DESPACHO SEGURO

Guía de Ciberseguridad con Inteligencia Artificial

Equipo Redacción – derechoartificial.com

Edición: junio de 2025

Índice

1. Resumen ejecutivo

El despacho migra a un sistema basado en inteligencia artificial, lo que exige una ciberseguridad de nivel máximo. Este documento detalla un sistema completo construido sobre arquitectura Zero Trust, que se orienta al cumplimiento de las principales normativas aplicables: RGPD/GDPR, NIS2 (con las salvedades de ámbito subjetivo explicadas en el punto 3), Reglamento de IA (UE) 2024/1689, CCPA/CPRA y NYDFS 23 NYCRR Part 500 (como estándar de referencia y en la medida en que resulte de aplicación indirecta). Se describen los componentes técnicos —firewalls de siguiente generación, SASE, XDR, SIEM/SOAR, PAM, KMS, DLP— y los procesos organizativos —CISO, formación, simulacros, plan de respuesta a incidentes—. Se incluye un plan de implementación realista para un despacho mediano de entre 20 y 100 profesionales.

2. Arquitectura de Confianza Cero (Zero Trust)

2.1. Introducción y justificación para un despacho de abogados

La arquitectura tradicional de seguridad basada en perímetro —firewall en la entrada de la oficina, VPN para acceso remoto, red interna plana— resulta insuficiente para un despacho que gestiona datos sensibles de clientes, herramientas de IA y abogados que trabajan desde cualquier ubicación. El modelo Zero Trust parte de un principio radical: *nunca confiar, siempre verificar*. No existe una red interna "segura" por defecto.

Para un bufete, esto implica lo siguiente:

- Un abogado que accede desde su portátil en la oficina está sujeto a los mismos controles que si accede desde una cafetería.
- Un asistente de IA (agente automatizado) tiene permisos estrictamente limitados y monitorizados, del mismo modo que cualquier empleado humano.
- Un atacante que compromete un equipo no puede moverse lateralmente hacia los expedientes de otros clientes.

2.2. Componentes clave de la arquitectura Zero Trust

A continuación se describen los elementos técnicos necesarios, con independencia del proveedor, que deben implantarse.

2.2.1. Zero Trust Network Access (ZTNA) — Sustitución de la VPN

- Función: conectar a usuarios y dispositivos directamente con las aplicaciones autorizadas, sin integrarlos en la red interna general.
- Funcionamiento: el usuario inicia sesión mediante un cliente ligero o navegador; el conector se autentica ante un broker ZTNA; este verifica identidad, estado del dispositivo, ubicación y políticas; si todo es correcto, se establece un túnel cifrado directo hacia la aplicación específica (no hacia la red completa).
- Requisitos para el despacho: MFA obligatoria en cada conexión; políticas por área de práctica (p. ej., solo el área fiscal accede a la aplicación de cálculos tributarios); registro de todas las conexiones para auditoría.

2.2.2. Microsegmentación de red

- Función: dividir la red interna en segmentos lógicos aislados de forma que, aunque un atacante acceda a uno de ellos, no pueda alcanzar los demás.
- Implementación práctica: plano de control centralizado que gestiona políticas de firewall distribuidas; cada segmento se define mediante etiquetas (p. ej., "cliente-X", "sistema-contable", "IA-entrenamiento"); el tráfico entre segmentos solo se permite mediante reglas explícitas.

Ejemplo de segmentos para el despacho:

Segmento	Contenido	¿Con quién se comunica?
CORP-RRHH	Datos de empleados y nóminas	Solo con CORP-AD y backup
LEG-A	Expedientes del cliente A (confidencial)	Solo con el segmento de impresión autorizada y con el ZTNA gateway
LEG-B	Expedientes del cliente B	Sin comunicación con LEG-A
IA-SANDBOX	Entorno de pruebas de IA	Solo con internet controlado (sin acceso a LEG-*)
IA-PROD	IA en producción que trata datos reales	Con LEG-A y LEG-B bajo políticas muy restrictivas
ADMIN	Controladores de dominio y backups	Solo con los segmentos que lo necesiten

La microsegmentación ayuda a demostrar ante reguladores que se han aplicado medidas de contención de brechas, lo que resulta relevante a efectos de NIS2 y, en su caso, de NYDFS.

2.2.3. Identity-Aware Proxy (IAP) para aplicaciones web

- Función: todas las aplicaciones web internas (intranet, gestor documental, portal del cliente) se publican detrás de un proxy que verifica identidad y contexto antes de permitir el acceso.
- Beneficio: las aplicaciones no quedan expuestas directamente; el atacante no puede ni detectar el puerto de la aplicación si no supera la verificación.
- Configuración típica: el IAP recibe la petición HTTPS; redirige al proveedor de identidad (IdP) para autenticación MFA; comprueba el estado del dispositivo; si todo es correcto, inyecta cabeceras con el usuario autenticado y reenvía la petición a la aplicación interna.

2.2.4. Verificación continua de salud del dispositivo (Posture Checking)

- Función: antes de conceder acceso, y periódicamente durante el mismo, se verifica que el dispositivo del usuario cumple las políticas de seguridad.

Políticas mínimas:

- Sistema operativo con parches de seguridad aplicados en los últimos 30 días.
- Antivirus/EDR activo y con definiciones actualizadas.
- Disco duro cifrado (BitLocker, FileVault o equivalente).
- Pantalla bloqueada con contraseña o biométrico.
- Ausencia de software no autorizado (lista de permitidos).

Frecuencia estándar del despacho: re-verificación cada 30 minutos durante una sesión activa. En entornos de alto riesgo (acceso desde redes no corporativas) puede reducirse a 15 minutos. El intervalo de 30 minutos equilibra seguridad y usabilidad.

2.2.5. Red de confianza cero para oficinas físicas y sedes

- Oficina principal y sucursales: no existe una VLAN "todo permitido". Cada toma de red asigna al usuario a su segmento lógico según su rol, independientemente del puerto físico.
- Wi-Fi corporativo: se configura un SSID único con asignación dinámica de VLAN mediante 802.1X con certificados. Un abogado del área fiscal y uno del área laboral se conectan a la misma red Wi-Fi pero terminan en segmentos diferentes.
- Dispositivos invitados o personales (BYOD): solo se permite acceso a internet a través de un portal cautivo; nunca a la red interna. Si un abogado desea usar su dispositivo personal, debe pasar por ZTNA con MFA y comprobación de posture limitada.

2.3. Flujo de trabajo de un acceso típico bajo Zero Trust

- El abogado enciende su portátil corporativo.
- El agente de posture verifica: disco cifrado, antivirus actualizado.
- El abogado abre el navegador e intenta acceder al gestor documental (URL interna).
- El IAP intercepta la petición y redirige al IdP (Azure AD, Okta, etc.) para autenticación MFA.
- El broker ZTNA comprueba el contexto: hora laboral, ubicación geográfica. Si el acceso procede de una zona de riesgo, se exige verificación adicional.
- Si todo es correcto, el broker crea un túnel cifrado exclusivamente hacia el gestor documental. El abogado no puede hacer ping a otros equipos ni escanear la red.
- Cada 30 minutos, el agente de posture re-verifica el dispositivo. Si el antivirus se desactiva, el acceso se corta automáticamente.

2.4. Relación con el cumplimiento normativo

- NIS2: exige medidas de segmentación y control de accesos para las entidades en su ámbito de aplicación. Zero Trust proporciona los mecanismos técnicos adecuados.
- NYDFS (23 NYCRR 500): exige inventario de sistemas, MFA y monitorización. Zero Trust los proporciona.
- RGPD: el principio de minimización de datos se extiende a la red; solo se expone lo estrictamente necesario.

2.5. Consideraciones de implementación para un despacho mediano

- Tamaño típico: 20-100 abogados y personal administrativo.
- Presupuesto estimado de esta capa (sin marcas): entre 30.000 € y 80.000 € anuales en licencias ZTNA + microsegmentación + IAP, dependiendo del proveedor.
- Personal necesario: un responsable de TI (interno o externalizado) con conocimientos en Zero Trust. Se recomienda formación específica certificada.
- Fase piloto: iniciar con un grupo de 5 abogados voluntarios y un segmento no crítico.

2.6. Posibles riesgos y mitigaciones

Riesgo	Mitigación
El broker ZTNA falla y deja sin acceso a todos los usuarios	Desplegar el broker en alta disponibilidad (mínimo 2 nodos) y mantener un plan de contingencia con acceso de emergencia muy limitado.
Los abogados perciben el sistema como lento	Optimizar políticas: verificar posture cada 30 minutos; usar caché de autenticación; elegir un proveedor con baja latencia.
Dificultad para integrar aplicaciones antiguas (legacy)	Usar un conector ZTNA que sirva la aplicación como si fuera web, o mantener un segmento especial legacy con controles adicionales.

