

EL GENAI DIVIDE · QUÉ MIDE REALMENTE EL 95%

La Cifra Es Aproximadamente Correcta. La Lectura Es Errónea.

Por qué el 95 % del MIT mide una brecha de comprensión, no una falla tecnológica — y por qué las brechas de comprensión y las de cálculo exigen respuestas opuestas

AUDIENCIA

CIOs, VPs Ops, directores mineros

FOCO

Leer el 95 % del MIT por tipo de falla

REGIÓN

Industria pesada LATAM · Chile

Contenido

01	El MIT No Midió Lo Que Dice el Titular	3
02	La Cifra en la Lámina del Directorio, Traducida a un Portafolio de Cobre	4
03	El Año en Que Tres Pilotos Cerraron Con Un Solo Retorno Reportado	5
04	Lo Que el Comprador Creyó Comprar, y Lo Que la Tecnología Cobra	6
05	Por Qué el Argumento de Wall Street No Llega a Tu Planta	7
06	Una Señal Direccional, Leída Como Veredicto	8
07	El Ejecutivo Que No Ve el Costo, y el Que Ya Lo Calculó	9
08	Lo Que Esconde un Éxito Reportado	10
09	Qué Preguntar Antes de la Próxima Revisión de Portafolio	11
10	Los Remedios Que Tratan la Enfermedad Equivocada	12
11	Qué Confirmar en Tu Propio Portafolio Este Trimestre	13
12	Ideas Clave	14
13	Fuentes	16

01

EL MIT NO MIDió LO QUE DICE EL TITULAR

El "95 % sin ROI" del MIT no es una medición del fracaso de la IA. Es una medición de distancia: la brecha entre lo que una empresa creyó comprar y lo que la tecnología cobra antes de pagar. La cifra es aproximadamente correcta. La lectura estándar es errónea, y la diferencia entre ambas vale varios trimestres de tu presupuesto de IA.

La cifra proviene de *The GenAI Divide: State of AI in Business 2025*, publicado en agosto de 2025 por Project NANDA ("Networked Agents and Decentralized AI"), un grupo de investigación del MIT Media Lab. Antes de construir nada encima, conviene leer qué es. El estudio descansa sobre cincuenta y dos entrevistas estructuradas a ejecutivos, 153 respuestas de encuesta recogidas entre líderes senior en conferencias de industria, y una revisión de más de 300 iniciativas de IA divulgadas públicamente cuyo método de síntesis nunca se publicó. El "ROI medible" (retorno sobre la inversión) lo definió cada respondente, no un estándar común, y el 95 % se refiere a organizaciones que no reportan impacto medible en el estado de resultados — no a pilotos contados como unidades aisladas. Los propios autores lo advirtieron: el reporte lleva el descargo "hallazgos preliminares, anonimizados por confidencialidad", y describe sus conclusiones como direccionalmente precisas sobre la base de entrevistas individuales, no de reportes oficiales de las compañías. Es una fotografía tomada temprano sobre una curva de capacidad que se mueve en trimestres, no en décadas.

Esa distinción es todo el juego. Una señal direccional te dice hacia dónde corre el río. Una medición precisa te dice cuánta profundidad hay en un punto con nombre. El retelling viral de la cifra del MIT trata una señal direccional como una medición precisa, y de ahí extrae una conclusión precisa: *el 95 % de los pilotos de IA fracasa, por lo tanto la IA no funciona*. La premisa es blanda. La conclusión es dura. El razonamiento entre ambas está roto.

Una compañía que construya una tesis sobre la cifra del MIT tiene que sostener dos cosas a la vez. La señal es real: la mayoría del gasto empresarial en GenAI (IA generativa) de 2025 no devolvió nada que la función de finanzas pudiera ver. Y la señal es gruesa: no te dice, para ningún portafolio individual, *por qué* el retorno fue invisible. El "por qué" no está en el reporte. El "por qué" está en tu propia operación, y hay al menos dos, con remedios opuestos.

Este documento defiende una sola tesis: el 95 % mide una brecha de comprensión, no un veredicto sobre la tecnología. Un cero en la línea de ROI es un síntoma. El mismo síntoma lo producen dos condiciones distintas que para un CFO se ven idénticas y exigen respuestas opuestas. Nombrar cuál de las dos tienes es el entregable. Todo lo demás es comentario.

02

LA CIFRA EN LA LÁMINA DEL DIRECTORIO, TRADUCIDA A UN PORTAFOLIO DE COBRE

La estadística del MIT llega al directorio como una sola frase en un artículo reenviado, normalmente por un director que lo leyó en un vuelo. Para cuando aterriza frente al ejecutivo que firmó el portafolio de IA, ya viene comprimida en una sola pregunta: *¿deberíamos pausar?*

Esa pregunta no puede responderse a la altitud a la que se formula. "Deberíamos pausar el programa de IA" supone que hay un programa con un solo estado de salud. Un portafolio real en una productora de cobre no es un programa. Es una concentradora grande corriendo un optimizador de control de intervalo corto en el circuito de molienda, una oficina de planificación piloteando un copiloto de mantenimiento, un asistente de balance metalúrgico que un superintendente construyó sin avisarle a nadie, y tres relaciones con vendedores en tres etapas distintas de decepción. Cada una de esas piezas tiene su propia respuesta a "¿está funcionando?". El directorio ve el promedio. El promedio es el número menos útil de la sala.

