo es parte del producto.** Al finalizar, el jurista debe poder reconstruir en un párrafo qué hizo cada agente, con qué instrucción y qué se conservó del resultado. Ese registro es lo que hace auditable la práctica.

Quinta: **la firma final es siempre humana y personal.** Ninguna arquitectura compuesta, por sofisticada que sea, transfiere la responsabilidad última del jurista al sistema. El Principio 1 del método —la instrumentalidad— gobierna también, y sobre todo, aquí.

Meta-prompting, cadenas y agentes son formas de aumentar la potencia del método. La misma sofisticación, sin embargo, multiplica los modos en que el método puede fallar. ¿Cuáles son los errores recurrentes que la práctica documenta con mayor frecuencia, y cómo se identifican antes de que produzcan daño? Ese es el objeto del capítulo siguiente.

Patologías frecuentes del prompt mal diseñado

Los capítulos previos describieron el método en su forma básica y en sus arquitecturas compuestas. El presente capítulo invierte la perspectiva: cataloga los errores recurrentes que la práctica documenta con mayor frecuencia en el uso jurídico de la inteligencia artificial. Cada patología se presenta en tres registros: la forma típica en que aparece, el riesgo que introduce y la regla correctiva que la neutraliza.

El listado no es exhaustivo, pero abarca los siete defectos más documentados en la práctica forense reciente. Funciona como autodiagnóstico antes que como manual: el lector puede revisar sus propios prompts contra este catálogo y verificar, en pocos minutos, cuáles de las patologías le son habituales.

Patología 1 — El prompt-pregunta

Forma típica. "¿Qué procede en un caso de despido injustificado?"

Riesgo. El sistema responde con un panorama general que no se ajusta al caso concreto. La respuesta es enciclopédica, no operativa.

Corrección. Reformular toda consulta como instrucción dirigida a un producto específico. No preguntar qué procede: ordenar qué se necesita.

Patología 2 — El terreno vacío

Forma típica. "Mi cliente fue despedido y necesito ayuda."

Riesgo. Al carecer de hechos concretos, el sistema reconstruye el contexto a partir de patrones promedio. La respuesta es plausible pero ajena al caso real, y su corrección posterior cuesta más tiempo que haber suministrado el contexto desde el inicio.

Corrección. Aplicar la primera capa del Capítulo 6. Antigüedad, fecha, hechos, salario, omisiones declaradas. Mínimo verificable, máximo controlable.

Patología 3 — La sobrecarga

Forma típica. "Redáctame la demanda completa con todos los argumentos posibles."

Riesgo. El sistema entrega un producto extenso, formalmente correcto, pero estratégicamente plano. La sobrecarga impide afinar cada componente; la respuesta es la suma de promedios.

Corrección. Descomponer el trabajo en fases discretas y prompts encadenados: viabilidad, fundamentación, anticipación, redacción, revisión. La calidad final es función de la calidad de cada fase.

Patología 4 — La cita fantasma

Forma típica. "Cita toda la jurisprudencia relevante en la materia."

Riesgo. El sistema completa la solicitud generando tesis con número, rubro y sala plausibles, pero potencialmente inexistentes. La verosimilitud disfraza el error.

Corrección. Indicar explícitamente que sólo se citen referencias verificables; donde no haya certeza, ordenar al sistema que lo declare y proponga cómo localizar la fuente.

Patología 5 — El prompt unilateral

Forma típica. "Construye los argumentos más sólidos para sostener mi posición."

Riesgo. El resultado es retóricamente fuerte pero estratégicamente ciego. Ignora la respuesta previsible del adversario y deja al litigante inerme ante la primera objeción seria.

Corrección. Incorporar en el prompt la instrucción de modelar la posición contraria, identificar los puntos débiles propios y mapear zonas posibles de convergencia. La solidez argumentativa proviene de haber previsto el ataque, no de haberlo ignorado.

Patología 6 — La delegación de la decisión

Forma típica. "Dime qué debo hacer en este caso."

Riesgo. El sistema produce una recomendación con apariencia de decisión, pero sin acceso a la información extralingüística que la decisión requiere: las condiciones reales del cliente, las dinámicas locales del tribunal, las consecuencias humanas del curso de acción. La delegación de la decisión transfiere a la máquina una responsabilidad que estructuralmente no puede asumir.