2.7. Resumen de entregables para este punto

- ☐ Diagrama de red Zero Trust de alto nivel.
- ☐ Mapa de segmentos (mínimo 5 segmentos).
- ☐ Política de posture checking estándar.
- ☐ Procedimiento de incorporación (onboarding) de un nuevo abogado en ZTNA.
- ☐ Plan de pruebas de la arquitectura en un entorno aislado.

3. Cumplimiento normativo detallado (Europa y EE. UU.)

3.1. Introducción al marco normativo aplicable

Un despacho de tamaño medio que gestiona datos de clientes en Europa y en Estados Unidos, y que incorpora además herramientas de inteligencia artificial, debe observar un entramado normativo complejo y en evolución. El incumplimiento no solo conlleva sanciones económicas de gran cuantía, sino también pérdida de confianza de los clientes e, incluso, consecuencias disciplinarias en el ámbito colegial.

Las normativas que se abordan en este punto son:

- Unión Europea: Reglamento General de Protección de Datos (RGPD/GDPR), Ley Orgánica 3/2018 de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD), Directiva NIS2 (Directiva (UE) 2022/2555), y el Reglamento de Inteligencia Artificial (Reglamento (UE) 2024/1689).
- Estados Unidos: Ley de Privacidad del Consumidor de California (CCPA/CPRA) y la normativa de ciberseguridad del Departamento de Servicios Financieros de Nueva York (NYDFS 23 NYCRR Part 500), esta última como estándar de referencia y en la medida en que resulte de aplicación indirecta o contractualmente exigida.
- Normativa colegial: Código Deontológico de la Abogacía Española (aprobado por el Consejo General de la Abogacía Española) y las guías sobre uso de tecnología e IA publicadas por los Colegios de Abogados.

3.2. Reglamento General de Protección de Datos (RGPD/GDPR) y LOPDGDD

3.2.1. Ámbito de aplicación

El RGPD aplica cuando el despacho trata datos personales de residentes en la UE, con independencia del lugar de establecimiento del despacho. La LOPDGDD (Ley Orgánica 3/2018) complementa al RGPD en el ordenamiento español, con disposiciones específicas sobre datos de condenas penales, datos laborales y derechos digitales. El cumplimiento de ambos textos es obligatorio.

3.2.2. Principios clave y su implementación técnica

Principio RGPD	Implementación en el sistema
Licitud, lealtad y transparencia (art. 5.1.a)	El despacho informará a los clientes sobre el uso de IA. Se redactarán cláusulas informativas actualizadas. Todo tratamiento de datos para entrenar o ajustar modelos de IA exige una base jurídica explícita (consentimiento o interés legítimo con balance documentado).
Limitación de la finalidad (art. 5.1.b)	Los datos recabados para un litigio no pueden emplearse para fines incompatibles (p. ej., marketing). El sistema DLP y la microsegmentación impedirán movimientos de datos no autorizados.

Minimización de datos (art. 5.1.c)	Solo se procesan los datos estrictamente necesarios. Las herramientas de IA se configurarán para no almacenar datos históricos más allá de lo preciso para la tarea.
Exactitud (art. 5.1.d)	Se implementarán procedimientos de revisión humana de las salidas de IA y mecanismos para que los clientes corrijan datos erróneos.
Limitación del plazo de conservación (art. 5.1.e)	Políticas automáticas de retención: los expedientes se eliminan tras el plazo legal aplicable (generalmente 5-10 años según el tipo de asunto). La IA no retendrá datos personales más allá de la sesión de procesamiento, salvo justificación documentada.
Integridad y confidencialidad (art. 5.1.f)	Cifrado en reposo y en tránsito, controles de acceso ZTNA, auditoría continua.
Responsabilidad proactiva (art. 5.2)	El despacho debe poder demostrar el cumplimiento. El sistema generará informes de auditoría, registros de acceso y evaluaciones de impacto (DPIA) firmadas.

3.2.3. Delegado de Protección de Datos (DPD/DPO)

El artículo 37 RGPD exige la designación de un DPO cuando las actividades principales del responsable implican el tratamiento a gran escala de categorías especiales de datos (art. 9 RGPD: datos de salud, condenas penales, origen racial, etc.). La LOPDGDD (art. 34) amplía esta obligación a colegios de abogados y asesorías jurídicas.

Para un despacho mediano, la obligatoriedad del DPO depende de si el tratamiento de datos especiales constituye una actividad principal y se realiza a gran escala. Los indicadores de "gran escala" incluyen el número de interesados afectados, el volumen de datos, la duración del tratamiento y el ámbito geográfico (considerando 91 RGPD). Despachos que tramiten exclusivamente asuntos individuales pueden no alcanzar el umbral de "gran escala"; no obstante, la LOPDGDD establece la obligatoriedad para los colegios de abogados, lo que hace recomendable —y para muchos despachos obligatorio— la designación de un DPO. Se designará un DPO interno o externo desde el inicio del proyecto, independientemente del análisis de umbrales, como medida de cumplimiento proactivo y de diligencia debida.

3.2.4. Evaluación de Impacto relativa a la Protección de Datos (DPIA)

Es obligatoria antes de emplear IA que evalúe aspectos personales a gran escala o con probable alto riesgo (art. 35 RGPD). Se realizará una DPIA para cada herramienta de IA, que incluirá: descripción del tratamiento y sus fines; evaluación de necesidad y proporcionalidad; identificación de riesgos para los interesados; medidas de mitigación; consulta al DPO y, si persiste un riesgo residual alto, consulta previa a la AEPD.

3.2.5. Notificación de brechas de seguridad en 72 horas

- El SIEM genera una alerta crítica ante la detección de un acceso no autorizado.
- El equipo de seguridad (SOC) evalúa el alcance en un máximo de 2 horas.
- El DPO determina si la brecha entraña riesgo para los derechos de los interesados.
- Si es así, se notifica a la autoridad de control (AEPD u autoridad competente) mediante el formulario oficial en menos de 72 horas desde que se tuvo conocimiento (art. 33 RGPD).
- Si el riesgo es alto para los interesados, se les notifica sin dilación indebida (art. 34 RGPD).

Requisito técnico: los registros deben ser inmutables y disponer de marca de tiempo precisa (NTP sincronizado). Se requiere un procedimiento documentado, simulado semestralmente.

3.2.6. Derechos de los interesados (ARCO+)

- Acceso: exportar todos los datos personales de un cliente en formato legible, incluidos los datos tratados por IA.
- Rectificación: corregir datos erróneos en todos los sistemas, incluidos modelos ajustados con esos datos.
- Supresión (derecho al olvido): eliminar los datos cuando ya no sean necesarios o falte base jurídica, incluyendo la eliminación en bases de datos de ajuste de modelos (técnicamente complejo; se implementarán políticas de retención y se evitará el ajuste continuo con datos personales no anonimizados).
- Oposición: cuando el tratamiento se base en interés legítimo.
- Limitación del tratamiento: marcar los datos como bloqueados para determinados usos.

Procedimiento: se habilitará un buzón de privacidad y un formulario web. El sistema debe disponer de APIs o consultas parametrizadas para localizar y extraer o suprimir datos de un cliente específico en todos los repositorios (archivos, bases de datos, registros, cachés de IA). El plazo de respuesta es de un mes, prorrogable a tres (art. 12.3 RGPD).

3.2.7. Transferencias internacionales de datos

Si el despacho utiliza proveedores de nube estadounidenses (AWS, Azure, Google, etc.), debe garantizar un nivel de protección equivalente al europeo. Las vías disponibles son:

- Marco de Privacidad de Datos UE-EE. UU. (EU-US Data Privacy Framework, Decisión de Ejecución (UE) 2023/1795 de la Comisión, de 10 de julio de 2023): los proveedores certificados en este marco se consideran destino adecuado sin necesidad de Cláusulas Contractuales Tipo adicionales. Verificar que el proveedor concreto figure en la lista actualizada del DPF (www.dataprivacyframework.gov).
- Cláusulas Contractuales Tipo (SCC, Decisión de Ejecución 2021/914): aplicables cuando el proveedor no esté certificado en el DPF. Se complementan con una evaluación de las leyes locales (p. ej., FISA 702, EO 12333) para verificar que no existe acceso desproporcionado por agencias de inteligencia, y con medidas técnicas suplementarias (cifrado con claves gestionadas por el despacho).

3.3. Directiva NIS2 (UE) — Seguridad de redes y sistemas de información

3.3.1. Ámbito de aplicación: ¿aplica a un despacho de abogados?

