PLANTILLA P16 · LINAJE DE DATOS Y MODELOS

Linaje de datos y modelos

Registra de dónde proceden los datos, cómo se transforman y qué modelos y versiones intervienen, para poder explicar, reproducir y revertir cualquier resultado del sistema.

Fernando García Varela Versión 0.1 (borrador de trabajo) 16-09-2026 Borrador para revisión

EN ESTE DOCUMENTO

01 Instrucciones de uso

02 Identificación

03 Fuentes de datos

04 Conjuntos de datos y transformaciones

05 Modelos y componentes

06 Recorrido de extremo a extremo

07 Reproducibilidad y trazabilidad de resultados

08 Tratamiento en caso de reversión o retirada

09 Criterios de calidad

10 Aprobación y verificación

11 Control de versiones

Aviso legal y exención de responsabilidad. SEVEN-G es un marco metodológico de referencia que se ofrece «tal cual» y con fines exclusivamente informativos. No constituye asesoramiento jurídico, regulatorio, financiero ni profesional, ni garantiza el cumplimiento de ninguna norma. Las referencias a regulación general (como el Reglamento Europeo de IA, el RGPD, DORA o NIS2), a normas técnicas y a regulación específica de cada sector o jurisdicción pueden ser incompletas, no aplicar a un caso concreto o quedar desactualizadas por cambios normativos, interpretaciones o criterios de las autoridades posteriores a su fecha de consulta. **Cada organización que use SEVEN-G es la única responsable de identificar la normativa que le aplica, verificar su vigencia y certificar su propio cumplimiento regulatorio**, con el asesoramiento cualificado que corresponda. Esta metodología es una ayuda genérica y gratuita, compartida con la comunidad para que nadie tenga que empezar desde cero; cada persona u organización puede y debe adaptarla a su propio uso. No debe entenderse que sus partes jurídicamente sensibles hayan sido revisadas por una asesoría jurídica: esas revisiones, para cada empresa o sector, son responsabilidad última de la empresa, el consultor o la organización que la use. Aunque se procura mantenerla al día, alguna norma puede haber cambiado sin que se recoja aquí. En la máxima medida permitida por la ley, el autor no asume responsabilidad alguna por los efectos de su aplicación en ninguna organización ni por su aplicabilidad completa. La metodología no otorga certificación de ningún tipo. El autor no asume responsabilidad alguna por el uso que se haga de este contenido ni por las decisiones que se adopten con él. Los datos, cifras, compañías y casos de los ejemplos son ficticios o ilustrativos.

01

Instrucciones de uso

Aspecto	Indicación
Cuándo se usa	Fase 4 (Diseño de la solución), antes de solicitar G4. Se actualiza en la fase 5 con las versiones realmente desplegadas y en la fase 6 con cada cambio de datos, modelo, instrucciones o base de conocimiento.
Quién la rellena	Responsable técnico de IA, con los responsables de los datos de origen.

Aspecto	Indicación
Quién la verifica	Oficina de IA en Lite; auditor de IA en Enterprise (01 §7.5). El responsable de riesgos revisa la base legal y las restricciones de uso.
En qué *gate* se revisa	G4. Se comprueba de nuevo en G5 (versiones desplegadas) y en cada revisión de continuidad R6.
Lite frente a Enterprise	Lite puede omitir los campos marcados **(Enterprise)**. Si el sistema está clasificado como de alto riesgo, se completan todos los campos con independencia de la intensidad.
Relación	P05 Ficha de inventario · P11 Clasificación regulatoria · P15 Registro de arquitectura · P19 Plan de reversión · P27 Registro de incidentes y cambios · P30 Decisión de escalado o retirada · P64 Ficha de conjunto de datos e inventario de fuentes de conocimiento · documentos 51 y 53.

Reglas de cumplimentación:

- Cada fuente, conjunto de datos, modelo y componente tiene **identificador, versión y responsable**. No se admiten descripciones genéricas como "datos de clientes": se identifica la tabla, el fichero, la API o el servicio.
- Los modelos de terceros, incluidos los de propósito general consumidos por API, se registran con proveedor, nombre comercial y versión exacta.
- En IA generativa y agentes, las **instrucciones del sistema, las bases de conocimiento y las herramientas conectadas** son componentes versionados, igual que un modelo.
- Todo cambio de versión en producción se registra como cambio en P27 y actualiza esta plantilla.
- "Sin dato" no es cero: si un campo no se conoce, se indica *Sin dato* y el motivo.

