



La estrategia que crea valor

**El invitado que quiere ser el
anfitrión:**

por qué la inteligencia artificial

no es una revolución más

Por qué los trabajadores tienen razones para que sus proyectos fracasen

El invitado que quiere ser el anfitrión:

Por qué la inteligencia artificial no es una revolución más

La idea en síntesis

La adopción de IA está revelando un fracaso silencioso. Solo 5% de los pilotos genera retorno medible. La explicación estándar habla de inmadurez tecnológica; falta una más incómoda: los trabajadores no tienen incentivo para que estos proyectos funcionen.

Esta es la primera revolución tecnológica que ataca el último activo del trabajador del conocimiento: su juicio. Y es la primera que necesita aprender de él para instalarse.

La metáfora del «AI coworker» es engañosa: la IA no puede ser leal ni desleal, no asume responsabilidades, no firma contratos. Tratarla como persona es parte del problema.

Mientras la asimetría entre quien aporta el dato y quien captura el valor no se resuelva, los trabajadores seguirán cooperando lo justo para no ser sancionados, y saboteando lo justo para no quedar expuestos.

Resumen ejecutivo

Las cifras de fracaso de los proyectos IA son persistentemente altas y notablemente uniformes a través de industrias y geografías. Un reporte del MIT de 2025 encontró que solo alrededor del 5% de los pilotos de IA generativa logra aceleración medible de ingresos.¹ En servicios financieros, análisis de FinTellec AI indica que el 80% de los proyectos de IA falla en alcanzar producción, y de los que sí llegan, el 70% no entrega valor de negocio medible.² La industria explica esto como inmadurez tecnológica, calidad de datos, curva de aprendizaje, falta de alineamiento estratégico, problemas de integración. Explicaciones que son verdaderas y a la vez insuficientes.

Hay otras explicaciones más incómodas pero que deben empezar a aparecer en los análisis. Los trabajadores cuyos flujos están siendo modelados por proyectos de IA no tienen ningún incentivo para que esos proyectos funcionen. Las organizaciones les piden que documenten sus procesos, que expliciten sus criterios de decisión, que aporten datos sobre cómo resuelven los casos difíciles, información necesaria para entrenar los sistemas que, una vez maduros, justificarán optimizaciones de personal. El trabajador es simultáneamente la fuente del dato y el destinatario probable del impacto. Es un trato visiblemente injusto por lo que no debería sorprender que parte del fracaso de los proyectos de IA sea —al menos en parte, probablemente de manera inconsciente— intencional.

Esta hipótesis es incómoda al cuestionar la imagen benévola de la transformación digital, pero no es infundada y empieza a contar con respaldo en la literatura organizacional más reciente. Varias investigaciones recientes —entre las que destacan Bankins, Ocampo, Marrone y Restubog (2024), Kinowska y Sienkiewicz (2023), y un meta-análisis publicado en *Frontiers in Psychology* bajo el nombre «Algorithmic anxiety»— revelan que los empleados no perciben la introducción de IA como un cambio estratégico más, sino como una violación de confianza cualitativamente distinta.³ El meta-análisis, basado en mil cuatrocientas cincuenta y cuatro narrativas digitales de trabajadores, registra que estos reportan sentirse «traicionados por un algoritmo»: las decisiones impulsadas por IA se perciben impersonales y arbitrarias, careciendo del elemento humano que tradicionalmente amortiguaba las transiciones organizacionales difíciles.

Lo que las revoluciones hasta ahora habían respetado

Vale la pena hacer el ejercicio histórico, porque ayuda a calibrar de qué tamaño es el cambio que tenemos enfrente.

La revolución agrícola sustituyó la fuerza humana por la fuerza animal y, después, por la fuerza mecánica. La revolución industrial sustituyó la habilidad manual por la precisión de la máquina. La línea de ensamblaje sustituyó la coordinación humana por la coordinación del proceso. La automatización electromecánica del siglo veinte sustituyó la velocidad humana por la velocidad de la máquina. La revolución digital sustituyó la memoria humana por la memoria de los sistemas, y los cálculos humanos por los cálculos del computador.