Considera lo que ese promedio esconde. Una operación polimetálica en el sur de Perú, con tres vendedores de IA y un equipo de datos interno, puede tener una iniciativa produciendo valor real, una produciendo la apariencia de valor y una produciendo nada, y reportar las tres a través de una sola línea de "ROI del portafolio de IA" que netea aproximadamente cero. El cero es aritméticamente correcto. Como insumo de decisión no vale nada, porque las tres iniciativas necesitan tres decisiones distintas: acelerar, re-medir, detener.

Esto es el 95 % del MIT renderizado a la resolución en la que un operador realmente trabaja. El reporte aplana millones de estos portafolios en un solo número. La lectura estándar aplana después tu portafolio de la misma manera, y te invita a tomar una sola decisión para iniciativas que necesitan tres. El aplanamiento es el error. El instinto del directorio de preguntar "¿pausar o acelerar?" es el aplanamiento en acción.

El trabajo, entonces, es recuperar la distribución detrás del promedio. No la distribución de la industria: la tuya. ¿Qué fracción de tu portafolio está produciendo valor que el libro contable no puede ver? ¿Qué fracción está produciendo la apariencia de valor sin línea base detrás? ¿Qué fracción genuinamente no devuelve nada y, de esa fracción, cuánto se arregla con mejor información versus cuánto se arregla solo cambiando quién decide qué? Esas son preguntas contestables. La cifra del MIT no es una de ellas. Es el empujón para empezar a hacerlas.

03

EL AÑO EN QUE TRES PILOTOS CERRARON CON UN SOLO RETORNO REPORTADO

En las revisiones de portafolio de IA que hemos corrido con productoras de cobre en LATAM, el patrón se repite con una regularidad que ya no nos sorprende. Una de esas revisiones lo resume todo. Era un jueves de fines de marzo, el cierre del año fiscal, cuando la oficina de planificación integrada de un grupo minero top-cinco de la región presentó al directorio su portafolio de IA de tres años. El CIO que había firmado ese portafolio había leído la estadística del MIT la semana anterior, en un reenvío de un miembro del directorio, y llegó a la revisión preparado para recomendar una pausa. Al final de la reunión la recomendación había cambiado, porque la revisión —corrida con nuestro equipo en la sala— sacó a la superficie algo que el resumen del portafolio tenía enterrado.

Tres pilotos habían corrido durante tres años contra un gasto total de aproximadamente USD 2,6 millones.

El primero era un optimizador de control de intervalo corto en el circuito de molienda de una concentradora de 95.000 toneladas por día. Reportaba el único número duro de ROI del portafolio: una ganancia declarada de 1,8 % de throughput, contabilizada en unos USD 9,2 millones al año. Ese número era el titular de la presentación. También era el único número que alguien cuestionó, porque otras dos cosas habían cambiado en los mismos dos trimestres en que el optimizador entró en servicio: un cambio de revestimientos programado del molino que restauró el desempeño mecánico, y una nueva receta de mezcla de mineral del grupo de metalurgia. Tres causas, un indicador movido, y el piloto de IA se había atribuido el delta completo.

El segundo era un copiloto de planificación de mantenimiento. Su métrica de éxito era adopción: el 80 % de los planificadores lo usaba al menos semanalmente. La presentación reportaba ese 80 % como impacto. Nadie había medido el tiempo de ciclo de planificación antes de que llegara el copiloto, así que no existía línea base contra la cual convertir “uso semanal” en horas ahorradas o detenciones evitadas. La actividad era real. Si era un resultado, nadie podía decirlo, porque la pregunta no se instrumentó antes del despliegue.

El tercer piloto no estaba en el portafolio de IA. Un superintendente de balance metalúrgico, trabajando con una cuota de créditos de modelo y sin comité directivo, había construido un asistente que extraía las cifras de producción de fin de mes desde el historial de planta y el LIMS (sistema de gestión de información de laboratorio) y redactaba la conciliación que cerraba las cuentas metalúrgicas. El cierre había tomado seis días durante los dos años anteriores. Ahora tomaba dos. Tres analistas recuperaron aproximadamente cuatro días cada mes. El valor era modesto y real, quizás USD 180.000 al año en tiempo de analista recuperado más un cierre más rápido y más limpio. Se había contabilizado bajo “mejora de procesos de finanzas”. Nunca se etiquetó como IA, así que nunca llegó a la contabilidad de IA.