La Directiva (UE) 2022/2555 (NIS2) amplió el ámbito subjetivo de su predecesora, pero no incluye a los despachos de abogados en sus Anexos I y II. Los sectores regulados son: energía, transporte, banca, infraestructuras de mercados financieros, sector sanitario, agua potable, aguas residuales, infraestructuras digitales, gestión de servicios TIC, administración pública y espacio (entidades esenciales, Anexo I); correos y mensajería, gestión de residuos, química, alimentación, fabricación, proveedores digitales e investigación (entidades importantes, Anexo II).

En consecuencia, un despacho de abogados no está sujeto directamente a NIS2 por el mero hecho de superar 50 empleados o 10 millones de euros de facturación. Sin embargo, NIS2 puede afectarle de dos formas:

- De forma indirecta, cuando presta servicios a entidades esenciales o importantes que, a su vez, están obligadas a evaluar la seguridad de su cadena de suministro, lo que incluye los servicios jurídicos externos que reciben (art. 21.3 NIS2).
- A través de la transposición nacional española: la Ley que transpone NIS2 en España (pendiente de publicación en la fecha de redacción de esta guía) puede ampliar el ámbito a sectores no previstos en los Anexos europeos. Se deberá verificar su contenido una vez publicada.

En todo caso, se recomienda que el despacho adopte las medidas de NIS2 como estándar de referencia, tanto por la presión regulatoria indirecta de sus clientes como por la elevada calidad técnica que dichas medidas implican.
Sanciones previstas por NIS2 (para las entidades en su ámbito directo):

- Entidades esenciales: hasta 10.000.000 € o, si es superior, el 2 % del volumen de negocio anual mundial (art. 34.2 NIS2).
- Entidades importantes: hasta 7.000.000 € o, si es superior, el 1,4 % del volumen de negocio anual mundial (art. 34.3 NIS2).

3.3.2. Obligaciones principales y su traducción técnica

Obligación NIS2	Implementación
Políticas de análisis de riesgos y seguridad de la información	Política documentada, revisada anualmente, que incluya ciberseguridad, continuidad de negocio, gestión de incidencias y cadena de suministro. Aprobada por el órgano de administración (socio director o consejo).
Gestión de incidencias	Procedimiento de detección, registro, análisis y notificación: alerta temprana en 24 h, notificación completa en 72 h, informe final en 1 mes.
Continuidad del negocio y gestión de crisis	Plan de continuidad (BCP) que contemple fallos del sistema de IA, ataques de ransomware y caídas de conectividad. Prueba anual documentada.

Seguridad de la cadena de suministro	Evaluación de la seguridad de los proveedores de servicios cloud, software de IA y conectividad. Cláusulas contractuales con obligaciones de notificación de incidentes.
Formación en ciberseguridad de los directivos	Los socios y directores deben recibir formación específica (mínimo cada dos años) sobre riesgos y responsabilidades.
Supervisión y control de la alta dirección	El órgano de administración debe aprobar las medidas de seguridad y supervisar su implementación. Se designará un CISO con recursos suficientes.

3.3.3. Medidas técnicas mínimas exigidas por NIS2

El sistema desarrollado debe cubrir las siguientes áreas:

- Análisis de riesgos y seguridad de los sistemas.
- Gestión de incidencias.
- Continuidad de negocio y recuperación ante desastres.
- Seguridad de la cadena de suministro.
- Seguridad en la adquisición, desarrollo y mantenimiento de sistemas.
- Ciberseguridad básica: gestión de vulnerabilidades, criptografía, control de accesos, monitorización.

3.3.4. Registro y notificación de incidentes

- El despacho debe registrar todos los incidentes de seguridad que afecten a la disponibilidad, autenticidad, integridad o confidencialidad de los datos.
- Los incidentes con impacto significativo se notifican al CSIRT nacional en los plazos indicados.
- El sistema generará automáticamente un informe de incidente con: fecha, duración, sistemas afectados, causas, medidas adoptadas y datos personales comprometidos.

3.4. Reglamento de Inteligencia Artificial (UE) 2024/1689

3.4.1. Clasificación de las herramientas de IA del despacho

El Reglamento de Inteligencia Artificial (RIA) clasifica los sistemas de IA en cuatro niveles de riesgo: riesgo inaceptable (prohibido), alto riesgo, riesgo limitado y riesgo mínimo. Para un despacho legal, la clasificación aplicable es la siguiente:

Uso de IA	Nivel de riesgo	Consecuencias y obligaciones
Herramienta de búsqueda de jurisprudencia (sin evaluación de personas)	Riesgo mínimo	Sin obligaciones específicas, salvo buenas prácticas voluntarias y formación básica del personal.
Asistente de redacción de documentos (propone borradores; no toma decisiones)	Riesgo limitado	Obligación de transparencia: informar al usuario de que está interactuando con IA; posibilidad de revisión y modificación.
Análisis predictivo de éxito de litigios (evalúa probabilidades sobre datos de casos)	Alto riesgo (si se usa para apoyar decisiones en procedimientos judiciales, Anexo III, punto 8, letra a, RIA)	Sistema de gestión de calidad, documentación técnica, supervisión humana efectiva, registro en la base de datos de la UE, evaluación de conformidad.
Selección de personal (filtrado de candidaturas)	Alto riesgo (Anexo III, punto 4, RIA)	Ídem.

Nota sobre la evaluación de conformidad para sistemas de alto riesgo: El RIA distingue entre proveedores (fabricantes) y responsables del despliegue (deployers). La obligación de realizar la evaluación de conformidad recae sobre el proveedor del sistema, no sobre el despacho como deployer. Para la mayoría de los sistemas de alto riesgo del Anexo III, la evaluación puede efectuarse mediante autoevaluación interna del proveedor (art. 43.2 RIA), sin necesidad de un organismo notificado. Solo determinadas categorías —como sistemas biométricos de identificación remota en tiempo real— exigen evaluación por tercero certificado. El despacho, como deployer, debe exigir al proveedor la documentación técnica y la declaración de conformidad UE, y registrar el sistema en la base de datos de la UE cuando así lo exija el RIA (art. 71).

3.4.2. Calendario de aplicación del Reglamento de IA

El RIA entró en vigor el 1 de agosto de 2024 con la siguiente aplicación progresiva:

- 2 de febrero de 2025: prohibición de las prácticas de IA inaceptables (art. 5).

- 2 de agosto de 2025: obligaciones para sistemas de IA de propósito general (GPAI) y gobernanza.
- 2 de agosto de 2026: plena aplicación de las obligaciones para sistemas de alto riesgo del Anexo III.
- 2 de agosto de 2027: para sistemas de IA integrados en productos regulados por otra normativa sectorial.

3.4.3. Obligaciones transversales para todo sistema de IA

- Alfabetización en IA (art. 4 RIA): el despacho debe garantizar que el personal que usa o supervisa la IA tenga conocimientos suficientes sobre su funcionamiento, limitaciones y riesgos. Se impartirá formación específica (vid. punto 7).
- Transparencia: el sistema debe indicar claramente que el contenido ha sido generado o asistido por IA, salvo que un ser humano lo haya revisado y asuma la autoría.
- Supervisión humana efectiva (art. 14 RIA): para sistemas de alto riesgo, la supervisión debe ser documentada, efectiva y con capacidad de intervenir o detener el sistema. El RIA no cuantifica el número de revisores, sino que exige que la supervisión sea adecuada al riesgo concreto del sistema.

3.4.4. Documentación necesaria

Para cada sistema de IA en uso, el despacho debe conservar:

- Descripción del sistema y su propósito.
- Clasificación de riesgo justificada.
- Declaración de conformidad del proveedor (para sistemas de alto riesgo).
- Instrucciones de uso para los abogados.
- Registro de incidentes relacionados con la IA.
- DPIA (para sistemas que traten datos personales con riesgo alto).

3.5. Ley de Privacidad del Consumidor de California (CCPA/CPRA)

3.5.1. ¿Aplica a un despacho europeo?

La CCPA/CPRA aplica a cualquier entidad con fines de lucro que haga negocios en California y que cumpla al menos uno de los siguientes umbrales:

- Ingresos brutos anuales superiores a 25 millones de dólares.
- Compra, venta, recepción o intercambio de datos personales de 100.000 o más consumidores o unidades domésticas (umbral elevado por la CPRA desde el 1 de enero de 2023; el umbral previo de 50.000 consumidores quedó derogado).
- Derivación del 50 % o más de los ingresos anuales de la venta o intercambio de datos personales de consumidores.

