



MINISTERIO
DE HACIENDA



Fondos Europeos



red.es



COIICO
INGENIEROS
INDUSTRIALES
COLEGIO CANARIAS ORIENTAL

—

Medir para decidir: métricas y analítica para saber si tu IA funciona



01

El problema

Por qué la mayoría de las empresas no saben si su IA funciona

0'-20'

03

Taller en vivo

Construye tu cuadro de mando de IA

60'-90'

02

El framework

Qué medir, cómo medirlo y cuándo preocuparse

20'-50'

04

Casos reales y cierre

Errores fatales, casos reales y conclusiones

90'-120'

01.

El problema

Por qué la mayoría de las empresas no saben si su IA funciona

Todo el mundo usa IA, casi nadie sabe si funciona

72%

Destruyendo valor

De las inversiones en IA destruyen valor por ineficiencia (Larridin, 2025)

42%

Proyectos abandonados

De empresas han abandonado sus proyectos de IA en 2025, vs 17% el año anterior (S&P Global)

25%

ROI esperado

Solo 1 de cada 4 iniciativas de IA entrega el ROI esperado (IBM CEO Study)

79%

Reportan ganancias

De organizaciones reporta ganancias de productividad. Pero ¿justifican lo que cuestan?

⚠ La pregunta no es si la IA hace algo. La pregunta es si lo que hace justifica lo que cuesta.

El síndrome del «va bien, creo»: métricas basadas en vibraciones

Cómo miden la mayoría

«Tenemos X usuarios activos.»

«Se han generado X documentos con IA este mes.»

«El equipo dice que le ahorra tiempo.»

Lo que ninguna de estas métricas responde

Ninguna responde a la pregunta que importa: ¿cuánto valor de negocio está generando esta inversión?

- Medir actividad no es medir impacto
- Contar documentos generados no es medir calidad
- Que la gente lo use no significa que funcione

Por qué falla la medición: 5 trampas habituales

1

Sin línea base

No definir el «antes» antes de desplegar. Sin un punto de partida documentado, no hay forma de calcular el «después».

2

Actividad $\neq$ resultados

Número de prompts, llamadas a la API o documentos generados no dicen nada sobre el impacto en el negocio.

3

Correlación $\neq$ causalidad

Que las ventas suban el mismo trimestre que se despliega la IA no significa que la IA las haya subido.

4

Costes ocultos

Licencias, integración, formación, supervisión, retrabajos. El coste real casi nunca está solo en la factura del proveedor.

5

Impuesto de retrabajo

Si el equipo dedica un tercio del tiempo ahorrado a corregir lo que la IA genera, el ROI real es mucho menor del que crees.

La pregunta equivocada

«¿Usamos IA?»

Esta pregunta mide adopción, no valor. Es el punto de partida, no el destino.

La pregunta correcta

Antes de escribir una sola línea de código, la pregunta debería ser: ¿qué indicador de negocio necesito mover? Y después: ¿puedo demostrar que la IA lo ha movido?

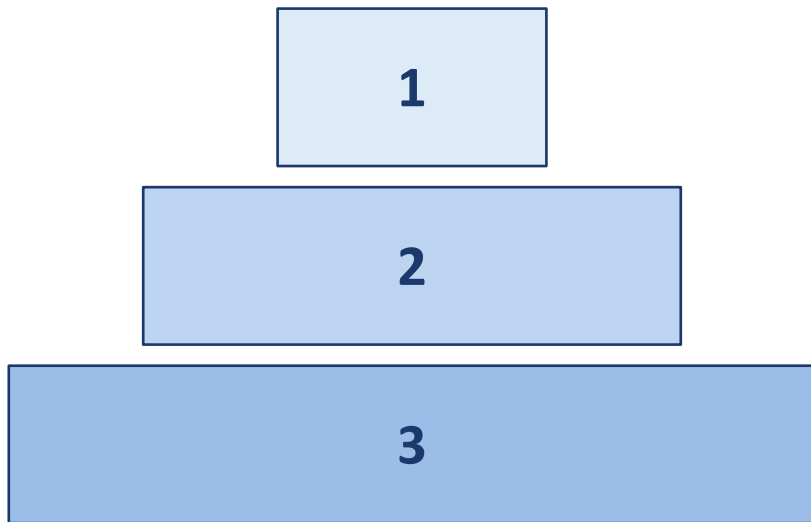
✗ Si no puedes conectar tu proyecto de IA con un resultado de negocio medible (coste, tiempo, calidad, ingresos, satisfacción), no tienes un proyecto de IA. Tienes un experimento sin hipótesis.

02.

El framework

Qué medir, cómo medirlo y cuándo preocuparse

Framework de medición: 3 niveles, de lo básico a lo estratégico



Nivel 3 — Impacto en negocio

¿Mueve una métrica que importa? Reducción de coste, aumento de ingresos, satisfacción del cliente, reducción de errores, time-to-market.