Cada una de estas revoluciones desplazó a una porción enorme de la fuerza laboral. Pero ninguna de ellas traspasó la frontera de la inteligencia, esa capacidad que distingue al ser humano de cualquier otra especie y de cualquier máquina construida hasta ahora. Cada revolución sustituyó una capacidad específica y periférica a esa frontera: fuerza, precisión, velocidad, capacidad de repetición, capacidad de almacenamiento, capacidad de cálculo. La inteligencia, entendida como la capacidad de interpretar situaciones nuevas, de ejercer juicio bajo incertidumbre, de articular sentido y de tomar decisiones que requieren ponderación, permaneció intocada precisamente porque era el rasgo distintivamente humano que ninguna tecnología sabía replicar. Más aún: cada revolución hizo a la inteligencia humana progresivamente más valiosa, porque al liberarla de las tareas que la máquina absorbía, la concentraba en el territorio donde la máquina no podía competir.

Peter Drucker anticipó este movimiento con notable lucidez. Desde finales de los años cincuenta, y luego con creciente sistematicidad en *The Age of Discontinuity* (1969), empezó a hablar de los «trabajadores del conocimiento» como la categoría laboral que iba a definir el siglo veintiuno.⁴ Su descripción era precisa: el trabajador del conocimiento posee un activo radicalmente distinto

al del trabajador industrial. Ese activo no está en la fábrica ni en las máquinas ni en los procesos de la empresa. Está en su cabeza. Es portátil, le pertenece, puede llevárselo cuando cambia de empleo. Esto, según Drucker, alteraba la estructura misma de la relación laboral: la empresa ya no era dueña del medio de producción de su trabajador más valioso, y eso exigía un nuevo trato. Más autonomía, más movilidad, más reconocimiento del conocimiento como propiedad del trabajador y no de la organización.

Durante cuarenta años, las economías desarrolladas se reorganizaron progresivamente alrededor de los trabajadores del conocimiento. Sus condiciones laborales mejoraron, su movilidad aumentó, su poder de negociación creció.

La inteligencia artificial es la primera tecnología en la historia que busca traspasar la frontera

Ese conocimiento, y por extensión los trabajadores que lo encarnaban, constituían la frontera entre el recurso humano y el no humano dentro de las organizaciones. Las tecnologías de la información comenzaron a rondar esa frontera de manera cada vez más amenazante, pero nunca pasaron de ahí: una amenaza. La IA es la primera tecnología que la cruzó —y no con ánimo exploratorio, sino con la intención de asentarse en ese dominio. Eso es lo que vuelve a esta revolución distinta de todas las anteriores: no por la magnitud de la amenaza, sino por la ausencia de precedente.

Una revolución que necesita del trabajador para instalarse

Hay un segundo rasgo distintivo de esta revolución, y es probablemente el más decisivo para entender por qué se está atascando. Todas las revoluciones anteriores compartían una característica que se daba por evidente: no necesitaban del trabajador desplazado para instalarse. La máquina industrial no aprendía del operario manual: simplemente lo reemplazaba mediante un diseño exterior a él. La línea de ensamblaje no consultaba al artesano: imponía un nuevo proceso por encima del antiguo. El computador no le pedía al contable que le enseñara a calcular: traía el cálculo desde fuera. En cada caso, la tecnología llegaba ya formada y el trabajador podía resistirla, pero no condicionar su instalación.

La inteligencia artificial es distinta. Por más que existan datos históricos abundantes y sistemas tecnológicos robustos detrás de ella, la IA llega a aprender. Necesita aprender cómo las personas realizan sus funciones dentro de la organización: qué consideran al decidir, qué excepciones manejan, qué criterios aplican para abordar lo que no existe en los manuales. La IA reproduce — de manera nunca exacta y solo después de varias iteraciones— la forma efectiva en que se opera,

no la forma declarada. La única fuente para obtener esa información son los trabajadores, los mismos que la IA está destinada a desplazar.

El caso más ilustrativo son los call centers. Las empresas anuncian que graban las llamadas para mejorar el servicio, registrando la voz del cliente, pero también la del trabajador, porque con esa voz entrenan a los bots que la siguiente vez que el cliente llame contestarán automáticamente.⁵ El trabajador del call center es simultáneamente el operador del sistema actual, la fuente del dato que entrena al sistema futuro, y la víctima probable de la optimización que ese sistema futuro habilitará. Estos casos terminan rompiendo el contrato psicológico, que hace que la persona encuentre legítimo no comportarse de la forma más ética, ya que entiende que la empresa tampoco se ha comportado de esa manera.