Lee el portafolio como lo vio primero el directorio: un número grande de ROI bajo cuestionamiento, una pila de métricas de adopción, y un gasto total de USD 2,6 millones. Léelo como

lo dejó la revisión: un retorno declarado que tres causas podían explicar, un piloto sin línea base, y un retorno genuino que la contabilidad de IA no podía ver porque estaba archivado en otra parte. El fracaso aparente del portafolio y su único éxito real eran, ambos, producto de cómo se clasificaron las cifras, no de la tecnología.

El CIO no pausó el programa. Pausó dos pilotos, conservó el asistente de conciliación y le pidió a finanzas re-etiquetarlo, y mandó el optimizador de vuelta a re-medirse contra el cambio de revestimientos y la mezcla de mineral. Esa no es una historia sobre si la IA funciona en minería. La tecnología funcionó en el único lugar donde se la apuntó a un problema nítido con una línea base medida. Es una historia sobre un número que significaba lo contrario de lo que aparentaba.

04

LO QUE EL COMPRADOR CREYÓ COMPRAR, Y LO QUE LA TECNOLOGÍA COBRA

El optimizador de molienda, el copiloto de planificación y el asistente de conciliación fracasaron y triunfaron por la misma razón de fondo, expresada de tres maneras. Cada uno estaba gobernado por un contrato que casi nadie escribe, y la brecha dentro de ese contrato es lo que mide el 95 % del MIT.

Hay cuatro partes en el contrato cada vez que una organización compra GenAI, y cada una firma una versión distinta.

EL CONTRATO NO ESCRITO DE COMPRENSIÓN

EL VENDOR VENDE	"Inteligencia como servicio. Conecta tus datos, pregunta, recibe insight."
EL COMPRADOR COMPRA	"Una herramienta de productividad que se paga en doce meses sobre la línea que podemos ver."
LA PLANTA RECIBE	"Un motor de texto probabilístico sin memoria, sin autoridad para decidir y sin cableado a los sistemas donde corre el trabajo."
FINANZAS MIDE	"¿Se movió la línea del estado de resultados, y podemos atribuir el movimiento?"

La brecha entre dos cualesquiera de estas versiones es una tajada del 95% del MIT.

Figura 1 · Las cuatro versiones del contrato no escrito. El 95 % es la distancia entre lo que se vendió, se compró, se recibió y se midió — no una propiedad del modelo.

El vendedor vendió una capacidad que la tecnología todavía no tiene por sí sola. El comprador compró un payback cuyo andamiaje el presupuesto nunca financió. La planta recibió una herramienta que sugiere pero no puede imponer, y que no recuerda nada entre sesiones sin una arquitectura que nadie cotizó. Finanzas midió una línea del estado de resultados que ninguna de las otras tres partes estaba moviendo realmente. Cuando la línea medida marca cero, todas las partes se sorprenden, y cada una culpa a una distinta de las otras tres.

Por esto el 95 % del MIT es una burbuja de comprensión, no una burbuja financiera. El capital se gastó completo. El valor fue real en algunos bolsillos y ausente en otros. Lo que se infló no fue el precio de un activo. Lo que se infló fue la distancia entre lo que cuatro partes creían haber acordado. Una burbuja financiera se desinfla cuando un precio colapsa. Una burbuja de comprensión se desinfla cuando el comprador lee la especificación real del producto y la concilia contra el presupuesto, la cadena operativa y el marco de medición.

El contrato se aprieta cuando los cuatro roles convergen. El asistente de conciliación funcionó porque un superintendente sostenía los cuatro roles: no se vendió nada a sí mismo, compró una herramienta para un problema con el que había convivido, la desplegó donde los datos ya estaban, y la midió contra una línea base de seis días que había observado durante dos años. El optimizador y el copiloto aflojaron el contrato a través de cuatro funciones distintas, ninguna de las cuales había leído la versión de las otras. El número que resulta no es una medición del modelo. Es una medición de cuán lejos quedaron las cuatro firmas.

05

POR QUÉ EL ARGUMENTO DE WALL STREET NO LLEGA A TU PLANTA

Un lector puede colapsar este argumento en otro distinto, y un lector hostil lo hará. La tesis adyacente dice: la IA es una burbuja, las valoraciones están desconectadas de los fundamentos, la corrección viene en camino, así que el operador prudente espera. Esa es la tesis de la burbuja financiera. Es un argumento real. No es este, y la distinción importa porque una tesis puede ser cierta sin que la otra lo sea.

La tesis de la burbuja financiera es una afirmación sobre precios de mercado: que lo que los inversionistas pagan por exposición a IA excede lo que esos negocios van a ganar. Defenderla requiere evidencia que este documento no tiene y sobre la que no toma posición: datos de flujos de capital, múltiplos de valoración, la conversión de caja de los vendedores de modelos. Que esa tesis sea cierta o no, no afecta el argumento de aquí.