Un despacho mediano europeo con clientes en California podría estar sujeto si supera alguno de esos umbrales. Para muchos despachos medianos, lo prudente es aplicar voluntariamente los principios de transparencia de la CCPA, dada la complejidad de verificar el incumplimiento ex ante.

3.5.2. Derechos de los consumidores californianos

- Derecho a saber qué datos personales se recopilan, usan, comparten o venden.
- Derecho a eliminar los datos personales.
- Derecho a corregir datos inexactos.
- Derecho a optar por no participar en la venta o intercambio de datos personales. El despacho no vende datos, pero si los comparte con proveedores de IA en condiciones asimilables a "intercambio" (sharing), debe informar de ello y ofrecer la opción de exclusión.
- Derecho a limitar el uso de datos sensibles.
- Derecho a no ser discriminado por ejercer los derechos anteriores.

3.5.3. Obligaciones de transparencia

- Publicar un aviso de privacidad en el sitio web que incluya: categorías de datos recopilados; fuentes; fines del tratamiento (incluyendo uso de IA); categorías de terceros con quienes se comparten datos; enlace a "Do Not Sell or Share My Personal Information".
- Habilitar un canal de solicitud (teléfono, formulario web) para que los clientes ejerzan sus derechos en el plazo de 45 días.

3.5.4. Medidas técnicas

- El sistema debe permitir la búsqueda, extracción y eliminación de datos de un consumidor californiano específico, análogamente al RGPD.
- Se implementará un mecanismo de exclusión de venta/intercambio (opt-out) con soporte para la señal Global Privacy Control (GPC).

3.6. NYDFS 23 NYCRR Part 500 (Nueva York)

3.6.1. ¿Aplica al despacho?

La normativa NYDFS 23 NYCRR Part 500 es de aplicación directa a las "covered entities", es decir, personas o entidades sujetas a supervisión, licencia, registro o carta de NYDFS (bancos, aseguradoras, prestamistas hipotecarios, etc.). Un despacho de abogados no es, por sí mismo, una entidad supervisada por NYDFS y, en consecuencia, no está directamente obligado por esta normativa.

No obstante, la normativa puede afectar al despacho de dos formas:

- De forma indirecta, cuando sus clientes (bancos, aseguradoras u otras entidades reguladas por NYDFS) exigen al despacho, como "Third-Party Service Provider" (sección 500.11), que cumpla con unos requisitos mínimos de ciberseguridad. En ese caso, la obligación nace del contrato de servicios, no de la norma directamente.
- Como estándar de referencia técnica: el NYDFS 23 NYCRR Part 500 contiene requisitos técnicos muy avanzados que constituyen un modelo de madurez en ciberseguridad de primer nivel. El despacho que los adopte voluntariamente estará mejor posicionado para satisfacer auditorías de sus clientes financieros.

3.6.2. Obligaciones clave y su implementación

Sección	Obligación	Medida en el sistema
500.02	Política de seguridad de la información	Documento formal aprobado por el consejo, revisado anualmente.
500.03	Designación de un responsable de seguridad (CISO)	El mismo CISO del despacho.
500.04	Evaluación de riesgos	Realizada cada tres años o tras cambios significativos (p. ej., implantación de IA).
500.05	MFA obligatoria	Para cualquier acceso externo a la red del despacho y para accesos a datos confidenciales.
500.06	Acceso de mínimo privilegio	Microsegmentación + PAM.
500.07	Pruebas de penetración y análisis de vulnerabilidades	Pruebas de penetración anuales; análisis de vulnerabilidades trimestrales.
500.08	Monitorización continua	SIEM + SOC 24/7.
500.09	Plan de respuesta a incidentes	Documentado y probado anualmente.
500.10	Cifrado en reposo y en tránsito	AES-256 en reposo; TLS 1.3 en tránsito; gestión de claves interna.
500.11	Gestión de proveedores externos	Evaluación de seguridad de proveedores de nube e IA.
500.12	Notificación de incidentes en 72 horas	Procedimiento equivalente al RGPD.
500.15	Registros de actividad (audit trails)	Logs inmutables con retención de 6 años.
500.16	Inventario de sistemas y datos	Herramienta de inventario automático, actualización mensual.

3.7. Secreto profesional y confidencialidad

3.7.1. Fundamento

Además de las leyes de protección de datos, los abogados están sujetos al deber de secreto profesional consagrado en el artículo 542.3 de la Ley Orgánica del Poder Judicial, en el artículo 22 del Estatuto General de la Abogacía Española (Real Decreto 135/2021) y

en los códigos deontológicos colegiales. Su vulneración puede conllevar sanciones disciplinarias, incluida la inhabilitación.

3.7.2. Implicaciones para la arquitectura de IA

- Prohibición de usar asistentes de IA públicos (p. ej., versiones de consumo de ChatGPT, Gemini, etc.) para datos confidenciales de clientes. Solo se emplearán modelos on-premise o en nube privada con acuerdos de confidencialidad y sin retención de datos para entrenamiento.
- El abogado es el responsable final del contenido generado por IA: el sistema debe garantizar la revisión humana completa antes de que cualquier output llegue al cliente.
- Los registros de acceso a expedientes deben restringirse a los abogados del caso y al personal expresamente autorizado. El DPO y el CISO pueden auditar, pero no revelar el contenido a terceros sin base jurídica.
- El Consejo General de la Abogacía Española (CGAE) ha publicado guías sobre el uso de IA que deben consultarse y observarse como parte del cumplimiento deontológico.

3.8. Matriz de cumplimiento unificada

Para facilitar la gestión, se creará una matriz de controles que relacione cada requisito normativo con una medida técnica concreta, un responsable y un plazo de revisión:

Requisito	Normativa	Medida técnica	Responsable	Frecuencia
Notificar brecha en 72 h	RGPD art. 33; NIS2 art. 23	SIEM + procedimiento de respuesta	CISO + DPO	Simulación semestral
MFA obligatoria	NYDFS s. 500.05; NIS2 best practice	ZTNA con IdP + MFA	CISO	Revisión trimestral
DPIA para IA	RGPD art. 35; RIA art. 9-10	DPIA + evaluación de conformidad del proveedor	DPO + CISO	Antes de implantar IA
Supervisión humana de IA	RIA art. 14	Interfaz de revisión, logs de validación	Socio director	Por cada uso de IA de alto riesgo
Política de gestión de riesgos	NIS2 art. 21; NYDFS s. 500.04	Documento aprobado	Dirección	Anual
Inventario de sistemas	NYDFS s. 500.16	Herramienta de inventario automático	TI	Mensual

3.9. Riesgos específicos de incumplimiento y su mitigación

Riesgo	Consecuencia	Mitigación
Uso de IA de alto riesgo sin evaluación del proveedor	Sanciones del RIA: hasta 15.000.000 € o 3 % de la facturación global anual del proveedor (art. 99 RIA); responsabilidad del deployer por incumplimiento de sus propias obligaciones	Clasificar todas las IA antes de implantar. Exigir al proveedor la documentación técnica y la declaración de conformidad UE.
No designar DPO cuando es obligatorio	Multa RGPD de hasta 10.000.000 € o 2 % de la facturación global (art. 83.4 RGPD)	Nombrar DPO desde el inicio del proyecto.
Brecha no notificada en 72 h	Multa + daño reputacional	Automatizar la alerta y disponer de un procedimiento documentado y simulado.
Proveedor cloud sin garantías adecuadas	Riesgo de transferencia internacional inválida; posibles sanciones RGPD	Verificar certificación en el EU-US DPF o celebrar SCCs; aplicar cifrado con claves propias.

4. Componentes técnicos y topología de servidores

4.1. Introducción y visión general

Este punto describe la infraestructura física y lógica necesaria para alojar el sistema completo, incluyendo servidores, almacenamiento, red, componentes de seguridad e integración con herramientas de IA. Se detallan las especificaciones mínimas, las opciones de despliegue y la topología de red interna, alineadas con la arquitectura Zero Trust y los requisitos de cumplimiento normativo. Para un despacho de entre 20 y 100 profesionales se recomienda un modelo híbrido:

- Nube privada o nube virtual soberana (europea) para los datos sensibles y los sistemas de IA en producción.
- Componentes de borde y seguridad en las oficinas físicas (firewalls NGFW, SD-WAN).
- Almacenamiento local para copias de seguridad críticas (regla 3-2-1).