Esta plantilla no constituye asesoramiento jurídico. La base legal del tratamiento debe validarla protección de datos.

02

Identificación

Campo	Contenido	Guía
Código del caso de uso		IA-AAAA-NNN del registro de casos de uso (T01).
Nombre del caso de uso		Nombre corto y comprensible.
Sistema(s) de IA		Identificador en el inventario de sistemas de IA (T02, P05).

Campo	Contenido	Guía
Intensidad		Lite · Enterprise (P04).
Clasificación regulatoria		Prohibido · Alto riesgo · Obligaciones de transparencia · Riesgo mínimo · Fuera de ámbito · Pendiente de clasificar (P11).
Versión del documento		0.1, 0.2…; 1.0 al aprobarse. Cada cambio en producción incrementa la versión.
Autor		Rol y nombre.
Fecha		DD-MM-AAAA.
Estado del documento		Borrador · Presentado para verificación · Verificado · Aprobado · Sustituido.

03

Fuentes de datos

Una fila por fuente. **Tipo de datos:** Personales · Categorías especiales · Confidencial · Interno · Público · Sintético. **Base legal (RGPD art. 6):** Consentimiento · Contrato · Obligación legal · Intereses vitales · Interés público · Interés legítimo · No aplica (sin datos personales).

ID	Fuente y sistema de origen	Responsable del dato	Tipo de datos	Base legal	Restricciones de uso	Frecuencia de actualización	Calidad verificada (Sí · No · Parcial)
F-01 ***(ejemplo ilustrativo)***	Histórico de incidencias de mantenimiento, tabla de órdenes de trabajo del sistema de gestión de activos	Jefatura de mantenimiento	Interno	No aplica	Solo uso interno; no se comparte con proveedores	Diaria	Parcial: 4 % de registros sin código de avería

Campo complementario	Contenido	Guía
Datos de terceros o adquiridos		Proveedor, licencia y condiciones de uso para IA.

Campo complementario	Contenido	Guía
Transferencias fuera del Espacio Económico Europeo (Enterprise)		Destino, garantía aplicada y referencia a la evaluación de proveedor (P14).
Evaluación de impacto en protección de datos		Realizada · No requerida (con justificación) · Pendiente. Enlace a P11.

04

Conjuntos de datos y transformaciones

Uso: Entrenamiento · Ajuste · Validación · Prueba · Base de conocimiento (recuperación) · Inferencia · Monitorización.

ID conjunto	Uso	Fuentes (ID)	Transformaciones principales	Versión o huella	Periodo cubierto	Volumen	Controles de sesgo y representatividad (Enterprise)
CD-01 ***(ejemplo ilustrativo)***	Entrenamiento	F-01, F-02	Eliminación de duplicados; imputación de fechas; seudonimización del técnico	v3 · huella registrada en el repositorio de datos	01-2023 a 12-2025	180.000 registros	Distribución por planta y tipo de activo comparada con la población

05

Modelos y componentes

Tipo de modelo: Propio entrenado · Propio ajustado sobre modelo base · De terceros por API · De terceros desplegado en infraestructura propia · Modelo de propósito general · Reglas (no es IA).
Estado: En desarrollo · Validado · En producción · Retirado.

ID modelo	Tipo	Proveedor y versión exacta	Conjuntos usados	Finalidad	Métricas de referencia	Limitaciones conocidas	Estado
M-01 ***(ejemplo ilustrativo)***	Propio entrenado	Interno · v1.4	CD-01, CD-02	Estimar probabilidad de avería a 30 días	Métrica de discriminación y calibración registradas en P22	Bajo rendimiento en activos con menos de 6 meses de histórico	Validado

Componentes de IA generativa y agentes (si aplica):

Componente	Identificador y versión	Responsable	Ubicación	Qué exige un cambio
Instrucciones del sistema				Prueba de regresión y registro en P27.
Base de conocimiento				Revisión de contenido, fecha de corte y permisos de acceso.
Herramientas conectadas del agente				Revisión de permisos en P18.
Parámetros de generación (Enterprise)				Prueba de regresión.
Modelo de respaldo por coste (documento 52 §10.3), si se prevé				Proveedor y versión exacta; evaluado antes de G5 con el conjunto de referencia y las pruebas de sesgo del principal. Incorporarlo o cambiar sus condiciones es un cambio (documento 52 §6.4).