La tesis de la burbuja de comprensión es una afirmación sobre el entendimiento del comprador: que la distancia entre lo que las empresas creen que compraron y lo que la tecnología exige para entregar valor es grande, y que esa distancia — no la capacidad de la tecnología — explica la mayor parte del 95 % del MIT. Se puede sustanciar desde dentro de un solo portafolio, con el tipo de revisión que corrió la oficina de planificación. No necesita ni un dato de mercado.

Esta es la frase que las separa. *La burbuja de comprensión seguiría explicando tu línea de ROI cero incluso si las acciones de IA estuvieran bien valoradas, y una corrección de la burbuja financiera no cerraría, por sí sola, la brecha de comprensión de tu operación.* Supón que el mercado está racionalmente valorado y la IA es una inversión sólida de largo plazo. Tu portafolio igual puede marcar cero, porque tu contrato de cuatro partes está flojo. Supón que el mercado está salvajemente sobrevalorado y una corrección barre a la mitad de los vendedores. Tu asistente de conciliación sigue ahorrando cuatro días-analista al mes, porque su valor nunca dependió de una valoración.

Las dos tesis se fusionan con frecuencia en la conversación de directorio, y la fusión sale cara. Amarra una afirmación defendible sobre tu propia realidad operativa a una afirmación indefendida sobre timing de mercado, e invita a un escéptico a descartar ambas pinchando la que nunca estabas haciendo. Mantenlas separadas. No tomes posición sobre la burbuja financiera. No la necesitas. La burbuja de comprensión está sentada en tu última revisión fiscal, y puedes actuar sobre ella el lunes sin un solo dato sobre la capitalización de mercado de nadie.

06

UNA SEÑAL DIRECCIONAL, LEÍDA COMO VEREDICTO

La cifra del MIT se viralizó por una razón que vale la pena nombrar, porque la razón es también la trampa. Una estadística que suena precisa y confirma una sospecha viaja más rápido que una cuidadosa que la complica. “El 95 % de los pilotos de IA fracasa” confirmó lo que un ejecutivo cansado ya medio creía después de un año de demos decepcionantes. Dio permiso para parar. El permiso viaja.

Los críticos metodológicos llegaron en días, no en meses. El Marketing AI Institute publicó su réplica el 26 de agosto de 2025 — ocho días después del lanzamiento de prensa del reporte — y planteó el problema con precisión: el retelling eliminó todos los calificadores que los autores mantuvieron, y una estadística sin calificadores es un objeto distinto del que los investigadores publicaron. La propia barra de éxito era inusualmente estricta: despliegue más allá del piloto con ROI medible dentro de seis meses, una definición que archiva como fracaso las ganancias de eficiencia y la reducción de costos. La misma semana, el profesor de Wharton Kevin Werbach lanzó el desafío académico más filosófico que registra el episodio: si Project NANDA respalda sus afirmaciones, debería publicar los datos completos que las sostienen; si no, debería retractar el reporte. A mediados de 2026, NANDA no ha hecho ninguna de las dos cosas. Los investigadores midieron una tendencia y lo dijeron. Internet reportó una ley, y los autores dejaron circular la ley.

Hay un problema de segundo orden debajo incluso de la lectura cuidadosa, y afila el caso en lugar de debilitarlo. Puede que el ROI sea el instrumento equivocado para esta medición desde el principio. El grupo de educación ejecutiva de UC Berkeley argumentó en septiembre de 2025 que el retorno sobre la inversión, calibrado para proyectos de capital con curvas de payback predecibles, mide mal una capacidad que mejora sobre una curva empinada y entrega valor en formas que el estado de resultados tarda en reconocer. Si la propia métrica

está mal calibrada para el activo, entonces un portafolio de ceros te dice tanto sobre la regla de medir como sobre la cosa medida.

Sostén las tres observaciones juntas. El estudio es direccional, no preciso, por el framing de sus propios autores. El retelling eliminó los calificadores que lo hacían honesto, y el desafío sobre los datos de fondo sigue sin respuesta. Y la métrica en la que el retelling se apoya puede ser la métrica equivocada para el activo. Nada de esto rescata a un programa de IA genuinamente fracasado. Todo esto descalifica la inferencia perezosa de que 95 % significa que la tecnología no funciona.

Nombrar estos límites no es cubrirse. Es el movimiento de credibilidad que separa un análisis serio de una lámina que hizo pattern-matching sobre un número viral. Un ejecutivo que entra a la reunión de directorio y dice "el estudio del MIT es direccional; esto es lo que puede y no puede decirnos sobre nuestro portafolio" ya ganó la discusión contra el director que solo leyó el titular. La respuesta de 90 segundos a "¿por qué el MIT dice que el 95 % fracasa?" es esta: no dice eso; dice algo más angosto y más útil, y así es como lo leemos dentro de nuestra propia operación.