Corrección. Pedir información, análisis comparativo, evaluación de escenarios. No pedir la decisión. La decisión, en derecho como en medicina, permanece en el agente humano responsable.

Patología 7 — El abandono iterativo

Forma típica. Aceptar el primer resultado del sistema sin refinarlo ni someterlo a prompts correctivos sucesivos.

Riesgo. El primer resultado suele ser un borrador estructurado, pero genérico. Su utilidad real depende de las iteraciones siguientes —correcciones, profundizaciones, depuraciones— que la pereza profesional o la presión de tiempo invitan a omitir.

Corrección. Tratar el primer resultado como boceto, no como producto final. La regla operativa: ningún prompt jurídico serio produce su mejor resultado en la primera respuesta. La calidad emerge en la iteración disciplinada.

Las siete patologías como sistema

Las siete patologías comparten una propiedad estructural: cada una desplaza, sobre el sistema, una responsabilidad que correspondía al jurista. El prompt-pregunta delega la formulación; el terreno vacío delega el contexto; la sobrecarga delega la planificación; la cita fantasma delega la verificación; el prompt unilateral delega la dialéctica; la delegación de la decisión delega el juicio; el abandono iterativo delega la depuración. La inteligencia artificial responde, en cada caso, lo que puede responder con los recursos que se le suministran. Las fallas no son del modelo: son de quien lo opera.

Ese diagnóstico tiene una consecuencia inmediata. Si las patologías nacen del desplazamiento de responsabilidad, su prevención exige una ética operativa explícita —no un conjunto de buenas intenciones generales, sino reglas concretas que orienten el uso responsable de la herramienta.

¿Qué principios deben gobernar el uso responsable de la inteligencia artificial en el ejercicio jurídico, cuando la herramienta es ya parte de la práctica cotidiana? La respuesta articula la última parte del libro.

PARTE IV

Implicaciones

Diez principios para una práctica responsable

Los capítulos anteriores describieron el método. El presente formula sus límites. Una herramienta poderosa, ejercida sin reglas, produce daños proporcionales a su poder. Las diez reglas que siguen no son un código deontológico —los códigos son tarea de los colegios profesionales y de los legisladores— sino un decálogo razonado: cada principio se enuncia como regla operativa y se justifica como derivación de las tesis filosóficas examinadas a lo largo del libro.

El orden no es jerárquico. Los diez principios funcionan como sistema, no como secuencia.

Principio 1 — Instrumentalidad y trazabilidad

Regla. La inteligencia artificial es herramienta, no agente. Toda decisión jurídica con consecuencias para terceros permanece bajo responsabilidad humana identificable. Cuando el método emplea arquitecturas compuestas —cadenas de prompts o agentes especializados—, el operador debe conservar registro mínimo del flujo: qué agente produjo qué componente, con qué instrucción.

Fundamento. Como se mostró en el Capítulo 4, los modelos generativos producen textos, no decisiones. La aparente fluidez con que un sistema formula recomendaciones puede inducir, en el operador, una transferencia silenciosa de responsabilidad. El principio neutraliza esa tentación: la firma que respalda el documento, el patrocinio que sostiene el escrito, la asesoría que orienta al cliente, son siempre humanos. La

trazabilidad de la responsabilidad es condición de la práctica jurídica civilizada, y se dificulta —sin desaparecer— cuando varios agentes intervienen en un mismo producto.

Principio 2 — Verificación

Regla. Ninguna referencia normativa, jurisprudencial o doctrinal producida por la IA puede invocarse en un documento jurídico sin verificación independiente en la fuente oficial.

Fundamento. La alucinación —fenómeno descrito en los Capítulos 5 y 7— produce referencias plausibles pero potencialmente inexistentes. La verosimilitud del resultado disfraza el error y lo vuelve peligroso, particularmente en contextos donde el destinatario confía en la apariencia formal de la cita. La verificación no es prudencia: es obligación.

Principio 3 — Transparencia

Regla. El uso significativo de IA en la producción de un documento jurídico debe poder declararse cuando sea relevante para el destinatario.

