

DIMENSIONES DE LA SEGURIDAD

Ian Santillán (CEO) - Miguel Zamora (CISO) - Aitor Jiang (CDO)

Borja Rodriguez (CIO) - Alejandro Zhou - Iván Moràn



ÍNDICE

Índice.....	1
Resumen ejecutivo.....	3
1. Introducción.....	3
1.1 Objetivo	3
1.2 Alcance y limitaciones.....	3
1.3 Metodología	3
2. Contexto del caso.....	4
2.1 Empresa y activos de información	4
2.2 Uso de IA generativa y ausencia de normas	4
3. Descripción del incidente	4
3.1 Narrativa del incidente	4
3.2 Línea temporal del incidente	4
3.3 Punto de la brecha y datos comprometidos	5
4. Análisis causal	5
4.1 Causas inmediatas	5
4.2 Causas raíz (personas, procesos y tecnología)	5
4.3 Controles que fallaron o no existían	5
5. Impacto y evaluación de riesgos	6
5.1 Impacto económico	6
5.2 Impacto reputacional y operacional	6
5.3 Riesgos para el cliente	6
5.4 Matriz simplificada de riesgos	6
6. Marco jurídico aplicable.....	7
6.1 Confidencialidad contractual (NDA).....	7
6.2 Protección de datos: RGPD y LOPDGDD (cuando proceda)	7
6.3 Secreto empresarial	7
6.4 Propiedad intelectual (código fuente)	7
6.5 Responsabilidad y deber de diligencia	8
7. Plan de acción	8
7.1 Gestión del incidente (respuesta).....	8
7.2 Comunicación y notificación	8
7.3 Acciones técnicas (corto plazo).....	8
7.4 Acciones organizativas (medio plazo).....	9
7.5 Cronograma y responsables (RACI simplificado)	9
8. Propuesta de política interna de uso de IA	10
8.1 Principios	10
8.2 Datos prohibidos en herramientas públicas	10
8.3 Herramientas permitidas	10
8.4 Procedimiento de autorización y revisión.....	10
9. Conclusiones y lecciones aprendidas	10
10. Referencias normativas.....	11
Anexo A. Checklist rápido para desarrolladores	11

RESUMEN EJECUTIVO

Este documento analiza un incidente de seguridad ocurrido en ATHENEA S.A. durante el desarrollo de un proyecto para un cliente. Ante la presión por plazos y un problema técnico complejo, un desarrollador utilizó una herramienta de IA generativa pública e introdujo fragmentos reales de código fuente y documentación técnica del cliente. Posteriormente, el cliente detectó fragmentos muy similares publicados como ejemplos accesibles públicamente, lo que activó la investigación interna.

El informe identifica el punto de la brecha, la información potencialmente comprometida, las decisiones incorrectas y el marco jurídico aplicable (confidencialidad contractual, protección de datos cuando proceda, secreto empresarial y propiedad intelectual). Finalmente se propone un plan de acción con medidas organizativas y técnicas para la contención, corrección y prevención de incidentes similares, incluyendo una política interna de uso de IA.

1. INTRODUCCIÓN

ATHENEA S.A. es una empresa tecnológica dedicada al desarrollo de software a medida para terceros. Su operativa habitual implica el acceso a repositorios de código, documentación técnica, especificaciones funcionales y, en ocasiones, datos de carácter personal relacionados con usuarios o empleados del cliente. Por ello, la confidencialidad y la seguridad de la información constituyen elementos críticos del servicio.

1.1 Objetivo

El objetivo del presente proyecto es documentar el incidente, evaluar su impacto y analizar el marco legal afectado, proponiendo un plan de acción para responder al incidente y reducir la probabilidad de recurrencia.

1.2 Alcance y limitaciones

El análisis se centra en la gestión interna del incidente, los activos de información implicados y las medidas correctivas y preventivas. No se realiza un peritaje forense completo ni se atribuye la filtración a un proveedor concreto; el documento trabaja con la información disponible en el caso práctico y con hipótesis razonables de riesgo.

1.3 Metodología

La metodología aplicada combina:

1. Reconstrucción de la línea temporal
2. Identificación de activos y amenazas
3. Análisis causal (personas, procesos y tecnología)
4. Revisión de obligaciones legales y contractuales
5. Propuesta de medidas alineadas con buenas prácticas de gestión de incidentes y seguridad de la información.