07

EL EJECUTIVO QUE NO VE EL COSTO, Y EL QUE YA LO CALCULÓ

El cero en la línea de ROI lo producen dos condiciones que para finanzas se ven idénticas y que son opuestas al nivel del remedio. Distinguir las es el acto central del diagnóstico, porque el remedio equivocado aplicado a la condición correcta es una de las maneras más confiables de fabricar otro cero.

La primera condición es una **falla de comprensión**. El ejecutivo dueño de la iniciativa todavía no ve la brecha entre lo que se compró y lo que la tecnología exige. El costo de modelo operativo de capturar el valor — los derechos de decisión que hay que mover, los datos que hay que hacer alcanzables, la línea base que hay que medir — nunca se hizo visible, así que nunca se financió. Esta condición es contingente y corregible. Se disuelve cuando el costo sale a la superficie. Muéstrale al dueño el contrato de cuatro partes, nómbrale el andamiaje que el presupuesto saltó, y el dueño puede redefinir el alcance, re-medir y continuar. La enfermedad está aguas abajo de información faltante, y la información es la cura. Este ejecutivo responde a evidencia y argumento.

La segunda condición es una **falla de cálculo**. El ejecutivo ve la brecha perfectamente. Los derechos de decisión están fragmentados entre tres funciones, los datos viven en una función que no los va a compartir, la responsabilidad de actuar sobre el output está indefinida, y el ejecutivo ya hizo la aritmética: rediseñar el modelo operativo para arreglar esto cuesta más, en capital político y disrupción, que el ROI que el programa de IA habría devuelto. Así que se deja que el programa fracase en silencio. Esta condición es estructural, enraizada en incentivos y poder, no en un punto ciego. No responde a evidencia, porque la evidencia no falta. El ejecutivo ya la tiene y eligió racionalmente el ROI perdido por sobre el costo del rediseño. Esta condición responde solo a incentivos cambiados o a autoridad cambiada.

No son dos puntos de un espectro. Son diagnósticos opuestos con curas opuestas, y la cura de uno es inerte contra el otro.

Tabla 1 La clave diagnóstica. El mismo síntoma de ROI cero; condiciones opuestas; primeros movimientos opuestos.

Señal	Falla de comprensión	Falla de cálculo
¿El dueño ve el costo de modelo operativo?	No — nunca se hizo visible	Sí — y ya lo comparó contra el ROI perdido
Qué necesita el dueño	Información: el costo, nombrado y cuantificado	Incentivo o autoridad: una razón para pagar el rediseño
Respuesta a un briefing claro	Redefine el alcance y acelera	Asiente, está de acuerdo, no cambia nada
Pregunta "¿quién decide si se actúa sobre el output?"	Un nombre, hoy demasiado cauto	Silencio, o tres funciones con poder de veto
Primer movimiento correcto	Acelerar, con el contrato reescrito	Detener el programa de IA; arreglar primero la cadena operativa o el incentivo

La razón de que esta distinción rara vez se trace es que incomoda. Una falla de comprensión es una historia sobre una brecha de entendimiento, que un buen briefing cierra y que no avergüenza a nadie de forma permanente. Una falla de cálculo es una historia sobre un modelo operativo cuyos incentivos hacen que el ejecutivo racional elija el fracaso, lo que implica a la estructura de la organización y a las personas que la construyeron. Los directorios echan mano del remedio de comprensión porque es el barato, y lo aplican a fallas de cálculo, donde no produce nada salvo otro año de pilotos y otro cero. El remedio barato aplicado a la condición estructural es la mayor parte de lo que está hecho el 95 % del MIT.

08

LO QUE ESCONDE UN ÉXITO REPORTADO

Si el 95 % no es una sola población, el 5 % tampoco. La cohorte que reporta ROI aparente es también una mezcla, y un lector técnicamente alfabetizado debería ser tan escéptico de un triunfo reportado como de una pérdida reportada. Tres patrones producen resultados que parecen éxito sin serlo, y cada uno tiene una sola pregunta que lo expone.

El primero es **actividad disfrazada de resultado**. Una métrica de adopción se reporta como impacto: el 80 % de los planificadores usa el copiloto semanalmente, por lo tanto el copiloto triunfó. El movimiento confunde uso con valor y sobrevive solo porque nadie midió el resultado antes del despliegue. La pregunta detectora es *¿cuál era la línea base?* Si no existe una medición pre-despliegue de aquello que la herramienta debía mejorar, el triunfo reportado es un conteo de actividad usando el disfraz de un resultado. El copiloto de planificación de la historia central es exactamente este patrón.

El segundo es **anécdota nunca multiplicada**. Un caso atractivo se recontó hasta quedar en lugar del portafolio. El asistente de conciliación del superintendente es real y valioso, y un vendor o un campeón interno lo mostrará feliz en cada lámina mientras las otras once iniciativas que no devolvieron nada se quedan fuera de la presentación. El caso es verdadero. La generalización a partir de él es falsa. La pregunta detectora es *¿cuál es el desempeño del portafolio?* Un éxito individual prueba que la tecnología puede funcionar; no dice nada sobre si tu portafolio, en promedio, funciona. Exige el denominador, no el highlight reel.