4.2. Catálogo de componentes técnicos (por capas)

4.2.1. Capa de perímetro y red

Componente	Descripción	Especificación mínima
Firewall de siguiente generación (NGFW)	Inspección de tráfico entrante y saliente, descifrado SSL, prevención de intrusiones (IPS), filtrado de aplicaciones.	Rendimiento de al menos 2 Gbps con inspección SSL; alta disponibilidad (2 unidades activo-pasivo).
ZTNA Gateway	Punto de entrada para usuarios remotos y dispositivos, con autenticación MFA, posture checking y enrutamiento a aplicaciones internas.	Capacidad para al menos 200 conexiones simultáneas (adecuado para el tamaño previsto de 20-100 usuarios).
SD-WAN	Conectividad optimizada entre oficinas y con la nube, con conmutación automática ante caídas y enrutamiento basado en políticas.	Mínimo 2 enlaces de banda ancha (fibra + 5G redundante).
Switches de acceso con 802.1X	Asignación dinámica de VLAN según identidad del usuario.	Switches gestionables con soporte RADIUS.
Microsegmentación por software	Controlador central que define políticas entre servidores y cargas de trabajo.	Agente en cada hipervisor o en cada VM; consola central.

4.2.2. Capa de servidores y almacenamiento

Componente	Descripción	Especificación mínima (on-premise)	Alternativa en nube
Hipervisores (3 nodos)	Virtualización de servidores de aplicaciones, bases de datos y controladores de dominio.	CPU: 2 × 16 núcleos; RAM: 256 GB DDR4; almacenamiento compartido.	Instancias reservadas (p. ej., AWS M6i, Azure D16s).
Servidor de gestión de identidades (IdP)	Proveedor de identidad (Active Directory o equivalente).	VM con 8 vCPU, 32 GB RAM.	Azure AD, Okta o Google Workspace.
Servidor de base de datos principal	Datos relacionales de expedientes, clientes y facturación.	16 vCPU, 64 GB RAM, SSD rápido, replicación síncrona.	RDS o equivalente en alta disponibilidad.
Servidores de aplicaciones (incluidas IA)	Gestor documental, CRM, asistentes de IA (inferencia).	8 vCPU, 32 GB RAM (escalable).	Instancias con GPU opcional para IA.
NAS (documentos legales)	Sistema de archivos compartido con cifrado en reposo, versionado y retención legal.	Redundante con 10-20 TB (RAID 6) + snapshots horarios.	Buckets S3 con bloqueo de objetos.
Almacenamiento de backups	Copias de seguridad fuera de línea o inmutables.	2× la capacidad total (p. ej., 40 TB) en disco o cinta.	Glacier Deep Archive o similar.
Servidor de gestión de claves (KMS)	Almacenamiento de claves de cifrado con HSMs opcionales.	VM aislada con acceso restringido.	KMS gestionado (AWS KMS, Azure Key Vault) con claves del cliente.

4.2.3. Capa de seguridad y monitorización

Componente	Descripción	Despliegue
SIEM	Centralización de logs con correlación y alertas.	VM dedicada o servicio cloud.
SOAR	Orquestación de respuestas automáticas (p. ej., aislar un endpoint infectado).	Integrado con SIEM.
XDR / EDR	Agente en cada endpoint que monitoriza comportamientos anómalos y ransomware.	Consola central + agentes.
DLP (Data Loss Prevention)	Políticas para evitar fugas de datos por correo, USB, impresión y portapapeles.	Gateway de correo + agente en endpoints.
PAM (Privileged Access Management)	Gestión de cuentas de administrador y de servicio. Rotación automática de credenciales; registro de sesiones.	VM + agente en servidores críticos.
Gestor de vulnerabilidades	Escaneo periódico de servidores, aplicaciones y contenedores.	Herramienta de red o agente interno.

4.3. Topología de red detallada (con microsegmentación)

4.3.1. Segmentos (VLANs o subredes lógicas)

Segmento	VLAN	Propósito	Acceso desde otros segmentos
WAN-EDGE	10	Firewalls externos, ZTNA gateway, SD-WAN.	Solo hacia internet y hacia DMZ-PUB.
DMZ-PUB	20	Servidores que publican servicios web (portal del cliente).	Desde WAN-EDGE, hacia CORP-AD y CORP-APP autorizados.
CORP-AD	30	Controladores de dominio, IdP, KMS, PAM.	Solo desde segmentos que necesiten autenticación, por puertos específicos.
CORP-APP	40	Aplicaciones internas (gestor documental, CRM, facturación).	Desde ZTNA gateway; desde IA-PROD de forma limitada.
CORP-DB	50	Bases de datos relacionales y documentales.	Solo desde CORP-APP e IA-PROD por puertos específicos.

LEG-A	101	Datos del cliente A (expedientes, correos).	Solo desde CORP-APP autorizado y desde IA-PROD con política de solo lectura, mediante API REST o HTTPS, nunca mediante protocolos de uso compartido de archivos (SMB).
LEG-B	102	Datos del cliente B (aislado de LEG-A).	Ídem; sin comunicación directa con LEG-A.
IA-SANDBOX	60	Entorno de pruebas de IA sin datos reales.	Acceso a internet controlado; sin acceso a LEG-* ni CORP-DB.
IA-PROD	70	Servidores de inferencia IA con datos reales anonimizados o con permisos.	Acceso restringido a LEG-A, LEG-B, CORP-DB. Sin internet salvo listas blancas.
BACKUP	80	Servidores de backup (repositorios inmutables).	Solo desde el motor de backup (segmento ADMIN) y desde BACKUP secundario.
ADMIN	90	Administración (CISO, TI).	Acceso a todos los segmentos con MFA y sesiones grabadas mediante PAM.
GUEST	999	Wi-Fi para invitados.	Solo internet; sin acceso a ningún segmento interno.

4.3.2. Políticas de firewall entre segmentos (ejemplos)

Nota importante sobre el acceso de IA a documentos: el segmento IA-PROD no debe acceder a los expedientes del segmento LEG mediante el protocolo SMB (TCP/445). SMB ha sido el principal vector de propagación de ciberataques masivos (WannaCry, NotPetya). El acceso debe realizarse exclusivamente a través de una API REST autenticada con tokens de corta vida, o mediante HTTPS hacia el gestor documental, conforme a los principios Zero Trust.

Origen	Destino	Protocolo:Puerto	Permiso	Justificación
ZTNA Gateway	CORP-APP	TCP/443, TCP/8443	Sí	Acceso a aplicaciones web internas.
CORP-APP	CORP-DB	TCP/5432 (PostgreSQL)	Sí	Lectura/escritura de datos de expedientes.
IA-PROD	LEG-A	TCP/443 (API REST del gestor documental, solo lectura)	Sí	Leer documentos para análisis; acceso controlado por token; sin protocolo SMB.

IA-PROD	CORP-DB	TCP/5432 solo lectura	Sí	Obtener metadatos de casos.
IA-PROD	Internet	TCP/443 (listas blancas)	Solo a destinos expresamente autorizados	Descargar actualizaciones de modelos desde repositorios oficiales.
LEG-A	LEG-B	Cualquiera	No	Aislamiento estricto entre clientes.
IA-SANDBOX	CORP-*	Cualquiera	No	El entorno de pruebas no toca datos reales.
ADMIN	Cualquier segmento	Cualquiera, bajo PAM	Sí, con restricciones	Los administradores usan cuentas just-in-time.

4.4. Opciones de despliegue

4.4.1. Escenario recomendado: híbrido con soberanía de datos

- On-premise reducido (rack en oficina principal): controladores de dominio y servicios de identidad; backups críticos; segmentos LEG-A y LEG-B (datos más sensibles).
- Nube soberana europea: servidores de aplicaciones (gestor documental, CRM); bases de datos principales (replicadas); servidores IA-PROD (con GPU bajo demanda); SIEM y XDR centralizados.

Ventajas: flexibilidad, escalabilidad y cumplimiento normativo (datos sensibles en Europa).

4.4.2. Opción completamente on-premise

Inversión inicial elevada (hardware, climatización, electricidad, personal). Mayor control, pero requiere redundancia total y un SOC externo para monitorización 24/7.

4.4.3. Opción completamente en nube pública

Solo viable con proveedores que ofrezcan regiones en la UE y acuerdos de procesamiento de datos (SCC o certificación DPF). Se debe cifrar todo con claves propias. Requiere un equipo de ingeniería cloud cualificado.

4.5. Especificaciones de servidores (despacho de 50 usuarios)

4.5.1. Servidores on-premise

Servidor	CPU	RAM	Almacenamiento	Red	Observaciones
Hypervisor 1	2× Intel Xeon Gold 5318Y (24c/48t)	256 GB DDR4	2× 960 GB NVMe (OS), 4× 3,84 TB SSD (VMs)	4× 10 GbE	VMs de AD, backups, LEG-A.
Hypervisor 2 (igual)	Ídem	256 GB	Ídem	Ídem	Alta disponibilidad.
NAS principal	—	—	12× 16 TB HDD (RAID 6) + 2× 1,92 TB SSD caché	2× 25 GbE	Almacenamiento compartido iSCSI/NFS.
Backup Server	1× Intel Xeon Silver 4310	64 GB	4× 18 TB HDD (RAID 10)	2× 10 GbE	Repositorio de backups inmutables.