06

Recorrido de extremo a extremo

Describe, en orden, cómo un dato de origen se convierte en un resultado del sistema. Una fila por paso; sin diagramas en texto.

Paso	Entrada	Proceso	Salida	Registro que lo evidencia
1 ***(ejemplo ilustrativo)***	F-01	Extracción nocturna	Copia bruta fechada	Registro del proceso de carga
2 ***(ejemplo ilustrativo)***	Copia bruta	Transformaciones de CD-01	CD-01 v3	Huella del conjunto
3 ***(ejemplo ilustrativo)***	Datos del día	Inferencia con M-01 v1.4	Lista priorizada de activos	Registro de inferencias con versión del modelo

07

Reproducibilidad y trazabilidad de resultados

Campo	Contenido	Guía
¿Puede vincularse cada resultado a la versión de modelo, datos e instrucciones?		Sí · Parcialmente · No. Si no, es un hallazgo para G4.
Registros automáticos que se conservan		Qué se registra (entrada, salida, versión, fecha, usuario). En sistemas de alto riesgo, alinear con las obligaciones de registro del Reglamento Europeo de IA (documento 34).
Plazo de conservación de registros		Plazo fijado por la compañía, no inferior al que exija la regulación aplicable.
Capacidad de reentrenar o regenerar una versión anterior (Enterprise)		Qué se guarda para reconstruir una versión (código, datos, parámetros).
Referencia para detectar deriva		Conjunto y periodo que sirve de referencia para la monitorización (P25). En IA generativa y agentes, referencia de uso: temas, intenciones y alcance validados en G5 (documento 52 §4.2.6).

08

Tratamiento en caso de reversión o retirada

Campo	Contenido	Guía
Versión estable anterior disponible		Identificador del modelo y componentes a los que se vuelve (P19).
Datos que deben conservarse		Por obligación legal, auditoría o defensa jurídica.
Datos que deben borrarse o anonimizarse		Incluye copias en proveedores.
Obligaciones del proveedor al terminar (Enterprise)		Devolución, borrado certificado, no reutilización para entrenamiento. Se detalla en P30.

09

Criterios de calidad

El verificador comprueba, como mínimo, lo siguiente. Los criterios formales de G4 están en el documento 21 y la lista de verificación LV-G4 en el documento 22.

#	Comprobación	Estado (Cumple · No cumple · No aplica · Pendiente)
1	Todas las fuentes tienen responsable, tipo de datos y base legal cuando hay datos personales.	
2	Cada conjunto de datos tiene versión o huella y periodo cubierto.	
3	Cada modelo y componente tiene versión exacta, incluidos los de terceros.	
4	El recorrido de extremo a extremo permite vincular un resultado a sus versiones.	
5	Las limitaciones conocidas están declaradas y trasladadas a P17 y P24.	
6	Existe una versión anterior identificada para la reversión.	
7	La plantilla existía y estaba verificada antes de la solicitud de G4.	

10

Aprobación y verificación

Función	Rol	Nombre	Fecha	Firma o referencia
Autor	Responsable técnico de IA			
Revisión de base legal	Protección de datos o responsable de riesgos de IA			
Verificador	Oficina de IA (Lite) · Auditor de IA (Enterprise)			
Decisor del *gate*	Según 01 §7.5 (registrado en P29)			

Separación de funciones: el verificador no forma parte del equipo que construye y no es la misma persona que el autor (01 §7.4 y §8.2).

11

Control de versiones

Versión	Fecha	Cambios
0.1	16-09-2026	Primera versión de la plantilla.
0.1	18-09-2026	Modelo de respaldo por coste en los componentes y referencia de uso para la deriva en IA generativa y agentes (documento 52 §4.2.6 y §10.3).

Fuentes externas citadas

Enlaces a la fuente oficial o primaria, comprobados antes de publicar. Verifique siempre la versión vigente en la fecha de uso.

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE

(Reglamento general de protección de datos) · https://eur-lex.europa.eu/eli/reg/2016/679/oj/spa · Parlamento Europeo y Consejo de la UE · 27-04-2016

Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial) · https://eur-lex.europa.eu/eli/reg/2024/1689/oj/spa · Parlamento Europeo y Consejo de la UE · 13-06-2024