2. CONTEXTO DEL CASO

2.1 Empresa y activos de información

ATHENEA desarrolla soluciones internas críticas para sus clientes. Los principales activos de información relacionados con el caso son:

- Código fuente del cliente (repositorios, ramas de desarrollo, librerías internas).
- Documentación técnica interna procedimientos, credenciales documentadas).
- Documentación funcional y de negocio (requisitos, procesos, diagramas).

2.2 Uso de IA generativa y ausencia de normas

En la empresa se debate el uso de IA generativa para aumentar la productividad, pero no existen normas claras. Esta ausencia de gobernanza provoca riesgos recurrentes: uso de herramientas no autorizadas, envío de información sensible a terceros, y falta de trazabilidad sobre qué datos se comparten y con qué finalidad.

3. DESCRIPCIÓN DEL INCIDENTE

3.1 Narrativa del incidente

Durante un proyecto importante, el equipo se enfrenta a un retraso acumulado y a presión por parte del cliente para obtener resultados. Un desarrollador se encuentra con un problema técnico complejo. Con el fin de acelerar la resolución, decide usar una herramienta de IA generativa pública. Para obtener una respuesta más precisa, copia y pega fragmentos reales de código y documentación del cliente. Semanas después, el cliente detecta fragmentos muy similares publicados como ejemplos accesibles públicamente, lo que sugiere una pérdida de control sobre la información compartida.

3.2 Línea temporal del incidente

Paso	Qué ocurre
1	Aparece una necesidad o presión por el tiempo
2	Se decide usar una IA generativa pública para acelerar la resolución técnica.
3	Se introduce información real y sensible
4	La información sale del control de la empresa (riesgo de almacenamiento/redistribución externa).
5	Se detecta el posible incidente.

3.3 Punto de la brecha y datos comprometidos

La brecha de seguridad se materializa en el paso 3: la introducción de información real y sensible en un servicio externo no autorizado. A partir de ese momento, ATHENEA pierde el control sobre la confidencialidad del contenido compartido. La información potencialmente comprometida incluye fragmentos de código, documentación técnica interna y cualquier dato derivado que permita inferir lógica de negocio, arquitectura o vulnerabilidades.

4. ANÁLISIS CAUSAL

4.1 Causas inmediatas

- Uso de una IA generativa pública sin autorización ni evaluación de riesgos.
- Compartición de información confidencial del cliente (código y documentación).
- Presión por plazos y falta de alternativas internas para soporte técnico.

4.2 Causas raíz (personas, procesos y tecnología)

El incidente no se explica solo por el error individual; existe una combinación de causas estructurales:

- Personas: falta de formación específica sobre uso seguro de IA, confidencialidad y clasificación de la información.
- Procesos: ausencia de una política de uso de IA y de un procedimiento de escalado cuando aparece un bloqueo técnico.
- Gobernanza: falta de supervisión y control de avances que permita detectar bloqueos antes de que se conviertan en decisiones de riesgo.
- Tecnología: inexistencia de herramientas corporativas (IA privada, repositorios de conocimiento internos) que reduzcan la necesidad de recurrir a servicios públicos.

4.3 Controles que fallaron o no existían

- Control preventivo: política y formación sobre tratamiento de información confidencial (no implementados).

- Control preventivo: herramientas autorizadas para asistencia al desarrollo (no disponibles o no definidas).
- Control detectivo: monitoreo de fugas de código (p. ej., búsquedas de similitud, DLP) (no evidenciado en el caso).
- Control correctivo: procedimiento de respuesta a incidentes (no formalizado o no activado a tiempo).

5. IMPACTO Y EVALUACIÓN DE RIESGOS

5.1 Impacto económico

- Coste de contención y análisis (tiempo del equipo técnico, legal y de seguridad).
- Posibles penalizaciones contractuales e indemnizaciones al cliente por incumplimiento de confidencialidad.
- Coste de reescritura o refactorización del código comprometido para evitar reutilización pública.
- Posibles sanciones administrativas si existe afectación de datos personales y procede notificación a la autoridad competente.

5.2 Impacto reputacional y operacional

- Pérdida de confianza del cliente y deterioro de la relación comercial.
- Riesgo de rescisión de contrato o reducción de proyectos futuros.
- Afectación al clima interno y a la productividad por interrupciones derivadas de la gestión del incidente.

5.3 Riesgos para el cliente

- Exposición de lógica de negocio y arquitectura interna.
- Posible ventaja competitiva para terceros si el código refleja procedimientos propios.
- Incremento del riesgo de ataques (si el material filtrado facilita análisis de vulnerabilidades).