Nivel 2 — Eficiencia

¿Ahorra tiempo o recursos? Tiempo por tarea antes vs. después, tasa de retrabajo, coste por tarea, horas reasignadas.

Nivel 1 — Actividad

¿Se usa? Usuarios activos, frecuencia de uso, tareas ejecutadas. Es el mínimo, pero no dice nada sobre el valor.

❗ La mayoría de las empresas se quedan en el Nivel 1. El valor real está en el Nivel 3.

Nivel 1: Métricas de actividad — el mínimo necesario

Qué medir

- Número de usuarios activos (diarios, semanales, mensuales)
- Frecuencia de uso por usuario
- Distribución por departamento, equipo o rol
- Tasa de adopción vs. licencias disponibles

Para qué sirve

Detectar si la herramienta se está usando o si se ha convertido en shelfware. Identificar equipos que necesitan más formación o soporte.

Limitación crítica

⚠ Un 80% de adopción con 0% de impacto en negocio sigue siendo un fracaso. Que algo se use no significa que funcione.

Nivel 2: Métricas de eficiencia — ¿realmente ahorra?

Tiempo medio por tarea

Antes vs. después de la IA. Ejemplo: «Redactar un informe pasó de 4 horas a 45 minutos.»

Rework Tax (Impuesto de retrabajo)

Porcentaje del tiempo ahorrado que se dedica a corregir el output de la IA. Si la IA ahorra 3h pero genera 1h de correcciones, el ahorro real es 2.

Horas reasignadas

¿El tiempo ahorrado se dedica a trabajo de mayor valor o se evapora? Esta distinción es fundamental para el ROI real.

Coste por tarea

Comparar el coste total (licencia + tiempo humano + supervisión) antes y después del despliegue de la IA.

X Regla de alerta: si el equipo dedica más de un tercio del tiempo ahorrado a revisar y corregir, tu cálculo de ROI necesita revisión urgente.

Nivel 3: Métricas de impacto en negocio — lo que importa de verdad



Atribuible a la IA, en euros, no en porcentaje vago.



Si aplica al caso de uso concreto del proyecto.



CSAT, NPS. Métricas reales, no percepciones del equipo.



Incidencias, reclamaciones, devoluciones.
Cuantificadas en euros.



Ciclo de entrega, velocidad de producción,
resolución autónoma en agentes.

No digas «el modelo tiene un 94% de precisión». Di «la precisión del 94% evitó 3,2 M€ en fraude el último trimestre».

La métrica que nadie quiere ver: el Rework Tax

¿Qué es el Rework Tax?

El Rework Tax mide el tiempo humano necesario para verificar, editar y corregir los outputs generados por la IA. Es la métrica que separa el ahorro real del ahorro percibido.

Cómo calcularlo

Tiempo ahorrado por la IA
— Tiempo dedicado a revisar y corregir
= Ahorro neto real

Ejemplo práctico

Si Copilot genera un informe en 10 minutos (antes tardabas 2 horas), pero revisarlo y corregirlo te lleva 40 minutos, tu ahorro real es

no 110 minutos.

X Si el Rework Tax supera el 33% del tiempo ahorrado, la IA está redistribuyendo trabajo, no eliminándolo.

Sin línea base no hay medición: cómo documentar el «antes»

Antes de desplegar cualquier herramienta de IA, documenta estos elementos con datos reales, no estimaciones:



Tiempo por tarea

Cuánto tiempo lleva cada tarea clave. Medido, no estimado.



Coste de la tarea

Coste/hora del profesional × tiempo. Calculado con datos reales.



Tasa de error actual

Cuál es el nivel de error o reproceso antes de la IA.



Satisfacción actual

Nivel de satisfacción del cliente o usuario interno en el momento cero.

❗ Mídalo durante 2-4 semanas antes del despliegue. Si no tienes línea base, cualquier mejora que reportes es una opinión, no un dato.

03.

Taller en vivo

Construye tu cuadro de mando de IA

Ejercicio 1: Documenta tu línea base

Instrucciones

Piensa en 3 tareas de tu trabajo diario que podrían beneficiarse de IA (o que ya uses con IA). Para cada una, completa la siguiente tabla:

Tarea / Frecuencia / Tiempo medio actual / Coste estimado (€) / Tasa de error (%)

Ejemplo

Redactar acta de reunión:

- Frecuencia: 3/semana
- Tiempo medio: 45 min
- Coste estimado: 37,50 €
- Tasa de error: 10% (datos incompletos)

Tu tarea 2 — ...

Tu tarea 3 — ...

⚠ Reto: Sé honesto con los tiempos. No pongas lo que crees, pon lo que mides. Si no lo has medido nunca, ese es el primer problema que debes resolver.