Fundamento. La auditabilidad —discutida en el Capítulo 5— es propiedad institucional, no técnica. Una sociedad que normaliza el uso oculto de inteligencia artificial en decisiones jurídicas pierde la posibilidad de evaluar, corregir y, en su caso, restringir la herramienta. La transparencia no exige declarar cada uso trivial; exige no negar el uso cuando alguien razonablemente lo pregunte.

Principio 4 — Simetría dialéctica

Regla. Todo análisis asistido por IA debe incluir la posición contraria con el mismo rigor que la posición propia.

Fundamento. El principio de anticipación del adversario, enunciado en el Capítulo 6 y desarrollado en el Capítulo 8, establece que la fortaleza argumentativa proviene de la anticipación, no de la unilateralidad. El operador que utiliza la IA exclusivamente para reforzar su propia

posición la convierte en cámara de eco, y desperdicia precisamente lo que el sistema hace mejor: explorar simultáneamente varias hipótesis con regularidad mecánica.

Principio 5 — Proporcionalidad

Regla. La complejidad del prompt debe corresponder a la complejidad del asunto. No usar la IA donde no agrega valor; no omitirla donde sí lo agrega.

Fundamento. El principio de economía cognitiva se aplica también al método mismo. Hay consultas que se responden mejor en cinco minutos con una memoria entrenada que con un prompt elaborado; hay otras donde la búsqueda manual perpetuaría una ignorancia evitable. El criterio profesional consiste en discriminar cuándo conviene cada herramienta.

Principio 6 — Protección del vulnerable

Regla. La IA jurídica debe utilizarse, prioritariamente, para reducir la asimetría de acceso a la justicia, no para profundizarla.

Fundamento. Los Capítulos 2 y 3 establecieron que el sistema jurídico contemporáneo presenta una asimetría estructural de acceso. La inteligencia artificial puede operar en dos direcciones contrarias: como recurso adicional del litigante poderoso o como herramienta de equilibrio para el litigante débil. La elección no es técnica: es política y profesional. Este libro toma posición explícita en la segunda dirección.

Principio 7 — Confidencialidad

Regla. La información de clientes, expedientes y comunicaciones privilegiadas no debe trasladarse a sistemas de IA cuyas políticas de retención y entrenamiento sean desconocidas o insuficientes para garantizar el secreto profesional.

Fundamento. El secreto profesional no es una formalidad: es un pilar funcional del sistema jurídico. Su violación, aun por descuido técnico, daña al cliente concreto y erosiona la confianza estructural que sostiene la profesión. La regla obliga al jurista a conocer, antes de usar una herramienta, las condiciones contractuales y técnicas bajo las cuales sus datos serán procesados.

Principio 8 — Iteración

Regla. Ningún producto jurídico asistido por IA se entrega en su primera versión. La depuración mediante iteraciones sucesivas forma parte constitutiva del método.

Fundamento. La séptima patología del Capítulo 10 documenta el riesgo del abandono iterativo: aceptar como definitivo el primer borrador del sistema. La calidad jurídica no emerge en la primera respuesta; emerge en la corrección sostenida. La iteración no es opcional: es parte del producto.

Principio 9 — Límite reconocido

Regla. El operador debe conocer, y poder enunciar, las tareas que la herramienta no puede realizar. Atribuirle capacidades inexistentes es tan grave como rehusar las que sí posee.

Fundamento. El Capítulo 4 distinguió con cuidado lo que la IA hace de lo que no hace. El operador profesional asume ambas listas. La sobreestimación produce resultados peligrosos; la subestimación, atraso competitivo. El equilibrio no se alcanza por intuición: se sostiene en formación técnica explícita.

Principio 10 — Formación continua

Regla. La competencia del jurista en el uso de IA no se adquiere una vez. La evolución constante de la herramienta exige una formación correlativa, sostenida y verificable.

Fundamento. Una tecnología en transformación rápida no permite competencias congeladas. El jurista que aprendió el método de hoy y deja de actualizarse será, en pocos años, técnicamente obsoleto. El principio no demanda erudición permanente: demanda la disposición a revisar regularmente lo que se sabe y a desaprender lo que dejó de servir.