4.5.2. Servidores en nube (recursos mensuales orientativos)

Componente	Tipo de instancia	vCPU / RAM	Almacenamiento	Observaciones
Controlador de dominio réplica	m6i.xlarge (mínimo recomendado)	4 vCPU / 16 GB	100 GB gp3	Enlace VPN con on-prem. Se recomienda m6i.xlarge (4 vCPU / 16 GB) como mínimo para DC de producción.
Aplicaciones (gestor doc)	m6i.2xlarge	8 vCPU / 32 GB	200 GB gp3 + EFS	Escalable.
Base de datos principal	db.r6i.xlarge	4 vCPU / 32 GB	500 GB io2 (5.000 IOPS)	Multi-AZ (2 zonas).
IA-PROD (inferencia)	g4dn.xlarge (1× GPU T4)	4 vCPU + 1 GPU / 16 GB	100 GB	Solo para inferencia; no para entrenamiento.
SIEM + XDR	m6i.4xlarge	16 vCPU / 64 GB	1 TB	Logs 90 días en línea.

Nota sobre el controlador de dominio réplica: la instancia t3.large (2 vCPU / 8 GB) que figuraba en versiones anteriores de este documento resulta insuficiente para un controlador de dominio de producción integrado con Kerberos, LDAP y ZTNA. Se recomienda como mínimo m6i.xlarge (4 vCPU / 16 GB).

4.6. Requisitos de conectividad y ancho de banda

- Oficina principal: dos conexiones de fibra simétrica de al menos 1 Gbps cada una (proveedores diferentes), con redundancia automática mediante SD-WAN.
- Oficinas secundarias: fibra de 500 Mbps + 5G de respaldo.
- Acceso remoto: al menos 10 Mbps por usuario (recomendado 25 Mbps) para ZTNA con inspección.
- Enlace VPN a la nube: IPSec redundante, ancho de banda mínimo de 500 Mbps.

4.7. Alta disponibilidad y recuperación ante desastres (DRP)

- RPO (pérdida de datos máxima admisible): 15 minutos para datos críticos.
- RTO (tiempo de restauración máximo): 4 horas para sistemas esenciales.
- Replicación síncrona de bases de datos entre dos zonas en la nube.
- Snapshots horarios para recuperación ante borrados accidentales.
- Backups completos diarios a almacenamiento inmutable con retención de 30 días y retención anual para cumplimiento legal.
- Prueba del plan cada seis meses simulando fallos de red, corrupción de datos y ransomware.

4.8. Presupuesto orientativo anual (despacho mediano)

Concepto	Rango estimado (€/año)
Hardware on-premise (servidores, NAS, switches, NGFW)	40.000 - 70.000 (amortizado a 5 años)
Licencias de software base (SO, hipervisor)	10.000 - 20.000
Nube (computación + almacenamiento)	30.000 - 60.000
Seguridad (SIEM, XDR, DLP, PAM, ZTNA)	25.000 - 50.000
Herramientas de IA (licencias + GPU)	20.000 - 80.000
Conectividad (fibra redundante, SD-WAN)	12.000 - 24.000
Total estimado anual	137.000 - 304.000 €

5. Seguridad e integración de las herramientas de IA

5.1. Introducción y enfoque general

La incorporación de inteligencia artificial en un despacho de abogados no solo aporta eficiencia, sino que introduce nuevos vectores de ataque y riesgos de cumplimiento. Este punto desarrolla un marco de seguridad específico para IA que cubre: el ciclo de vida completo de los modelos y datos; la protección de los *prompts* y sus respuestas; la integración segura con el resto del sistema (ZTNA, DLP, SIEM); y el cumplimiento con el RIA y las obligaciones de secreto profesional. Principio fundamental: la IA no es un sistema autónomo, sino una herramienta bajo el control y supervisión humana. El abogado sigue siendo el responsable último.

5.2. Clasificación interna de las herramientas de IA

Con anterioridad a la implantación de cualquier sistema de IA, debe realizarse un inventario y clasificación conforme al RIA. Para uso interno se adopta la siguiente escala de tres niveles:

Nivel	Descripción	Ejemplos en un despacho	Medidas exigidas
Nivel 1: Riesgo mínimo	La IA asiste en tareas sin evaluar personas ni tomar decisiones autónomas.	Corrector ortográfico, resumidor de jurisprudencia, transcripción de audios.	Transparencia al usuario; logs de uso básicos.
Nivel 2: Riesgo limitado	La IA interactúa con personas o genera contenido que puede requerir supervisión.	Chatbot de atención al cliente; asistente de redacción de borradores.	Transparencia explícita; posibilidad de desactivar la IA; supervisión recomendada.
Nivel 3: Alto riesgo (solo si es imprescindible)	La IA evalúa personas o apoya decisiones legales que afectan a derechos.	Herramienta predictiva de éxito de litigios; selección de personal.	Todas las medidas del RIA para alto riesgo + supervisión humana efectiva + documentación técnica completa + exigencia de conformidad al proveedor.

Recomendación estratégica: el despacho debe limitarse a los niveles 1 y 2 durante al menos el primer año de implantación de IA.

5.3. Modelos de despliegue según la sensibilidad de los datos

5.3.1. Modo A: IA local aislada (recomendado para datos confidenciales)

El modelo se ejecuta completamente dentro de la infraestructura del despacho. No hay comunicación con el exterior. Aplicación: procesamiento de expedientes con información sensible.

5.3.2. Modo B: IA externa con no retención y garantías contractuales

Se utiliza un servicio de IA en la nube, pero el proveedor se compromete contractualmente a no retener ni usar los datos para entrenamiento; debe estar en la UE o disponer de SCC o certificación en el EU-US DPF. Aplicación: tareas no confidenciales.

5.3.3. Modo C: IA externa con datos anonimizados (solo pruebas)

Se envían datos previamente anonimizados. La anonimización debe ser técnicamente verificada para eliminar el riesgo de reidentificación. Aplicación exclusiva: entorno IA-SANDBOX.
Decisión para este proyecto: el 90 % de los casos de uso se resolverán con Modo A. El Modo B se reserva para funciones específicas de bajo riesgo, previa evaluación formal. El Modo C se limita al entorno de pruebas.

5.4. Seguridad del modelo y de los datos de entrenamiento o ajuste

- Los datos empleados para ajuste (fine-tuning) deben estar anonimizados o seudonimizados.
- El entorno de ajuste es IA-SANDBOX, sin acceso a internet y con controles de acceso estrictos.
- Los pesos del modelo (archivos binarios) se consideran activos críticos: se almacenan cifrados y se versionan.
- Antes de pasar a producción, el modelo ajustado se somete a pruebas de sesgo (bias) con conjuntos de datos equilibrados.
- El procedimiento de borrado seguro de pesos sigue el estándar NIST SP 800-88 Rev. 1 (2014), que establece el uso de borrado criptográfico (crypto erase) como método preferente para soportes de estado sólido y otros medios modernos, complementado con destrucción física cuando la información es especialmente sensible. El estándar DoD 5220.22-M está derogado y no debe emplearse como referencia.

5.5. Protección de los prompts y las respuestas

Medida	Implementación
Cifrado de extremo a extremo	TLS 1.3 con certificados propios entre el dispositivo del abogado y el servidor de IA.
No almacenamiento persistente de prompts	Los logs de IA guardan solo metadatos (quién, cuándo, modelo, duración), no el contenido del prompt. Si se requiere conservar el prompt con fines legales, se almacena cifrado con control de acceso específico.
Detección de fugas de datos (data leakage)	El DLP monitoriza el tráfico hacia servidores de IA no autorizados y detecta datos especialmente sensibles en los prompts. Si se detectan, se bloquea y se genera alerta.
Prevención de inyección de prompts (prompt injection)	Validaciones de entrada: el sistema filtra comandos maliciosos y limita la longitud del prompt. Se emplea un modelo de seguridad adicional para detectar intentos de evasión (jailbreak).

5.6. Integración con la arquitectura de seguridad general

5.6.1. ZTNA

- Los servidores de IA se ubican en el segmento IA-PROD.
- El acceso a la API o interfaz web de IA se realiza a través del ZTNA Gateway, con MFA.
- Las políticas de ZTNA pueden restringir el acceso a herramientas de IA según el rol del abogado.

