PLANTILLA P22 · RESULTADOS DE VALIDACIÓN Y DEL PILOTO

# Resultados de validación y del piloto

Recoge los resultados de las pruebas y del piloto en condiciones reales, y los compara con los umbrales y la hipótesis de valor fijados antes de invertir.

**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 Resultados de las pruebas

04 Evaluación específica de IA generativa y agentes (Enterprise)

05 Diseño del piloto

06 Resultados del piloto frente a la hipótesis

07 Valor observado y proyección

08 Criterios de parada

09 Criterios de G5 según el nivel de ambición

10 Conclusión y recomendación del equipo

11 Criterios de calidad

12 Aprobación y verificación

13 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 5 (Entrega y validación), tras ejecutar las pruebas y el piloto y antes de solicitar G5. |
| **Quién la rellena** | Responsable de producto de IA (piloto y valor) y responsable técnico de IA (pruebas). |

| Aspecto | Indicación |
|---|---|
| **Quién la verifica** | Responsable de riesgos de IA en Lite; auditor de IA en Enterprise. El valor lo valida control de gestión cuando se declare como *validado*. |
| **En qué *gate* se revisa** | Puesta en producción. (G5) |
| **Lite frente a Enterprise** | Lite puede omitir los campos marcados **(Enterprise)**. |
| **Herramienta** | T11 Lienzo y calculadora de hipótesis de valor. |
| **Relación** | P08 Lienzo de hipótesis de valor · P09 Línea base · P17 Supervisión humana · P18 Diseño de seguridad · P19 Prueba de reversión · P20 Adopción · P21 Informe de entrega · P28 Seguimiento de valor · documento 40. |

Reglas:

- **Los umbrales se fijaron antes.** Las métricas, umbrales de éxito y criterios de parada son los aprobados en la fase 2 (P08). No pueden relajarse para evitar una decisión de parar sin aprobación del órgano que autorizó el caso de uso (01 §7.4, regla 6).
- **La producción es la única verdad.** Los resultados de laboratorio no sustituyen al piloto en condiciones reales.
- Todo importe tiene fórmula, es incremental frente a la línea base o el grupo de control y tiene estado (reglas 1 y 2 de medición).
- La extrapolación anual de un piloto es siempre **estimada** hasta que se mide en producción.

02

# Identificación

| Campo | Contenido | Guía |
|---|---|---|
| **Código del caso de uso** | | IA-AAAA-NNN (T01). |
| **Sistema(s) de IA y versión probada** | | Identificador (T02) y versión (P16). |
| **Nivel de ambición** | | Optimizar · Aumentar · Transformar. |
| **Intensidad** | | Lite · Enterprise. |

| Campo | Contenido | Guía |
|---|---|---|
| **Versión del documento** | | |
| **Autor** | | Rol y nombre. |
| **Fecha** | | DD-MM-AAAA. |
| **Estado del documento** | | Borrador · Presentado para verificación · Verificado · Aprobado · Sustituido. |

03

# Resultados de las pruebas

**Tipo de prueba:** Funcional · Rendimiento del modelo · Sesgo y no discriminación · Robustez · Seguridad · Inyección de instrucciones · Carga y latencia · Usabilidad · Supervisión humana · Reversión. **Resultado:** Superada · Superada con observaciones · No superada · Pendiente.

| Tipo de prueba | Métrica | Umbral de aceptación y fecha en que se fijó | Resultado obtenido | Estado | Evidencia |
|---|---|---|---|---|---|
| **Rendimiento del modelo *(ejemplo ilustrativo)*** | Proporción de clasificaciones correctas sobre conjunto de prueba independiente | ≥ 90 % · fijado en P08 el 03-03-2026 | 92,4 % | Superada | Informe de evaluación v2 |
| **Sesgo y no discriminación *(ejemplo ilustrativo)*** | Diferencia de tasa de error entre grupos de edad | ≤ 3 puntos | 4,1 puntos en el grupo de mayores de 65 años | No superada | Informe de sesgo v1 |
| **Inyección de instrucciones** | | | | | P18 bloque 11 |
| **Reversión** | Tiempo hasta estado seguro | Tiempo objetivo de P19 | | | P19 bloque 10 |

Toda prueba *No superada* indica su tratamiento y si bloquea la puesta en producción:

| Prueba no superada | ¿Afecta a un control crítico de seguridad, cumplimiento o supervisión humana? (Sí · No) | Tratamiento | Responsable | Fecha |
|---|---|---|---|---|
| | | | | |

04

# Evaluación específica de IA generativa y agentes (Enterprise)

| Aspecto | Método de evaluación | Muestra | Resultado | Umbral | Estado |
|---|---|---|---|---|---|
| **Respuestas fundamentadas en las fuentes autorizadas** | | | | | |
| **Contenido inventado o incorrecto** | | | | | |
| **Contenido inadecuado o contrario a la política de uso** | | | | | |
| **Fuga de información confidencial o personal** | | | | | |
| **Acciones del agente dentro de límites y finalidad** | | | | | |

05

# Diseño del piloto

| Campo | Contenido | Guía |
|---|---|---|
| **Periodo** | | Fechas de inicio y fin. |
| **Población y alcance** | | Área, casos, usuarios y personas afectadas incluidos. |
| **Método de atribución** | | El aprobado en P08: grupo de control · antes y después · otro justificado. |
| **Grupo de control o referencia** | | Composición y cómo se asignó. |
| **Tamaño de muestra y justificación (Enterprise)** | | Por qué el tamaño permite distinguir el efecto del ruido. |

| Campo | Contenido | Guía |
|---|---|---|
| **Nivel de autonomía durante el piloto** | | A0 · A1 · A2 · A3 (P17). |
| **Desviaciones respecto al diseño del piloto** | | Cambios durante el piloto y su efecto en la validez de los resultados. |