Ejercicio 2: Calcula tu Rework Tax real

Instrucciones

Si ya usas alguna herramienta de IA (Copilot, ChatGPT, Claude, Gemini), calcula para cada tarea:

Tarea / Tiempo sin IA / Tiempo con IA (generación) / Tiempo de revisión/corrección / Ahorro neto / Rework Tax (%)

Ejemplo

Informe técnico:

- Tiempo sin IA: 4h
- Tiempo con IA: 20 min
- Revisión/corrección: 50 min
- Ahorro neto: 2h 50 min
- Rework Tax: 71% del tiempo con IA es revisión

⚠ Reto: Si tu Rework Tax supera el 33%: necesitas mejorar los prompts, cambiar de herramienta o replantear si esa tarea es un buen caso de uso para IA.

Ejercicio 3: Diseña tu cuadro de mando de IA

Qué debes definir

Elige un proyecto de IA (real o hipotético) y define:

1. Objetivo de negocio — ¿qué métrica quieres mover?
2. Métrica Nivel 1 — ¿cómo medirás la adopción?
3. Métrica Nivel 2 — ¿cómo medirás la eficiencia?
4. Métrica Nivel 3 — ¿cómo medirás el impacto?

Continúa con

5. Línea base — ¿cuál es el valor actual de esas métricas?
6. Objetivo a 6 meses — ¿cuál es el valor esperado?
7. Frecuencia de revisión — ¿cada cuánto mides?

Frecuencia recomendada: mensual los primeros 6 meses, trimestral después.

 **Reto:** Comparte tu cuadro con el grupo. ¿Alguien tiene métricas que no pueda medir? Ese es un problema de datos, no de IA.

Ejercicio 4: Encuentra el fallo en este caso de ROI

Lo que la empresa presentó

«Hemos ahorrado 1.200 horas/año en generación de informes.»

«La satisfacción del equipo ha subido un 15%.»

«Tenemos 45 usuarios activos de 50 licencias (90% adopción).»

Preguntas para el grupo

- ¿Cuánto costaron las 50 licencias + integración + formación?
- ¿Las 1.200 horas se han reasignado a trabajo de valor?
- ¿Se midió la calidad de los informes antes y después?
- ¿Se ha medido el Rework Tax?
- ¿La satisfacción del equipo se traduce en algún resultado de negocio?

 **Reto:** Identifica al menos 3 métricas que faltan antes de validar este ROI.

Ejercicio 4: Encuentra el fallo en este caso de ROI (respuestas)

Lo que la empresa presentó

«Hemos ahorrado 1.200 horas/año en generación de informes.»

«La satisfacción del equipo ha subido un 15%.»

«Tenemos 45 usuarios activos de 50 licencias (90% adopción).»

Preguntas para el grupo

- 1. ¿Cuánto costaron las 50 licencias + integración + formación?→ **No se menciona en ningún momento. ROI incalculable.**
- 2. ¿Las 1.200 horas se reasignaron a trabajo de valor?→ No hay evidencia. **El ahorro puede ser solo tiempo libre no productivo.**
- 3. ¿Se midió la calidad de los informes antes y después?→ **No. Solo se mide velocidad, no precisión ni utilidad real.**
- 4. ¿Se ha medido el Rework Tax?→ **No. Si los informes generados requieren correcciones, el ahorro real es menor.**
- 5. ¿La satisfacción del equipo se traduce en algún resultado de negocio?→ **No. Un +15% de satisfacción sin impacto medible en retención o productividad no es ROI.**

 **Reto:** Identifica al menos 3 métricas que faltan antes de validar este ROI.

04.

Casos reales, errores fatales y cierre

Errores fatales, casos reales y conclusiones

Caso 1: Copilot en una ingeniería de 15 personas

INGENIERÍA INDUSTRIAL

15 LICENCIAS · 5.400 €/AÑO

80%

Adopción

12 de 15 usuarios activos semanales

2,5h

Ahorro por informe

Tiempo de redacción reducido de 4h a 1,5h por informe técnico

25%

Rework Tax

38 min de revisión por informe. Por debajo del umbral crítico del 33%

482%

ROI neto

Ahorro neto estimado de 31.400 €/año descontando licencias, formación y Rework Tax

⚠ Clave: se midió la línea base antes del despliegue. Sin ella, estos números serían una estimación, no un dato.

Caso 2: Chatbot de atención al cliente — métricas que engañan

EMPRESA DE SERVICIOS

INVERSIÓN: 18.000 €/AÑO

Métricas que presentaron

«El chatbot atiende el 70% de las consultas sin intervención humana.»

Sonaba bien. Pero faltaba la mitad de la historia.