*Esta vez, la tecnología no se impone al trabajador desde fuera:
depende de él para existir.*

Esto invierte un equilibrio que se daba por descontado en todas las revoluciones tecnológicas anteriores. La revolución no ocurre sin la colaboración del activo que pretende sustituir. Y eso le otorga al trabajador del conocimiento un poder de negociación que la mayoría de las empresas todavía no ha reconocido.

El invitado que quiere ser anfitrión

La industria tecnológica promovió durante años la imagen del «AI coworker»: la IA como compañero de trabajo, como nuevo integrante del equipo, como empleado digital al que basta con dar onboarding adecuado. En los últimos meses, esa metáfora ha empezado a extinguirse. *Harvard Business Review* publicó en mayo de 2026 un artículo que sostiene que tratar a los agentes de IA como empleados es precisamente donde la mayoría de los programas empresariales de agentes está fallando.⁶ Una investigación recogida por *Reworked* en marzo de 2026 lo formula con más precisión: el momento en que empezamos a preguntar a quién es «leal» una IA, ya diseñamos el sistema incorrectamente.⁷ La metáfora del compañero de trabajo implica reciprocidad: un compañero puede mirar el output, decir «esto no me cuadra» y negarse a firmarlo hasta que se revise. Un agente de IA no hace nada de eso: entrega, y la responsabilidad queda íntegra en quien lo usó.

Este vacío de reciprocidad no es solo conceptual; es ahora también jurídico. El Stanford CodeX, junto con law.MIT.edu, el Stanford HAI Digital Economy Lab y Consumer Reports, abrieron en 2025 un workshop específico sobre qué significa que un agente de IA sea «leal» a su usuario. La premisa explícita: a diferencia de la relación entre un individuo y un médico, un abogado o un

asesor financiero, no existe ningún requerimiento de que los sistemas de IA actúen de manera consistente con los mejores intereses del usuario.⁸ El trabajo que introduce el concepto técnico de «AI loyalty» lo plantea claramente: a diferencia de un médico o un abogado, cuya lealtad al paciente o al cliente viene impuesta por el código profesional y la ley, la de un agente de IA hay que escribirla línea por línea en el sistema. Si nadie la programa, no existe; y por defecto el agente sirve a quien lo desplegó, no a quien lo usa.⁹

La IA no es un compañero. No es un agente con voluntad, no puede asumir responsabilidades, no puede firmar un contrato, no puede ser leal ni desleal en el sentido humano de esas palabras. La IA es una tecnología que, por ahora, opera mejor cuando se la trata como tecnología y peor cuando se la trata como persona. Tratarla como persona es, paradójicamente, parte del problema: induce a las organizaciones a esperar de ella reciprocidades que no puede ofrecer, y a omitir respecto de ella controles que sí debería tener.

Pero hay una particularidad en esta tecnología que la distingue de cualquier otra. Una máquina industrial ocupa el lugar de la mano humana, pero la mano humana sigue siendo necesaria para operar la máquina. Un computador ocupa el lugar de la memoria humana, pero la memoria humana sigue siendo necesaria para programar el computador. La IA, en cambio, no es una herramienta que el trabajador del conocimiento usa para hacer mejor su trabajo. Es una tecnología que aspira a hacer el trabajo en lugar del trabajador del conocimiento. Y mientras más eficaz se vuelve, menos necesario se vuelve el trabajador.

El invitado no necesita querer ocupar el lugar del anfitrión para terminar ocupándolo. Basta con que el incentivo del anfitrión sea cederle progresivamente el espacio.

De ahí la figura del invitado que quiere ser el anfitrión. La organización invita a la IA a entrar como una capacidad técnica adicional, como un asistente, como un acelerador de procesos. Pero la lógica económica que rodea a la IA empuja en una dirección distinta: cada tarea humana que la IA puede absorber, la absorberá, porque el sistema económico premia esa absorción. Referirse a la IA como coworker no describe lo que realmente ocurre. La realidad muestra que la introducción de la IA se parece más a una sustitución gradual y silenciosa, donde la inteligencia humana es invitada a transferir progresivamente sus criterios, sus procedimientos y sus juicios a un sistema que, una vez que los ha internalizado, podrá prescindir de ella.

Por qué los trabajadores tienen razón en no cooperar

Si los trabajadores del conocimiento de una organización entienden —incluso intuitivamente, incluso sin formularlo en estos términos— que están siendo invitados a entregar el activo que Drucker identificó como portátil, propio y constitutivo de su valor laboral, ¿qué se puede esperar de su cooperación con los proyectos de IA?