5.4 Matriz simplificada de riesgos

Riesgo	Probabilidad	Impacto	Nivel	Tratamiento propuesto
Filtración adicional por uso repetido de IA pública	Media	Alto	Alto	Política IA + herramientas corporativas + formación
Reclamación contractual por incumplimiento de NDA	Media	Alto	Alto	Negociación, transparencia, plan correctivo, evidencia de medidas
Sanción por protección de datos (si hay datos personales)	Baja/Media	Alto	Medio/Alto	Evaluación DPO, registro brecha, notificación si aplica
Pérdida de cliente y daño reputacional	Media	Alto	Alto	Comunicación proactiva, remediación técnica, mejora de gobernanza
Uso del código filtrado por terceros	Baja	Alto	Medio	Reescritura de componentes críticos, control de exposición pública, vigilancia

6. MARCO JURÍDICO APLICABLE

6.1 Confidencialidad contractual

Los proyectos de ATHENEA suelen estar sujetos a acuerdos de confidencialidad (NDA) o cláusulas de confidencialidad en el contrato principal. La divulgación no autorizada de código y documentación puede constituir un incumplimiento contractual, habilitando al cliente a exigir la reparación del daño, aplicar penalizaciones pactadas o resolver el contrato, según lo establecido en el acuerdo.

6.2 Protección de datos: RGPD y LOPDGDD

Si el material compartido incluyera datos personales (por ejemplo, identificadores, datos de empleados/usuarios, registros, logs o información vinculada a personas), resultarían aplicables el RGPD y la LOPDGDD. En ese supuesto, ATHENEA debe garantizar medidas técnicas y organizativas apropiadas para proteger los datos (art. 32 RGPD) y gestionar la brecha conforme a los artículos sobre notificación a la autoridad de control y, en su caso, a los interesados (arts. 33 y 34 RGPD). Incluso cuando no haya datos personales, el incidente puede ser relevante en términos de seguridad de la información y obligaciones contractuales.

6.3 Secreto empresarial

El código fuente y la documentación técnica pueden constituir secretos empresariales cuando aportan valor por ser secretos y la empresa adopta medidas razonables para mantener su

confidencialidad. La divulgación no autorizada puede generar acciones civiles de protección del secreto empresarial, especialmente si un tercero obtiene, utiliza o divulga esa información sin consentimiento.

6.4 Propiedad intelectual

El código fuente suele estar protegido por derechos de propiedad intelectual como obra. Su publicación sin autorización puede afectar a los derechos del titular (habitualmente el cliente, según contrato). Además, la exposición pública puede complicar la estrategia de licenciamiento, reutilización y protección del software.

6.5 Responsabilidad y deber de diligencia

Desde la perspectiva de responsabilidad, la empresa debe actuar con diligencia para prevenir incidentes previsibles (políticas, controles y formación) y responder adecuadamente cuando ocurren. El empleado puede incurrir en responsabilidad disciplinaria interna. En escenarios agravados (por ejemplo, divulgación intencional o afectación grave), podrían existir implicaciones adicionales que requerirían valoración jurídica específica.

7. PLAN DE ACCIÓN

7.1 Gestión del incidente

Se recomienda estructurar la respuesta siguiendo fases reconocibles: identificación, contención, erradicación, recuperación y lecciones aprendidas.

1. Identificación: confirmar alcance, qué fragmentos se compartieron, dónde pudieron quedar almacenados y si incluyen datos personales.
2. Contención: detener el uso de IA pública para el proyecto, revocar accesos si procede, y limitar la exposición de repositorios/documentación.
3. Erradicación: eliminar o minimizar la causa (bloquear herramientas no autorizadas, aplicar DLP, definir herramientas corporativas).
4. Recuperación: reescribir componentes críticos si han quedado expuestos, rotar secretos/credenciales y validar el producto final.
5. Lecciones aprendidas: actualizar políticas, formación y controles; documentar el incidente y acciones.

7.2 Comunicación y notificación

- Comunicación con el cliente: informar de forma transparente, indicando hechos conocidos, medidas inmediatas y plan de remediación.
- Comunicación interna: implicar a dirección, seguridad (CISO), datos (CDO) y, si existe, delegado de protección de datos (DPO).

- Notificación a la autoridad: si hay datos personales y se confirma una brecha, valorar notificación a la AEPD dentro de plazos legales.
- Gestión de evidencias: registrar decisiones y acciones para poder demostrar diligencia.