06

# Resultados del piloto frente a la hipótesis

| Métrica | Línea base (P09) | Objetivo | Umbral de éxito | Resultado del piloto | Resultado del grupo de control | Diferencia atribuible | ¿Cumple el umbral? (Sí · No · Parcialmente) |
|---|---|---|---|---|---|---|---|
| **Métrica principal:** | | | | | | | |
| **Métrica secundaria:** | | | | | | | |
| **Adopción (P20)** | | | | | — | — | |
| **Tasa de anulación humana** | | | | | — | — | |

07

# Valor observado y proyección

**Tipo:** Eficiencias · Retorno · Coste recurrente. **Estado:** Validado · Declarado · Estimado.

| Concepto | Tipo | Fórmula (unidades × valor unitario) | Importe en el periodo del piloto (€) | Estado | Proyección anual (€) | Estado de la proyección |
|---|---|---|---|---|---|---|
| **Menor tiempo de tramitación materializado *(ejemplo ilustrativo)*** | Eficiencias | 2.300 casos × 3,10 € de coste externo evitado por caso | 7.130 (piloto de 3 meses) | Declarado | 28.520 (× 4 trimestres, mismo volumen) | Estimado |

| Concepto | Tipo | Fórmula (unidades × valor unitario) | Importe en el periodo del piloto (€) | Estado | Proyección anual (€) | Estado de la proyección |
|---|---|---|---|---|---|---|
| **Consumo de modelos y licencias** ***(ejemplo ilustrativo)*** | Coste recurrente | 3 meses × 1.900 € | 5.700 | Validado | 22.800 | Estimado |
| | | | | | | |

| Magnitud | Importe anual proyectado (€) | Estado |
|---|---|---|
| **Eficiencias** | | |
| **Retorno** | | |
| **Coste recurrente** | | |
| **Valor neto anual = eficiencias + retorno − coste recurrente** | | Estimado |
| **Capacidad liberada (horas, no suma)** | | |
| **Riesgo evitado y cumplimiento (no suman salvo fórmula)** | | |

08

# Criterios de parada

| Criterio de parada fijado en fase 2 | Umbral | Resultado | ¿Se ha activado? (Sí · No) |
|---|---|---|---|
| | | | |
| | | | |

Si algún criterio se ha activado, el equipo debe recomendar **Parar** o justificar ante el órgano que autorizó el caso de uso por qué propone otro resultado.

09

# Criterios de G5 según el nivel de ambición

| Nivel | Qué exige G5 (01 §7.6) | Evidencia aportada | Estado (Cumple · No cumple · No aplica · Pendiente) |
|---|---|---|---|
| **Optimizar** | Eficiencia validada frente a la línea base; plan para materializar la capacidad liberada. | | |
| **Aumentar** | Adopción real y mejora de rendimiento medidas. | | |
| **Transformar** | Evidencia de mercado o de cliente: uso, conversión, ingresos iniciales o cambio operativo verificado. | | |

10

# Conclusión y recomendación del equipo

| Campo | Contenido | Guía |
|---|---|---|
| **Conclusión** | | ¿Se sostiene la hipótesis en condiciones reales? |
| **Recomendación al *gate*** | | Continuar · Continuar con condiciones · Iterar · Parar. En G5 no cabe Pivotar (01 §7.3). Es una recomendación; decide el órgano competente en P29. |
| **Condiciones propuestas** | | Solo sobre aspectos no críticos, con plazo y responsable. |

11

# Criterios de calidad

Criterios formales de G5 en el documento 21; lista LV-G5 en el documento 22.

| # | Comprobación | Estado (Cumple · No cumple · No aplica · Pendiente) |
|---|---|---|
| **1** | Los umbrales y criterios de parada son los fijados en la fase 2, con fecha anterior al piloto. | |
| **2** | El piloto se ha realizado en condiciones reales con el método de atribución aprobado. | |
| **3** | Las pruebas de sesgo, robustez, seguridad e inyección de instrucciones se han ejecutado. | |

| # | Comprobación | Estado (Cumple · No cumple · No aplica · Pendiente) |
|---|---|---|
| **4** | Todo importe tiene fórmula, tipo y estado; la proyección anual figura como estimada. | |
| **5** | La capacidad liberada se informa por separado y no suma. | |
| **6** | Las pruebas no superadas sobre controles críticos están resueltas. | |
| **7** | La recomendación es coherente con los resultados y con los criterios de parada. | |

12

# Aprobación y verificación

| Función | Rol | Nombre | Fecha | Firma o referencia |
|---|---|---|---|---|
| **Autor (piloto y valor)** | Responsable de producto de IA | | | |
| **Autor (pruebas)** | Responsable técnico de IA | | | |
| **Validación de importes** | Control de gestión | | | |
| **Verificador** | Responsable de riesgos de IA (Lite) · Auditor de IA (Enterprise) | | | |
| **Decisor del *gate*** | Según 01 §7.5 (registrado en P29) | | | |

Separación de funciones: el verificador y quien valida los importes no forman parte del equipo que construye.

13

# Control de versiones

| Versión | Fecha | Cambios |
|---|---|---|
| **0.1** | 16-09-2026 | Primera versión de la plantilla. |