El tercero es **indicador mal atribuido**. Una métrica se movió genuinamente y la iniciativa de IA reclamó el crédito, pero varias causas cambiaron en la misma ventana. La ganancia de 1,8 % de throughput del optimizador de molienda es este patrón: el cambio de revestimientos y la nueva mezcla de mineral están sentados en los mismos dos trimestres, cada uno candidato a explicar el delta. La pregunta detectora es *¿qué más cambió en el mismo período?* Si dos o tres causas plausibles se superponen al despliegue de IA y nadie las aisló, la atribución es una adivinanza vestida de medición.

Estas tres preguntas convierten al lector de consumidor del 95 % en investigador del 5 %. *Cuál era la línea base. Cuál es el desempeño del portafolio. Qué más cambió en el mismo período.* Córrelas contra cualquier triunfo de IA que te presenten — el tuyo propio o la historia de victoria de un par en la última cena de industria — y una parte significativa del 5 % se revela como la misma brecha de comprensión vista desde el otro lado: valor reclamado sin la instrumentación para defenderlo.

09

QUÉ PREGUNTAR ANTES DE LA PRÓXIMA REVISIÓN DE PORTAFO- LIO

El diagnóstico no es teórico. Es un test que puedes correr sobre una sola iniciativa antes de tu próxima revisión de portafolio, con resultado binario: el retorno reportado es defendible o no lo es, y el tipo de falla es comprensión o cálculo. Elige una iniciativa estancada o una celebrada y corre cinco pasos en orden.

Paso uno. Nombra la línea base que se midió antes del despliegue. Si no existe un número pre-despliegue para el resultado que la herramienta debía mejorar, detente: el resultado reportado no está probado, y la condición es comprensión (nadie hizo visible el costo de medir). Si existe línea base, continúa.

Paso dos. Establece si el resultado reportado es a nivel de portafolio o un caso individual. Si es una anécdota representando al programa, el desempeño del portafolio no está probado por buena que sea la anécdota. Si es un número de portafolio, continúa.

Paso tres. Lista todo lo demás que cambió en la misma ventana: trabajos mecánicos, cambios de receta, movimientos de mercado, un supervisor nuevo. Si dos o más causas plausibles se superponen al despliegue y no se aislaron, la atribución no está probada. Si la causa está aislada, continúa.

Paso cuatro. Pídele al dueño que nombre, en una frase, el cambio de modelo operativo necesario para capturar el valor: qué decisión se mueve, qué dato se abre, quién pasa a ser responsable. Si el dueño no puede nombrarlo, la condición es comprensión: hazle el briefing, financia el andamiaje, acelera. Si el dueño lo nombra con precisión y luego explica por qué no ha ocurrido, continúa.

Paso cinco. Pregunta por qué el cambio nombrado no se ha hecho. Si la respuesta es “nadie nos mostró que hacía falta”, es comprensión después de todo. Si la respuesta es “lo sabemos, pero el costo de mover esos derechos de decisión excede lo que recuperaríamos”, la condición es cálculo: deja de tratarlo como un problema de IA, porque ningún briefing lo va a mover. Se mueve solo cuando cambia el incentivo o la autoridad.

Cinco pasos, una iniciativa, menos de una hora. El resultado es binario en cada compuerta y el camino a través de las compuertas clasifica la falla. Una iniciativa que pasa las cinco compuertas tiene un retorno defendible. Una que falla en el paso uno, dos o tres tiene un retorno no probado y una brecha de comprensión. Una que llega al paso cinco y se detiene ahí tiene una brecha de cálculo, y lo más caro que puedes hacer es entregarle otro año de pilotos.

Si no puedes correr este test sobre tu propio portafolio, todavía no sabes qué está haciendo tu programa de IA, y tu directorio tampoco. Ese es el hallazgo real que la cifra del MIT debería provocar.

10

LOS REMEDIOS QUE TRATAN LA ENFERMEDAD EQUIVOCADA

Con una clasificación en mano, los modos de falla cambian de “hicimos el proyecto equivocado” a “aplicamos el remedio equivocado al diagnóstico correcto”. Tres de estos son lo bastante comunes en la industria pesada de LATAM como para nombrarlos.

El primero es recetar arquitectura para una pregunta de modelo operativo. Aparece una falla de cálculo, y la respuesta es un lakehouse, una malla de agentes, una capa de retrieval, una plataforma nueva. La arquitectura se vende como la respuesta. No lo es, porque la condición nunca fue técnica. Los derechos de decisión siguen fragmentados después de que la plataforma se entrega, la función dueña de los datos sigue sin compartirlos, y el ejecutivo que ya cotizó el rediseño y lo declinó ahora tiene un programa más caro devolviendo el mismo cero. La arquitectura no puede comprar autoridad. Si el diagnóstico es cálculo, el primer movimiento es cambiar quién decide o qué se le premia, y solo después construir.