Los diez principios como sistema

Los diez principios pueden parecer, leídos por separado, recomendaciones de sentido común. Vistos en conjunto, sin embargo, configuran una ética operativa específica: la del jurista que utiliza la inteligencia artificial sin abdicar de la responsabilidad que le corresponde y sin renunciar al equilibrio entre la herramienta y el sujeto que la ejerce. Esa ética no resuelve, ni pretende resolver, los grandes debates regulatorios sobre la IA. Aspira a algo más modesto y más urgente: ofrecer, a quien hoy practica el derecho, una guía verificable sobre cómo no causar daño con la herramienta en la mano.

El epílogo recoge lo que estos once capítulos no resolvieron y propone, sin ánimo de respuesta, las preguntas que habrá que seguir formulando.

Epílogo

Preguntas abiertas

El prólogo prometió un ensayo de hipótesis, no de doctrinas. Llegado el final, conviene cumplir esa promesa en sentido inverso: nombrar las preguntas que el libro no resolvió y que el lector debe asumir como suyas si pretende continuar la discusión que estas páginas apenas abrieron.

No son lagunas de la exposición. Son problemas reales sobre los cuales el autor carece de respuesta satisfactoria, y respecto a los cuales una pretensión de respuesta habría sido deshonestas. La práctica jurídica responsable, frente a una tecnología en transformación, exige reconocer con claridad lo que aún no se sabe.

Primera pregunta abierta: la responsabilidad civil del prompt

Cuando un documento jurídico asistido por IA produce un daño —una demanda construida sobre jurisprudencia inexistente, un dictamen que omite una norma decisiva, una asesoría que conduce al cliente a una vía improcedente—, ¿quién responde? ¿El jurista que firmó? ¿El desarrollador del modelo? ¿La plataforma que lo distribuyó? ¿El proveedor institucional que lo integró en sus sistemas?

El Capítulo 11 estableció la responsabilidad última del operador humano. Esa respuesta sirve para casos ordinarios. No basta, sin embargo, cuando el daño deriva de comportamientos del sistema que el operador profesional no podía razonablemente prever. La doctrina civil de la responsabilidad por hecho ajeno y por riesgo creado ofrece herramientas conceptuales, pero su aplicación a herramientas autónomas adaptativas exige desarrollos que la academia jurídica apenas comienza.

Segunda pregunta abierta: la auditoría algorítmica obligatoria

El Capítulo 5 sostuvo que la auditabilidad no es propiedad del código sino del entorno institucional que lo rodea. La pregunta inmediata sigue sin respuesta: ¿quién audita, con qué autoridad, bajo qué procedimiento? El derecho comparado ofrece modelos incipientes —el Reglamento europeo sobre Inteligencia Artificial, las recomendaciones de la UNESCO, propuestas nacionales en distintos países latinoamericanos—, pero ninguno consolidado. La discusión sobre si la auditoría debe ser obligatoria, voluntaria, estatal o privada recae sobre los próximos años, y debe involucrar a la profesión jurídica como interlocutor activo, no como destinatario pasivo.

Tercera pregunta abierta: la concentración técnica

Los modelos generativos más capaces son producidos por un número reducido de empresas, mayoritariamente situadas en jurisdicciones extranjeras al lector hispanohablante. Esa concentración plantea problemas que el derecho doméstico no ha terminado de formular: la dependencia técnica de la administración de justicia, la soberanía sobre los datos jurídicos nacionales, la posibilidad de que las decisiones de diseño de un actor privado extranjero condicionen prácticas argumentativas locales. Estas cuestiones no son consecuencias futuras posibles: son tensiones presentes que el libro nombró sin pretender resolver.

Cuarta pregunta abierta: la pérdida de competencias formativas

Si la inteligencia artificial absorbe progresivamente las tareas rutinarias mediante las cuales el jurista joven aprendía el oficio —investigación jurisprudencial, redacción de borradores, síntesis de doctrina—, ¿qué ruta formativa sustituirá a esa experiencia perdida? La adquisición de juicio profesional no admite atajos: se construye en horas de trabajo aparentemente repetitivo que la herramienta hoy permite saltar. Cómo formar al jurista del futuro sin condenarlo a una competencia superficial es un problema pedagógico que la enseñanza del derecho aún no está discutiendo con la urgencia que merece.