5.6.2. DLP

- Salidas de IA: si una respuesta contiene datos personales y se intenta enviar a un dominio externo no autorizado, se bloquea.
- Entradas a IA: si un usuario pega una lista de clientes con datos sensibles en un prompt, el DLP puede advertir y exigir justificación.
- Copia de respuestas: si se copia al portapapeles el resultado de la IA y se pega en un documento no cifrado, se genera una alerta.

5.6.3. SIEM y SOC

Todos los eventos de seguridad relacionados con IA (intentos de acceso no autorizado, prompts bloqueados por DLP, cambios en la configuración de modelos) se envían al SIEM. El SOC tiene procedimientos específicos para incidentes de IA.

5.6.4. PAM para cuentas de servicio

Los agentes de IA autónomos tienen cuentas de servicio gestionadas por PAM: credenciales rotadas automáticamente cada 24 horas; permisos de mínimo privilegio; registro completo de actividad.

5.7. Ciclo de vida de la IA

5.7.1. Fase de desarrollo / adquisición

- Se exige certificación ISO 27001 o SOC 2 al proveedor externo, y un acuerdo de procesamiento que prohíba el uso de los datos del despacho para entrenar modelos.
- Si la IA es de desarrollo interno: pruebas de seguridad estáticas y dinámicas; análisis de dependencias.
- Análisis de riesgos específico (DPIA + evaluación de impacto en derechos fundamentales).

5.7.2. Fase de despliegue

- Despliegue inicial en IA-SANDBOX; paso a IA-PROD tras pruebas exitosas.
- Imágenes base hardened: SO minimizado, sin servicios innecesarios.
- Prueba de penetración específica sobre la interfaz de la IA.

5.7.3. Fase de operación continua

- Monitorización de métricas: precisión, sesgo, toxicidad (para LLMs).
- Actualizaciones y parches mensuales del software de IA, con validación previa en sandbox.

5.7.4. Fase de retirada

- Borrado seguro de los pesos del modelo conforme a NIST SP 800-88 Rev. 1: borrado criptográfico (crypto erase) como método preferente; destrucción física para información de máxima sensibilidad.
- Eliminación de todos los logs de interacción (contenido, si se almacenó).
- Supresión de las cuentas de servicio asociadas.
- Documentación de la retirada para auditorías.

5.8. Cumplimiento específico con el Reglamento de IA

Para sistemas de IA de alto riesgo, el sistema debe garantizar:

- Sistema de gestión de calidad por parte del proveedor: documentación de procesos de desarrollo, validación y monitorización post-comercialización (art. 17 RIA).
- Evaluación de conformidad: la realiza el proveedor, no el despacho como deployer. Para la mayoría de los sistemas del Anexo III puede ser autoevaluación interna; para sistemas biométricos específicos, evaluación por organismo notificado. El despacho exigirá la documentación técnica y la declaración de conformidad UE.
- Registro en la base de datos de la UE (art. 71 RIA), cuando resulte obligatorio para el tipo de sistema.
- Supervisión humana efectiva (art. 14 RIA): el RIA no impone un número mínimo de revisores; exige que la supervisión sea adecuada al riesgo concreto del sistema y que el supervisor tenga la formación y la capacidad de intervenir o detener el sistema.
- Ciberseguridad robusta: la arquitectura descrita en este documento cumple con los requisitos de "resiliencia, precisión y robustez" del art. 15 RIA.

5.9. Formación específica en IA

Módulo	Contenido	Duración	Frecuencia
Fundamentos de IA para abogados	Funcionamiento de LLMs, limitaciones, riesgos de alucinaciones.	2 horas	Al inicio; luego anual.
Seguridad y privacidad en el uso de IA	Prohibición de IAs públicas para datos confidenciales; reconocimiento de intentos de inyección.	1 hora	Anual.
Supervisión humana y validación de respuestas	Técnicas para revisar y corregir salidas de IA; responsabilidad legal del abogado.	2 horas	Anual.
Procedimiento ante incidentes de IA	Qué hacer si la IA genera una respuesta errónea o potencialmente dañina.	1 hora	Semestral.

5.10. Plan de respuesta a incidentes específico de IA

Escenario	Activación	Acciones inmediatas	Responsable
Fuga de datos a través de IA	Alerta DLP o queja de cliente.	1. Bloquear el acceso de ese usuario a la IA. 2. Solicitar al proveedor que elimine los datos. 3. Notificar al DPO. 4. Evaluar si constituye una brecha RGPD.	CISO + DPO
Modelo comprometido	Anomalía en sumas de verificación del modelo o comportamiento atípico.	1. Aislar el servidor IA-PROD. 2. Restaurar desde copia limpia. 3. Investigar el vector de ataque. 4. Notificar a los clientes si sus datos pudieron haberse visto expuestos.	CISO + TI
Alucinación grave con daño al cliente	Queja del cliente o detección interna.	1. Rectificar el consejo al cliente. 2. Revisar si el abogado omitió la supervisión obligatoria. 3. Acción disciplinaria si procede. 4. Reforzar las advertencias visibles del sistema de IA.	Socio director + CISO

5.11. Lista de verificación para cada nueva herramienta de IA

- ☐ Clasificación de riesgo según RIA (nivel 1, 2 o 3).
- ☐ DPIA completada y firmada por el DPO.
- ☐ Modo de despliegue (A, B o C) definido y justificado.
- ☐ El proveedor externo ha firmado el acuerdo de procesamiento de datos.
- ☐ Configuradas las políticas de DLP para esta herramienta.
- ☐ Acceso protegido por ZTNA con MFA.
- ☐ Procedimiento de supervisión humana establecido (niveles 2 y 3).
- ☐ Logs de interacción enviados al SIEM (solo metadatos, sin contenido sensible).
- ☐ Cuenta de servicio (si existe) gestionada por PAM.
- ☐ Formación específica impartida a los usuarios de esa herramienta.
- ☐ Prueba de penetración básica realizada sobre la interfaz.

6. Protección de datos, cifrado y secreto profesional

6.1. Política de cifrado

El cifrado es la primera línea de defensa frente a la exfiltración de datos. Se adopta el siguiente estándar:

- Cifrado en tránsito: TLS 1.3 para todas las comunicaciones internas y externas. Se prohíben TLS 1.0, TLS 1.1 y SSL 3.0.
- Cifrado en reposo: AES-256 para bases de datos (cifrado transparente TDE), almacenamiento de documentos (cifrado de volumen o de objeto) y copias de seguridad.
- Cifrado de endpoints: BitLocker (Windows) o FileVault (macOS) en todos los portátiles y equipos de escritorio corporativos.
- Correo electrónico: S/MIME o PGP para comunicaciones externas con clientes que manejen información especialmente sensible.
- Gestión de claves: todas las claves se gestionan a través del KMS. Las claves maestras se almacenan en HSMs (Hardware Security Modules). Se establece una política de rotación anual de claves y de revocación inmediata ante compromiso.

6.2. Gestión del ciclo de vida de las claves

El KMS centraliza el ciclo de vida de todas las claves criptográficas:

- Generación: mediante generadores de números aleatorios de hardware (HRNG) o HSMs certificados FIPS 140-2 nivel 2 o superior.
- Distribución: nunca en texto claro; siempre cifrada o mediante canales seguros fuera de banda.
- Rotación: anual para claves de cifrado de datos; inmediata en caso de compromiso o baja de personal con acceso.
- Custodia: las claves maestras se dividen mediante esquemas de umbral (p. ej., Shamir Secret Sharing) para evitar que una sola persona tenga acceso completo.
- Destrucción: registro auditado de la eliminación; los soportes que contenían claves se destruyen físicamente.

6.3. Secreto profesional y tecnología

El secreto profesional del abogado no es una obligación más de privacidad; es un derecho fundamental del justiciable y un deber deontológico irrenunciable. La tecnología debe estar al servicio del secreto, no en su detrimento. Las siguientes reglas son de obligado cumplimiento:

- Ningún dato de cliente en claro se almacena en servicios cloud no evaluados y aprobados expresamente por el despacho.
- Los sistemas de IA externos (Modo B) solo reciben datos previamente clasificados como no confidenciales o anonimizados.
- El despacho mantendrá un registro de los proveedores tecnológicos que tienen acceso a datos de clientes, con el contrato de confidencialidad y los acuerdos de procesamiento correspondientes.

- Ante cualquier requerimiento de autoridades para acceder a datos de clientes, el despacho invocará el secreto profesional y consultará al Consejo de Abogados competente antes de proporcionar información alguna.