El segundo es tratar el 95 % como homogéneo y pausarlo todo. El directorio lee la cifra, aplica una sola decisión a un portafolio que necesita tres, y detiene las iniciativas con brecha de comprensión — que estaban a un ajuste de alcance de empezar a componer — junto con las de brecha de cálculo que genuinamente deben parar. El costo de detener un programa con brecha de comprensión no es el presupuesto castigado; es el aprendizaje organizacional descartado, la fuerza laboral que ahora carga anticuerpos contra la próxima iniciativa de IA, y un reinicio desde cero contra un paisaje de vendors dos generaciones más adelante. Rebana el portafolio antes de pausarlo. El promedio no es una decisión.

El tercero es archivar el costo de modelo operativo como factor blando. “Cultura”, “gestión del cambio”, “formas de trabajo”: el lenguaje degrada una línea de costo dura a un taller. La persona que lee esto sabe que los derechos de decisión, la propiedad de los datos y la responsabilidad no son blandos. Son los determinantes de si algún output de IA se convierte en acción, y pertenecen al caso de negocio como un costo de captura cuantificado, no a un anexo sobre adopción. Cuando el cambio de modelo operativo se cotiza como línea de presupuesto, la falla de cálculo se vuelve visible como lo que es: una elección racional que el ejecutivo hizo porque nadie puso nunca el costo del rediseño y el ROI renunciado en la misma página.

El patrón a través de los tres es idéntico. El remedio se elige para la enfermedad para la que la literatura de management entrenó el reflejo, no para la enfermedad que el diagnóstico encontró. Trata la condición que tienes. El costo de tratar la otra es el cero del próximo año.

11

QUÉ CONFIRMAR EN TU PROPIO PORTAFOLIO ESTE TRIMESTRE

Cuatro verificaciones, ordenadas de la más barata a la más consecuente. Cada una clausura una manera en que el promedio del MIT miente sobre tu empresa.

Uno. Corre el test de cinco pasos de la Sección 09 contra cada iniciativa de IA activa, de a una, y deja el resultado por escrito: defendible o no probado, comprensión o cálculo. El costo es aproximadamente una hora por iniciativa. El producto es una clasificación de tu portafolio completo, por escrito, con evidencia — que es precisamente la distribución que el promedio del MIT te niega. Para el próximo ciclo de directorio puedes reemplazar “ROI del portafolio de IA = cero” por una frase que signifique algo.

Dos. Para cada iniciativa clasificada como comprensión, reescribe el contrato de cuatro partes en una sola página con el comprador, el que despliega y un usuario de línea en la sala: qué hace la herramienta, qué no hace, qué cuenta como éxito, quién actúa sobre el output. Medio día cada una. La página firmada es la cura de una brecha de comprensión, y es barata.

Tres. Para cada iniciativa clasificada como cálculo, sácala por completo de la conversación de IA y ponla frente al ejecutivo dueño de la cadena operativa, planteada como pregunta de costos: el costo del rediseño a un lado, el retorno renunciado al otro, en la misma página. El programa de IA no se reinicia hasta que esa página se haya leído y se haya tomado una decisión. La honestidad de esa escalada es el test. Un programa de IA no debe ser nunca la historia de tapa de un modelo operativo que la organización declinó arreglar.

Cuatro. Audita la contabilidad buscando valor registrado bajo el nombre equivocado. Para cada iniciativa que marca cero, pregunta si una mejora apareció en otra parte del estado de resultados — un cierre más rápido, menos detenciones, horas de analista recuperadas — y si esa mejora sigue la línea de tiempo del despliegue. Donde la siga, re-etiquétala a partir de ahora y explica el re-etiquetado en la próxima revisión. Cuenta con que finanzas se resista a reabrir un cierre ya firmado —por eso el re-etiquetado se aplica a partir de ahora,

no retroactivamente. Esta es la única verificación que puede revisar tu número reportado al alza este trimestre, y el asistente de conciliación de la historia central es la prueba de que el valor muchas veces ya está ahí, archivado bajo otra cosa.

Cuatro verificaciones, un trimestre. Al final sabrás qué iniciativas acelerar, cuáles re-medir, cuáles detener y cuáles re-registrar contra la línea de valor que siempre estuvieron produciendo. Gobernarás tu portafolio de IA por su distribución en lugar de por un promedio que no podías interpretar, que es la única manera honesta de responder la pregunta que el directorio realmente hizo.