7.3 Acciones técnicas

- Inventariar exactamente los fragmentos compartidos y evaluar si contienen secretos, claves, endpoints o datos personales.
- Realizar búsquedas de similitud en repositorios públicos (cuando sea posible) para estimar exposición.
- Rotar credenciales y secretos potencialmente expuestos (tokens, API keys, contraseñas documentadas).
- Revisar y reescribir módulos críticos o distintivos (si el cliente lo requiere).
- Aplicar controles de salida (DLP) o reglas de proxy para bloquear envíos a servicios no autorizados.

7.4 Acciones organizativas

- Definir y aprobar una Política de Uso de IA.
- Formación obligatoria para equipos técnicos (confidencialidad, clasificación de información, uso seguro de IA).
- Proceso de escalado: cuando un desarrollador se bloquee, solicitar apoyo interno en lugar de usar herramientas públicas.
- Revisión por pares y supervisión del avance (para detectar retrasos y reducir decisiones impulsivas).
- Cláusulas contractuales claras sobre herramientas y subprocesadores (cuando aplique).

7.5 Cronograma y responsables

Actividad	Responsable	Aprobador	Consultado	Informado
Detener uso de IA pública y comunicar directrices inmediatas	CISO	CEO	CIO, CDO	Equipo técnico
Inventario de datos compartidos y evaluación de impacto	CISO/CIO	CEO	CDO, Legal	Cliente
Evaluación legal (NDA/RGPD si aplica) y plan de comunicación	Legal/CDO	CEO	CISO	Cliente
Medidas técnicas de contención (DLP/bloqueos/rotación secretos)	CIO	CISO	Equipo DevOps	Equipo técnico
Definir política IA y plan de formación	CDO	CEO	CISO, CIO	Toda la empresa

8. PROPUESTA DE POLÍTICA INTERNA DE USO DE IA

8.1 Principios

- Necesidad y proporcionalidad: usar IA solo cuando aporte valor real.
- Minimización de datos: no compartir información real del cliente si no es imprescindible.
- Confidencialidad por diseño: priorizar herramientas corporativas o entornos privados.
- Trazabilidad: registrar qué herramientas se usan y con qué finalidad.

8.2 Datos prohibidos en herramientas públicas

Queda prohibido introducir en herramientas públicas o no autorizadas:

- Código fuente del cliente o de ATHENEA no publicado.
- Documentación interna, diagramas, especificaciones o tickets con información sensible.
- Credenciales, secretos, tokens, llaves criptográficas o configuraciones de infraestructura.
- Datos personales (nombres, emails, identificadores, logs con usuarios) y cualquier dato sujeto a confidencialidad.

8.3 Herramientas permitidas

- IA corporativa o privada con contrato, garantías de confidencialidad y configuración adecuada.
- Modelos locales o en entorno controlado para pruebas (sin datos reales).
- Repositorios de conocimiento internos

8.4 Procedimiento de autorización y revisión

6. Solicitud: el empleado solicita usar una herramienta nueva indicando finalidad y tipo de datos.
7. Evaluación: CISO/CIO/CDO evalúan riesgos, proveedor, ubicación del tratamiento y medidas.
8. Aprobación: se autoriza por escrito y se documenta el alcance.
9. Revisión periódica: auditoría trimestral de herramientas, logs y cumplimiento.

9. CONCLUSIONES Y LECCIONES APRENDIDAS

El incidente evidencia que el uso de IA generativa sin gobernanza puede derivar en pérdida de control sobre información crítica. La brecha se produce cuando se comparten datos reales y sensibles con un servicio externo no autorizado. Más allá del error individual, la causa raíz reside en la ausencia de políticas, formación y herramientas corporativas que permitan resolver problemas técnicos sin exponer información del cliente.

La respuesta recomendada combina acciones inmediatas de contención y evaluación (técnicas y legales), una comunicación transparente con el cliente, y la implantación de medidas preventivas (política de IA, formación, controles y supervisión). Con ello, ATHENEA puede reducir el riesgo de recurrencia y demostrar diligencia en la protección de la información que le confían sus clientes.

10. REFERENCIAS NORMATIVAS

- Reglamento (UE) 2016/679 (RGPD).
- Ley Orgánica 3/2018, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD).
- Ley 1/2019, de Secretos Empresariales.
- Texto Refundido de la Ley de Propiedad Intelectual (TR-LPI).
- Contrato y acuerdos de confidencialidad (NDA) firmados con el cliente (según proceda).