7. Gobernanza, gestión de accesos y formación

7.1. Estructura de gobernanza

La ciberseguridad del despacho requiere una estructura de responsabilidad clara:

- Socio Director / Consejo de Administración: aprobación anual de la política de seguridad; supervisión del CISO; responsabilidad última ante reguladores.
- CISO (Chief Information Security Officer): puede ser interno o externalizado. Coordina la implantación técnica, dirige el SOC y reporta a la dirección. Reporta incidentes en los plazos regulatorios.
- DPO (Delegado de Protección de Datos): supervisa el cumplimiento RGPD/LOPDGDD; aprueba DPIAs; es el punto de contacto con la AEPD y otras autoridades de control.
- Responsable de TI: gestión operativa de la infraestructura, aplicación de parches, gestión del inventario de sistemas.
- "Propietario de IA" (AI Owner) por cada herramienta: abogado designado responsable de supervisar el uso correcto de cada sistema de IA, reportar alucinaciones y solicitar revisiones.

7.2. Gestión de accesos e identidades (IAM)

- Principio de mínimo privilegio: cada usuario tiene acceso únicamente a los recursos necesarios para su función. Revisión trimestral de permisos.
- Onboarding: al incorporarse un nuevo abogado, el responsable de TI activa sus credenciales conforme a su rol predefinido. La cuenta está desactivada hasta el primer día de trabajo. No se conceden privilegios administrativos a ningún usuario estándar.
- Offboarding: al producirse la baja (voluntaria o no), las credenciales se revocan en el mismo día hábil. El PAM gestiona las cuentas de servicio y las cuentas administrativas con acceso just-in-time.
- Contraseñas: longitud mínima de 16 caracteres para cuentas estándar; 20 caracteres para cuentas privilegiadas. Se recomienda el uso de un gestor de contraseñas corporativo.
- MFA: obligatoria para todos los accesos, sin excepción. Se priorizan claves físicas (FIDO2/WebAuthn) sobre OTP por aplicación; se evita el OTP por SMS como único factor.

7.3. Programa de formación

La formación es una medida técnica de seguridad, no una actividad administrativa. El programa comprende:

Módulo	Destinatarios	Duración	Frecuencia
Concienciación básica en ciberseguridad	Todo el personal	2 horas	Al inicio; actualización anual
Phishing y ingeniería social	Todo el personal	1 hora + simulacros trimestrales	Anual + simulacros
Zero Trust y uso correcto de ZTNA y MFA	Abogados y personal con acceso a sistemas	1 hora	Al inicio y tras cambios de sistema
Seguridad y privacidad en el uso de IA	Todos los usuarios de IA	3 horas	Antes del primer acceso a IA; anual
Gestión de incidentes y respuesta	CISO, TI, DPO, socios	4 horas	Anual
Obligaciones normativas (RGPD, RIA)	Socios y directores	2 horas	Anual o ante cambios normativos

8. Monitoreo continuo, SOC y respuesta a incidentes

8.1. Centro de Operaciones de Seguridad (SOC)

Un despacho mediano raramente puede mantener un SOC interno 24/7. La opción recomendada es un SOC gestionado (Managed SOC / MSSP) con las siguientes características mínimas:

- Monitorización 24 horas, 7 días a la semana, 365 días al año.
- Tiempo de detección y notificación al despacho ante incidente crítico: máximo 15 minutos desde la generación de la alerta en el SIEM.
- Personal del SOC con acceso al entorno (con MFA y sesiones grabadas bajo PAM) a los logs del SIEM y al XDR. Sin acceso directo a los expedientes legales.
- Contrato que incluya SLA de respuesta, cláusulas de confidencialidad equivalentes al secreto profesional y obligación de notificación en caso de brecha.

8.2. Arquitectura de monitorización

- SIEM: centralización y correlación de todos los logs (firewalls, ZTNA, servidores, aplicaciones, agentes XDR, PAM, DLP). Retención mínima de 12 meses en línea; retención de 5 años en almacenamiento frío para cumplimiento legal.
- XDR/EDR: detección de comportamientos anómalos en endpoints (ejecución de procesos inesperados, conexiones a C2, cifrado masivo de ficheros).
- Integridad de logs: los registros deben ser inmutables (log forwarding a sistema independiente; firma digital o hash de cadena). Esto es obligatorio para que los logs sean admisibles como prueba ante un regulador.
- Correlación de eventos de IA: reglas específicas en el SIEM para detectar anomalías en el uso de IA (volumen inusual de prompts, accesos fuera de horario a herramientas de IA, intentos de inyección).

8.3. Plan de respuesta a incidentes (PRI)

El PRI debe estar documentado, aprobado por la dirección y ensayado al menos una vez al año mediante simulacros (tabla-top exercise o simulación técnica completa). Las fases del PRI son:

- Preparación: roles y responsabilidades asignados; contactos de emergencia actualizados; herramientas de respuesta disponibles; contratos con proveedores de respuesta a incidentes externos.
- Detección e identificación: el SIEM genera la alerta; el SOC valida y clasifica el incidente según severidad (P1-P4).
- Contención: aislamiento del sistema o usuario afectado (el SOAR puede hacerlo automáticamente para P1); bloqueo de accesos comprometidos.
- Erradicación: eliminación de la causa raíz (malware, credenciales comprometidas, vulnerabilidad explotada).
- Recuperación: restauración de servicios desde backups inmutables verificados; validación del entorno antes de volver a producción.

- Notificación regulatoria: el DPO determina si la brecha requiere notificación a la AEPD (72 horas, art. 33 RGPD) y/o a los afectados (art. 34 RGPD). Se documentan todas las decisiones.
- Lecciones aprendidas: revisión post-incidente en un plazo máximo de 15 días; actualización del PRI.

9. Plan de implementación por fases y métricas de éxito

9.1. Visión general del plan

La implantación del sistema completo se divide en cuatro fases secuenciales, con una duración total estimada de 18 meses para un despacho mediano. Cada fase tiene entregables verificables y métricas de éxito medibles.

9.2. Fases de implementación

Fase	Período	Objetivos principales	Entregables clave
Fase 0: Diagnóstico y planificación	Meses 1-2	Inventario de sistemas y datos; análisis de riesgos; selección de proveedores; designación de DPO y CISO.	Informe de diagnóstico; análisis de riesgos; hoja de ruta aprobada por la dirección.
Fase 1: Fundamentos de identidad y segmentación	Meses 3-6	Implantación de IdP y MFA; despliegue de ZTNA piloto (5 usuarios); microsegmentación inicial; primer nivel de logs en SIEM.	ZTNA operativo para grupo piloto; 5 segmentos de red documentados; SIEM recibiendo logs básicos.
Fase 2: Seguridad completa y cumplimiento	Meses 7-12	ZTNA para todos los usuarios; DLP; PAM; KMS; DPIAs completadas; plan de continuidad de negocio.	Certificación de cumplimiento RGPD interna; NYDFS readiness assessment; primera prueba de penetración.
Fase 3: IA y operación avanzada	Meses 13-18	Despliegue de primera herramienta de IA (nivel 1) en IA-PROD; SOC gestionado activo; simulacro de incidente; programa de formación completo.	Primera herramienta de IA en producción con checklist completo; SOC 24/7 activo; primer simulacro documentado.

9.3. Métricas de éxito

Métrica	Objetivo	Frecuencia de medición
Cobertura MFA	100 % de los usuarios con MFA activa	Mensual
Tiempo de detección de incidentes (MTTD)	< 30 minutos para incidentes P1	Por incidente
Tiempo de respuesta (MTTR)	< 4 horas para P1; < 24 horas para P2	Por incidente
Cobertura de parches críticos	100 % en 48 horas desde publicación	Semanal
Tasa de completitud de formación	> 95 % del personal con formación al día	Trimestral
DPIAs completadas antes de implantar IA	100 %	Por implantación
Simulacros de incidente realizados	Mínimo 1 al año con informe post-ejercicio	Anual
Backups verificados y restaurables	100 % de los backups probados en los últimos 30 días	Mensual

9.4. Recursos necesarios

- CISO / Responsable de seguridad: al menos a tiempo parcial desde el mes 1; a tiempo completo recomendado desde el mes 7.
- DPO: externo durante las fases 0-2; valorar internalización a partir de la fase 3.
- Responsable de TI: tiempo completo durante todo el proyecto.
- Proveedor de SOC gestionado: contrato desde el inicio de la fase 2.
- Proveedor de auditoría / pentesting externo: contratado una vez al año.

Nota final: esta guía debe revisarse como mínimo de forma anual y siempre que se produzca un cambio normativo relevante, una brecha de seguridad significativa o la incorporación de una nueva herramienta de IA. El entorno de ciberseguridad y el marco regulatorio de la IA evolucionan con rapidez; la vigilancia continuada no es opcional.