



Clase IAG: El Despacho Jurídico con IA

Diplomatura IAG | Abog. María Elisa Farías

Matriz de Esfuerzos: Ecosistema de Modelos & Routing (2026)

Routing Inteligente

En el ejercicio diario del derecho, usar el modelo más caro y lento para una tarea simple es una pérdida de tiempo; pero usar un modelo rápido para fundar un recurso extraordinario es una negligencia. Esta matriz clasifica los motores del ecosistema actual ([Claude](https://claude.ai) <https://claude.ai>), [OpenAI / ChatGPT](https://chatgpt.com) <https://chatgpt.com>) y [Google](https://gemini.google.com) <https://gemini.google.com>) según el nivel de esfuerzo de cómputo que exige cada escrito.

Selector Oficial de [Claude](https://claude.ai) <https://claude.ai> (Serie 5.5 & Fable 5.1)

Control de Esfuerzo: Bajo | Medio | Alto

Haiku 5.5

La más rápida para respuestas inmediatas

Cédulas, resúmenes de 1 foja, fechas y correos.

Sonnet 5.5

Más eficiente para tareas más sencillas

Mero trámite, contratos estándar e intimaciones.

Opus 5.5



Para trabajo complejo y tareas cotidianas

Demandas, contestaciones, agravios y litigio diario.

Fable 5.1

Créditos

Para tus desafíos más difíciles

Recurso extraordinario CSJN y arbitrajes complejos.

● Selector de **ChatGPT** (<https://chatgpt.com>) (modelo GPT-6)

Niveles de pensamiento: Instant · Medium · High · Extra High · Pro

En el menú elegís GPT-6 y un nivel de pensamiento; el menú no muestra los nombres internos. En los planes pagos, GPT-6 Sol responde de Instant a Extra High, y el nivel Pro usa GPT-6 Pro, que funciona con GPT-6 Astra. Los planes Free y Go usan Luna. En Plus no están Extra High ni Pro.

Instant

Respuesta instantánea para despacho operativo

Cédulas, extracción de fechas y datos, minutos y correos.

Medium / High

Equilibrio y fluidez para el trabajo cotidiano

Contratos tipo, cartas documento y escritos de trámite.

Extra High



Razonamiento extendido para litigación de fondo

Demandas, contestaciones, agravios y cruce de prueba.

Pro (Astra)

Máximo Cómputo

Razonamiento de máxima exigencia analítica

Auditoría de causas complejas y razonamiento de máxima exigencia.

Astra for Law es un producto aparte, con acceso restringido (Trusted Access) para estudios seleccionados de EE. UU.; no aparece en el menú de ChatGPT.

◆ Selector de **Gemini** (<https://gemini.google.com>) (modelos por nombre + nivel de pensamiento)

Pensamiento: Estándar · Extendido · Deep Think

Flash-Lite (3.5)

La opción más rápida y económica para consultas directas

Cédulas, resúmenes breves, extracción de datos y correos.

Flash (3.8)

Velocidad extrema y ventana de 1M de tokens

Ingesta de expedientes largos, cédulas, cronologías y correos.

Pro (3.1, versión preliminar) ✓

Precisión analítica para redacción jurídica

Contratos complejos, apelaciones y dictámenes procesales.

Argon (4)

gemini-4-argon

Modelo frontera con 1.048.576 tokens unificados

Auditoría de sentencias definitivas y recursos extraordinarios CSJN.

El nivel de pensamiento se elige aparte, sobre cualquier modelo. Deep Think requiere Pro y el plan Google AI Ultra.

NIVEL 1: RÁPIDOS & LIGEROS

0 Thinking / Esfuerzo Bajo

Claude Haiku 5.5 | Gemini 3.8 Flash (pensamiento bajo) | GPT-6 Luna

Inferencia ultrarrápida (en milisegundos), costo ínfimo y alta fidelidad para tareas directas sin divagaciones.

Tareas Jurídicas Típicas:

- Resumir una providencia simple o cédula de 1 foja.
- Extraer fechas, números de DNI o datos de un contrato.
- Redactar correos electrónicos formales a clientes.
- Escritos de mero trámite ("Téngase presente", "Librense oficios").

NIVEL 2: CABALLOS DE BATALLA

Esfuerzo Medio

Claude Opus 5.5 | Sonnet 5.5 | Gemini 3.8 Flash / 3.1 Pro | GPT-6 So1

El estándar para el 80% del trabajo cotidiano del despacho. En la serie 5.5, Opus asume el trabajo complejo cotidiano con máxima solidez argumentativa.

Tareas Jurídicas Típicas:

- Redacción de demandas y contestaciones de demanda con negativa categórica.
- Contratos de locación, convenios de liquidación y acuerdos societarios.
- Intercambio epistolar (Cartas Documento e intimaciones fehacientes).
- Recursos de apelación ordinarios y expresión de agravios.

NIVEL 3: RAZONAMIENTO PROFUNDO

Esfuerzo Alto / Thinking

Gemini 4 Argon | Claude Fable 5.1 | GPT-6 Astra

Modelos de frontera con razonamiento autorreflexivo nativo (CoT) y ventanas masivas para resolver contradicciones extremas.

Tareas Jurídicas Típicas:

- Recurso Extraordinario Federal ante la CSJN (Art. 14 Ley 48).
- Cruce de contradicciones en pericias contables y médicas complejas.
- Casos de arbitraje internacional o derecho societario corporativo (M&A).
- Auditoría de sentencias definitivas y detección de arbitrariedad judicial.



Ficha de Frontera: Gemini 4 Argon vs. Astra for Law (OpenAI)

Gemini 4 Argon (<https://gemini.google.com>) (Google DeepMind)

Ficha API Oficial: gemini-4-argon

- ✓ **1.048.576 tokens de contexto:** límite unificado de entrada y salida.
- ✓ **Razonamiento nativo:** piensa antes de responder.
- ✓ **Salida máxima:** 65.536 tokens por defecto y hasta 1.048.576 con decodificación larga.
- ✓ **Uso Principal:** Auditoría de mega-causas y detección de contradicciones probatorias entre cuerpos de prueba.

Astra for Law (<https://chatgpt.com>) (OpenAI / GPT-6 Astra)

Configuración de GPT-6 Astra para estudios jurídicos

✓ **Índice Especializado:** Anclado a 230 millones de URLs de fuentes jurídicas y contractuales.

⚠ **Acceso:** por ahora solo para estudios seleccionados de EE.UU. mediante Trusted Access.

⚠ **Alerta de Cátedra (Derecho Argentino):** Entrenada sobre precedentes anglosajones y tribunales de EE.UU. **No sustituye la compulsa en SAIJ ni InfoLEG;** su valor en Argentina se limita a redacción, síntesis y contratos internacionales.

⚠ **Alerta de Ineficiencia: El Mito del Sobrerazonamiento (Para el Aula):**

Muchos profesionales creen que *"más razonamiento siempre es mejor"*. **Falso:** activar el modo de razonamiento profundo (GPT-6 Astra, Fable 5.1 o el pensamiento extendido de Claude) para pedirle una cédula o redactar una cláusula estándar es un desperdicio: genera demoras innecesarias, agota rápidamente tus límites de mensajes en la suscripción y a veces produce sobre-análisis ("overthinking") en trámites procesales simples. Reservá el Thinking exclusivamente para casos con baches lógicos, contradicciones de prueba o recursos ante tribunales superiores.

Información verificada el 24/10/2026. Los modelos y las políticas cambian seguido: confirmá en la fuente oficial.