Quinta pregunta abierta: el tiempo regulatorio

La inteligencia artificial generativa evoluciona en ciclos de meses. La legislación, en ciclos de años. Esa asimetría temporal produce un problema estructural: cualquier norma promulgada será, en algún sentido decisivo, obsoleta al momento de su entrada en vigor. ¿Cómo regular un objeto que se transforma más rápido que el procedimiento legislativo? La respuesta no es renunciar a la regulación —ese es el reflejo perezoso— ni promulgar marcos rígidos que la realidad rebasará en seis meses. Se requieren formas regulatorias adaptativas que el constitucionalismo contemporáneo apenas comienza a imaginar. Mientras esas formas se construyen, la ética operativa del jurista —los diez principios del Capítulo 11— ocupa el vacío.

Una nota final

Estas cinco preguntas no agotan las que el libro deja abiertas; son las más visibles. Hay otras menos articuladas pero no menos importantes: la gobernanza de sistemas multiagente en el ejercicio profesional, el papel de la IA en sistemas jurídicos plurales y multiculturales, la transformación del rol del juez, el desplazamiento de la carga argumental entre las partes.

El libro se cierra, por tanto, en el mismo registro con que se abrió: como invitación a una conversación pendiente. La inteligencia artificial generativa es ya parte de la práctica jurídica cotidiana. Lo que aún está por decidirse es bajo qué condiciones, con qué reglas y con qué propósito.

Esa decisión no es del autor ni del lector individual. Es de la comunidad jurídica entera. Lo único que un libro de bolsillo puede aspirar a hacer es poner sobre la mesa, con la mayor claridad posible, los términos en que la conversación debe ocurrir.

— A. S. A.
Cuernavaca, Morelos, 2026

Fuentes consultadas

El siguiente apartado ofrece una guía mínima de las fuentes en que se apoyan las referencias empíricas, normativas y conceptuales del libro. No pretende exhaustividad bibliográfica: reúne las obras y documentos que el lector interesado en profundizar puede consultar directamente. Se organiza por capítulo, en el orden en que las referencias aparecen en el texto.

Sobre el Capítulo 1 (El derecho como tecnología social)

Aristóteles, *Ética a Nicómaco*, libro VI, sobre la distinción entre *episteme*, *techne* y *phronesis*.

Sobre la crítica a la pretensión científica del derecho: Karl Larenz, *Metodología de la ciencia del derecho*; Robert Alexy, *Teoría de la argumentación jurídica*; Manuel Atienza, *Curso de argumentación jurídica*.

Sobre la noción de tecnología social: Herbert Simon, *The Sciences of the Artificial*; Bruno Latour, *Reassembling the Social*.

Sobre el Capítulo 2 (La obsolescencia funcional del sistema jurídico)

World Justice Project, *Rule of Law Index*, ediciones anuales, disponible en worldjusticeproject.org.

Instituto Nacional de Estadística y Geografía (INEGI), Censo Nacional de Impartición de Justicia Estatal y Censo Nacional de Impartición de Justicia Federal, ediciones anuales.

Banco Mundial, informes sobre eficiencia judicial y acceso a la justicia en América Latina.

Movimiento internacional por el lenguaje claro aplicado al derecho: Red de Lenguaje Claro (Clarity International); en México, la iniciativa de la Suprema Corte de Justicia de la Nación sobre lenguaje ciudadano en resoluciones.

Sobre el Capítulo 3 (Tres aporías contemporáneas)

Sobre la paradoja del acceso: Mauro Cappelletti y Bryant Garth, *El acceso a la justicia. Movimiento mundial para la efectividad de los derechos*.

Sobre la relación entre confianza institucional y funcionamiento del sistema jurídico: Niklas Luhmann, *El derecho de la sociedad*; Tom Tyler, *Why People Obey the Law*.

Sobre el Capítulo 4 (La inteligencia artificial como tecnología epistémica)

Sobre tecnologías epistémicas: Luciano Floridi, *The Fourth Revolution. How the Infosphere is Reshaping Human Reality*; Don Ihde, *Technology and the Lifeworld*.