■ *El 95 % del MIT es un número real midiendo una cosa irreal: el promedio de dos condiciones opuestas que producen un mismo síntoma. Te dice que la IA no devolvió nada que pudieras ver. No te dice si la cura es un mejor briefing o un modelo operativo distinto — y esas curas son opuestas. El número es el empujón. El diagnóstico es el entregable.*

12

IDEAS CLAVE

Las tesis de este documento, cada una capaz de sostenerse sola — un bloque por idea, como ayuda de memoria.

1 El 95 % del MIT no es una medición del fracaso de la IA. Mide la distancia entre lo que una empresa creyó comprar y lo que la tecnología exige antes de pagar — una brecha de comprensión, no un veredicto tecnológico.

2 Lee el estudio antes de construir sobre él: 52 entrevistas, 153 encuestas de conferencia, 300+ iniciativas públicas con método de síntesis no publicado, ROI definido por cada respondente. Los propios autores lo llaman preliminar y direccional. Direccional es como debe usarse.

3 Los desafíos metodológicos al estudio — incluido un llamado público desde Wharton a publicar los datos o retractar — quedaron sin respuesta hasta mediados de 2026. Un lector serio cita la cifra con sus límites adjuntos, o no la cita.

4 El promedio de tu portafolio es el número menos útil de la sala. Una línea de ROI cero puede esconder tres iniciativas que necesitan tres decisiones opuestas: acelerar, re-medir, detener. El promedio no es una decisión.

5 Cada compra de GenAI firma cuatro versiones de un contrato no escrito — lo que el vendedor vende, el comprador compra, la planta recibe y finanzas mide. El 95 % es la distancia entre las cuatro firmas.

6 La burbuja de comprensión no es la burbuja financiera, y una puede ser cierta sin la otra: tu cero persistiría incluso con las acciones de IA bien valoradas, y una corrección de mercado no cerraría tu brecha operativa. Defiende la primera; no tomes posición sobre la segunda.

7 Una falla de comprensión responde a información: haz visible el costo de modelo operativo y el dueño redefine el alcance y acelera. Una falla de cálculo responde solo a incentivos o autoridad cambiados — el dueño ya vio el costo y lo declinó racionalmente.

8 El remedio barato aplicado a la condición estructural — el briefing al ejecutivo que ya hizo la aritmética — es la mayor parte de lo que está hecho el 95 % del MIT.

9 El 5 % tampoco es una sola población. Tres patrones fabrican éxito: actividad disfrazada de resultado, anécdota nunca multiplicada, indicador mal atribuido. Tres preguntas los exponen: cuál era la línea base, cuál es el desempeño del portafolio, qué más cambió en el mismo período.

10 Cinco compuertas, una hora, una iniciativa: línea base → nivel portafolio → causas aisladas → cambio de modelo operativo nombrado → razón por la que no ha ocurrido. El camino por las compuertas clasifica la falla y dicta el primer movimiento.

11 La arquitectura no puede comprar autoridad. Un lakehouse entregado a una falla de cálculo es un programa más caro devolviendo el mismo cero. Cambia quién decide o qué se le premia, y después construye.

12 Audita la contabilidad buscando valor registrado bajo el nombre equivocado — un cierre más rápido, horas de analista recuperadas, menos detenciones siguiendo la línea de tiempo del despliegue. Es la única verificación que puede revisar tu número al alza este trimestre.

13

FUENTES

Las afirmaciones públicas de este documento, con sus fuentes. La historia central y las cifras de portafolio son composiciones anonimizadas a partir de engagements de recursos naturales en LATAM; ilustran el diagnóstico, no a un cliente con nombre.

- MIT Project NANDA (“Networked Agents and Decentralized AI”), *The GenAI Divide: State of AI in Business 2025*, agosto 2025 — metodología, el hallazgo del 95 %, y los descargos de los propios autores (“preliminary findings”, “directionally accurate”) — MIT Project NANDA.
- Lanzamiento de prensa y cobertura ante audiencia CFO del reporte (18 de agosto de 2025) — Fortune.
- Mike Kaput, “That Viral MIT Study Claiming 95 % of AI Pilots Fail? Don’t Believe the Hype,” Marketing AI Institute, 26 de agosto de 2025 — la crítica del retelling sin calificadores y la objeción a la barra de éxito de seis meses — Marketing AI Institute.
- “Why We Don’t Believe MIT NANDA’s Weird AI Study,” Futuriom, 26 de agosto de 2025 — objeciones metodológicas, incluido el llamado del profesor de Wharton Kevin Werbach a publicar los datos completos o retractar — Futuriom.
- David Gallacher, “Beyond ROI: Are We Using the Wrong Metric in Measuring AI Success?,” UC Berkeley Executive Education, septiembre 2025 — el argumento de que el propio ROI está mal calibrado para la curva de capacidad GenAI — UC Berkeley Executive Education.

NTT DATA · Technology Consulting · 2026 · biztalksnttdata.com

Para más sobre este tema, contacta a: **Jaime Rebolledo**, Director Value Proposition Mining, NTT DATA · jaime.rebolledocampos@nttdata.com.