Métricas que faltaban

Tasa de resolución real: solo el 42% quedaba resuelto. El 28% restante abandonaba o llamaba después.

CSAT con chatbot: 3,1/5 vs. 4,2/5 con agente humano.

Coste por interacción resuelta: 2,80€ (chatbot) vs. 4,50€ (humano).
Ahorro real: 1,70€, no los 4,50€ reportados.

X La métrica de contención no es la métrica de resolución. Medir una por otra infla el ROI artificialmente.

Caso 3: Visión artificial en control de calidad

FABRICACIÓN DE COMPONENTES METÁLICOS

INVERSIÓN: 65.000 €

Nivel 2 — Eficiencia

Detección de defectos: 96% vs. 82% del control manual.
Tiempo de inspección: de 45 s/pieza a 3 s/pieza.

Nivel 3 — Impacto

Tasa de rechazo de cliente: del 4,2% al 0,8%.
Ahorro en devoluciones: 42.000 €/año.
Payback inicial: 18 meses.

El detalle crítico

El 4% de falsos positivos genera paradas de línea de 8 minutos. Coste: 6.200 €/año. Payback real: 22 meses, no 18.

⚠ Lección: medir solo la precisión sin medir el coste de los errores (falsos positivos y falsos negativos) distorsiona el ROI.

El cuadro de mando mínimo viable para tu proyecto de IA

Todo proyecto de IA debería tener al menos estas 7 métricas:

1

Adopción

% de usuarios activos vs. licencias/accesos disponibles.

2

Frecuencia

Número de usos por usuario por semana.

3

Tiempo ahorrado

Horas/semana liberadas por la IA. Medido, no estimado.

4

Rework Tax

% del tiempo ahorrado dedicado a corregir outputs de la IA.

5

Ahorro neto

$(\text{Tiempo ahorrado} - \text{retrabajo}) \times \text{coste/hora} - \text{coste de la herramienta}.$

6

Métrica de negocio

El KPI específico que el proyecto debía mover.

7

Satisfacción del usuario

¿El equipo percibe valor real o carga adicional?

Los 5 errores fatales en medición de IA

1. Medir solo la adopción

Un 90% de adopción con 0% de impacto en negocio sigue siendo un fracaso. La adopción es condición necesaria, no suficiente.

2. No documentar la línea base

Sin un «antes» medido, cualquier mejora reportada es una opinión disfrazada de dato.

3. Confundir contención con resolución

Que el chatbot «atienda» el 70% de las consultas no significa que las resuelva.

4. Ignorar los costes ocultos

Formación, integración, supervisión, retrabajo, coste de oportunidad. El ROI se calcula con todos los costes.

5. Reportar el ROI bruto sin descontar el Rework Tax

Si no descuentas el tiempo de corrección, estás inflando el ahorro y engañando a tu organización.

Quién mide, cuándo y para quién

Un cuadro de mando de IA sin dueño es un PDF que nadie mira. Define estos cuatro elementos de gobernanza:



Responsable de medición

Una persona concreta, no «el equipo». Idealmente alguien con acceso a datos de uso y a métricas de negocio.



Frecuencia

Mensual durante los primeros 6 meses, trimestral después. Sin excepciones.



Audiencia

El CTO quiere métricas técnicas. El CFO quiere euros. El CEO quiere impacto estratégico. Adapta el cuadro a quién lo ve.



Decisiones vinculadas

Cada revisión debe terminar con una decisión: escalar, ajustar, pivotar o cancelar. Si no produce decisiones, no es una revisión, es un trámite.

Checklist antes de reportar el ROI de tu proyecto de IA

- ☒ ¿Tienes una línea base documentada con datos reales (no estimaciones)?
- ☒ ¿Estás midiendo resultados de negocio, no solo actividad?
- ☒ ¿Has calculado el Rework Tax y lo has descontado del ahorro bruto?
- ☒ ¿Incluyes todos los costes (licencias, integración, formación, supervisión, retrabajo)?
- ☐ ¿Puedes separar el efecto de la IA de otros factores que hayan cambiado al mismo tiempo?
- ☐ ¿Tienes un responsable de medición con nombre y apellidos?
- ☐ ¿El cuadro de mando produce decisiones, no solo informes?
- ☐ ¿Estás midiendo la satisfacción real del usuario, no solo la adopción?



COIICO
INGENIEROS
INDUSTRIALES

COLEGIO CANARIAS ORIENTAL



Cofinanciado por
la Unión Europea



MINISTERIO
DE HACIENDA



Fondos Europeos



GOBIERNO
DE ESPAÑA

MINISTERIO
PARA LA TRANSFORMACIÓN DIGITAL
Y DE LA FUNCIÓN PÚBLICA

SECRETARÍA DE ESTADO
DE DIGITALIZACIÓN
E INTELIGENCIA ARTIFICIAL

red.es