Sobre las capacidades y límites técnicos de los modelos de lenguaje: Emily M. Bender, Timnit Gebru y otros, "On the Dangers of Stochastic Parrots" (2021); documentación técnica pública de los principales proveedores de modelos generativos.

Sobre el Capítulo 5 (Sesgos, opacidad y la ilusión de neutralidad)

Sobre sesgos algorítmicos: Cathy O'Neil, *Weapons of Math Destruction*; Safiya Umoja Noble, *Algorithms of Oppression*; Virginia Eubanks, *Automating Inequality*.

Sobre sesgos judiciales humanos: Daniel Kahneman, *Thinking, Fast and Slow*; Chris Guthrie, Jeffrey Rachlinski y Andrew Wistrich, "Blinking on the Bench: How Judges Decide Cases".

Marco normativo: Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial (AI Act); UNESCO, *Recomendación sobre la Ética de la Inteligencia Artificial* (2021).

Sobre los Capítulos 6 a 10 (Método)

Sobre ingeniería de prompts en contextos profesionales: literatura técnica en evolución constante; se recomienda consultar publicaciones actualizadas de los principales laboratorios de investigación en IA. En español, están emergiendo guías profesionales específicas para el ámbito jurídico que conviene revisar al momento

de leer este libro.

Sobre el fenómeno de la alucinación en modelos de lenguaje y su relevancia forense: reportes de sanciones judiciales por citación de jurisprudencia inexistente producida por IA en jurisdicciones estadounidenses y europeas, ampliamente documentados en prensa especializada desde 2023.

Sobre arquitecturas de agentes en sistemas de inteligencia artificial: literatura técnica reciente sobre *agentic AI* y sistemas multiagente en aplicaciones profesionales.

Sobre el Capítulo 11 (Diez principios para una práctica responsable)

Marcos éticos internacionales: UNESCO, *Recomendación sobre la Ética de la Inteligencia Artificial* (2021); OCDE, *Principios sobre la Inteligencia Artificial* (2019, actualizados 2024); Consejo de Europa, *Convenio Marco sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho* (2024).

Sobre deontología profesional aplicada al uso de IA en el ejercicio jurídico: opiniones del American Bar Association y de barras profesionales europeas y latinoamericanas emitidas a partir de 2023; en México, los pronunciamientos de la Barra Mexicana Colegio de Abogados y de otros colegios profesionales sobre el uso responsable de herramientas tecnológicas.

Sobre el Epílogo (Preguntas abiertas)

Sobre responsabilidad civil por daño causado por sistemas de IA: propuesta de Directiva europea sobre responsabilidad en materia de IA (2022) y sus versiones sucesivas.

Sobre auditoría algorítmica: literatura académica en desarrollo; se destacan los trabajos del AI Now Institute y del Ada Lovelace Institute.

Nota metodológica

Todas las fuentes citadas fueron consultadas y verificadas por el autor. Las referencias empíricas cuantitativas del libro (tiempos procesales, tasas de acceso a la justicia, comprensión ciudadana de resoluciones judiciales) provienen de las fuentes institucionales listadas y responden al orden de magnitud documentado por dichas fuentes al momento de la redacción. El lector que requiera cifras específicas actualizadas debe consultar las ediciones más recientes de los reportes citados.

Sobre el autor

Andrés Saavedra Avendaño es doctor en Derecho y Globalización por la Universidad Autónoma del Estado de Morelos. Litigante en ejercicio desde hace más de una década, con práctica en amparo federal, litigio constitucional, derecho de familia, derecho penal, derecho laboral, materia agraria y derecho condominal.

Es fundador y director de Artistas Legales A.C., organización sin fines de lucro dedicada a la aplicación de tecnología al acceso a la justicia y a la defensa de grupos en situación de vulnerabilidad.

Su investigación se ha centrado en la intersección entre derecho y tecnología, particularmente en el uso de sistemas automatizados y de inteligencia artificial como herramientas de democratización del acceso a la justicia.

Ejerce en Cuernavaca, Morelos, México.

Ingeniería jurídica.
Una introducción crítica a la justicia asistida
por inteligencia artificial
se terminó de editar en Cuernavaca, Morelos,
en el año 2026.