Clase IAG: El Despacho Jurídico con IA

Diplomatura IAG | Abog. María Elisa Farías

🔒 Blindaje de Cuentas: Evitar que los Chats Entrenen a la IA

Urgencia Procesal

Buena parte del riesgo no viene de un hackeo sino de **opciones activadas por defecto** en las plataformas comerciales. Cuando usás ChatGPT (https://chatgpt.com), Claude (https://claude.ai) o Gemini (https://gemini.google.com) sin configurar los controles de datos, tus causas pueden ser usadas para reentrenar futuros modelos o ser revisadas por empleados humanos.

ChatGPT (https://chatgpt.com) (OpenAI: GPT-6 Astra, Sol y Luna)

Riesgo Alto

En cuentas Free y Plus, el reentrenamiento viene **activo de fábrica**.

Fuente ↗ (https://help.openai.com/en/articles/7730893-data-controls-in-chatgpt)

Ruta de Blindaje:

1. Clic en tu Perfil > **Configuración**.
2. Menú **Control de Datos**.
3. Desactivar: **"Mejorar el modelo para todos"**.
4. Menú **Personalización > Memoria**: Borrar recuerdos cruzados acumulados.

⚠️ *Peligro de Memoria Cruzada:* La memoria persistente puede recordar datos del Cliente A mientras redactás para el Cliente B.

Claude (https://claude.ai) (Opus 5.5, Sonnet 5.5, Haiku 5.5)

Revisar

En Free, Pro y Max el entrenamiento depende del interruptor “Ayudar a mejorar Claude”. Activado: retención de hasta 5 años. Desactivado: 30 días.

Fuente ↗ (https://privacy.claude.com/en/articles/10023580-is-my-data-used-for-model-training)

Ruta de Blindaje:

1. Clic en Perfil > **Settings**.
2. Menú **Privacy & Data**.
3. Desactivar **Ayudar a mejorar Claude**.
4. En planes *Team* o *Enterprise* rigen los términos comerciales: no se entrena con tus datos. La retención cero no viene por defecto; se negocia aparte.

Google Gemini (https://gemini.google.com) (3.8 Flash, 3.1 Pro y 4 Argon)

Riesgo Crítico

En la app comercial (gemini.google.com (https://gemini.google.com)), **revisores humanos** leen muestras de chats. Fuente ↗ (https://support.google.com/gemini/answer/13594961)

Ruta de Blindaje:

1. Ir a **Actividad en las apps con Gemini**.
2. **Desactivar / Pausar** la actividad.
3. *Ojo con el nivel gratuito:* Google usa tus datos para mejorar sus productos y revisores humanos pueden leerlos. Solo cambia si vinculás facturación (nivel pago).

🚨 *Responsabilidad profesional:* el deber de confidencialidad (Ley 25.326 y normas éticas del Colegio) es personal del letrado. Cargar datos de una causa en una cuenta mal configurada es una decisión propia.

🛡️ El Flujo de Trabajo Seguro en 3 Pasos con VERIFICACIÓN EN CADA PASO

Protocolo Sanity Check

El mayor error del abogado novato es confiar ciegamente en el copiado y pegado. Un despacho profesional aplica **3 compuertas de auditoría estricta** antes de que un texto toque la mesa de entradas del juzgado.

El dato real nunca toca la IA

EN TU EQUIPO

1 Anonimización local

El anonimizador reemplaza nombres y documentos por marcadores, dentro de tu navegador.

```
[PARTE_ACTORA] · [DNI_1]
```

✓ Checkpoint 1 · Fuga Cero

FUERA DE TU EQUIPO

2 Procesamiento en LLM

Claude, ChatGPT o Gemini trabajan solo con marcadores. Verificás cada cita de la respuesta.

Pegar datos personales en una IA con servidores en el exterior es una transferencia internacional (art. 12, Ley 25.326): solo se admite hacia países con protección adecuada, dentro de las excepciones legales o con consentimiento expreso del titular.

Ver regulación (AAIP) ↗ (https://www.argentina.gob.ar/transferencias-internacionales)

```
[PARTE_ACTORA] · [DNI_1]
```

✓ Checkpoint 2 · Anti-Alucinación

EN TU EQUIPO

3 Reidentificación en Word

Con buscar y reemplazar vuelven los datos reales, sin dejar marcadores sueltos.

```
Juan Pérez · DNI 30.123.456
```

✓ Checkpoint 3 · Tokens Huérfanos

▢ Dato real ▢ Marcador anonimizado

PASO 1: Anonimización Local

Pegar el escrito o testimonial en **Anonimizador Web v5.1** (https://anonimizadorv5.mefarias.com.ar). Los Web Workers reemplazan nombres, DNI y patentes en la memoria RAM del navegador sin conexión a servidores.

✓ CHECKPOINT 1: VERIFICACIÓN DE FUGA CERO

Antes de copiar al LLM, hacé un escaneo visual rápido: verificá que no hayan quedado **apodos familiares, nombres de fantasía de comercios, números de póliza o firmas al pie** sin disociar.

PASO 2: Procesamiento en LLM

Ingresar el texto sanitizado (`[PARTE_ACTORA]`, `[DNI_1]`) en Claude Opus 5.5 / Sonnet 5.5 (https://claude.ai), Gemini 3.8 Flash (https://gemini.google.com) o GPT-6 Sol (https://chatgpt.com) para estructurar la demanda, contestación o análisis de prueba.

✓ CHECKPOINT 2: VERIFICACIÓN ANTI-ALUCINACIÓN

Auditá la salida del modelo: verificá que las **normas citadas estén vigentes** (no cite artículos del Código de Vélez ni leyes derogadas), que los artículos del CCyCN existan y que no haya inventado fallos de la CSJN de memoria.

PASO 3: Reidentificación en Word

Copiar la respuesta del modelo y en tu procesador de textos local (Word o LibreOffice) usar "Buscar y Reemplazar" para restituir los datos verdaderos del cliente.

✓ CHECKPOINT 3: VERIFICACIÓN DE TOKENS HUÉRFANOS

En Word ejecutá `Ctrl + F` y buscá el caracter `[`. **Certificá que no haya quedado ningún token huérfano** (`[COMPLETAR DATO]`, `[PARTE_DEMANDADA]`) en el escrito final antes de la firma electrónica.