La respuesta racional es: cooperación visible, ambigua y blanda. Visible porque la organización lo exige y porque resistir abiertamente tiene costos profesionales. Ambigua porque cada trabajador puede decidir cuánto de su conocimiento real entrega al sistema y cuánto retiene como ventaja personal. Y blanda porque los datos que se aportan pueden ser técnicamente correctos pero estratégicamente pobres: documentar el flujo formal pero no las excepciones, explicitar los criterios oficiales pero no los criterios reales, dar lo necesario para que el proyecto avance pero no lo suficiente para que reemplace al trabajador.

Un estudio en *Frontiers in Public Health* (2024), «Assist me or replace me?», documenta el patrón: el trabajador que se siente amenazado por la IA deja de invertir en la relación laboral —documenta menos, comparte menos, propone menos— incluso sin tomar la decisión consciente de hacerlo.¹¹ En septiembre de 2025, Cureus le dio nombre clínico: Artificial Intelligence Replacement Dysfunction (AIRD), el cuadro asociado a convivir con un sistema diseñado para sustituirte.¹²

Este fenómeno no requiere mala fe ni conspiración. Requiere simplemente que cada trabajador del conocimiento opere bajo la lógica que Drucker describió hace medio siglo: su conocimiento es su activo. Entregarlo voluntariamente y sin contraprestación a un sistema diseñado para sustituirlo no es una conducta cooperativa: es una conducta autodestructiva. Que los trabajadores no la ejecuten en su versión maximalista no es un problema técnico ni un problema de gestión del cambio, es una respuesta racional a un incentivo perversamente diseñado.

Como ninguna revolución tecnológica anterior había dependido del trabajador para instalarse, las empresas no tienen experiencia diseñando incentivos para esta situación. Aplican los manuales de la transformación digital tradicional —liderazgo ejecutivo, gestión del cambio, capacitación—, manuales que fueron escritos para tecnologías que se imponían desde fuera. La empresa puede invertir cien millones de dólares en una iniciativa de IA, pero si el trabajador que tiene el conocimiento real decide entregarlo a medias, el sistema resultante reproducirá una versión empobrecida de la operación y la promesa de eficiencia nunca aparecerá en la implementación.

Las empresas hablan de «upskilling» y «reskilling»: son paliativos cosméticos cuando el problema de fondo es la asimetría entre quién aporta el dato y quién captura el valor.

Si esto es así, parte de la baja tasa de éxito de los proyectos de IA no se resuelve con mejor tecnología, mejor metodología o mejor capacitación. Se resuelve solo si la organización logra ofrecer a sus trabajadores del conocimiento un trato que les dé una razón legítima para cooperar con un sistema que, sin ese trato, los va a perjudicar.

Mientras esta asimetría no se aborde, los trabajadores seguirán haciendo —consciente o inconscientemente— lo único racional: cooperar lo justo para no ser sancionados, y sabotear lo justo para no quedar expuestos.

Lo que viene

Hasta aquí, el problema descrito es económico y contractual: una asimetría entre quién aporta el dato y quién captura el valor, en una revolución tecnológica que por primera vez necesita la colaboración del trabajador desplazado.

Crear que los procesos de incorporación de IA son una variante más de los procesos de cambio que las organizaciones conocen es un error. Las empresas tienen experiencia gestionando transformaciones operacionales, despliegues tecnológicos, reestructuraciones, fusiones y programas de mejora continua. Cada uno de esos cambios cuenta con metodologías y experiencias extensas. Todas comparten supuestos comunes: el cambio se diseña desde la dirección de la empresa, se comunica a la organización, se implementa con los líderes y el apoyo de los equipos, y se mide por su grado de adopción. Ese modelo funciona porque la organización es, simultáneamente, dueña del cambio y dueña del conocimiento que el cambio requiere.

La incorporación de IA rompe el supuesto que sostenía a todos los demás: que la organización es dueña del conocimiento que el cambio requiere. En la IA no lo es. Ese conocimiento vive en sus trabajadores, y entregarlo no es neutral para ellos. Por primera vez, el cambio depende de que el activo que va a ser sustituido coopere con su propia sustitución.

Lo que esta tecnología requiere es un proceso propio. Un proceso que reconozca que está incorporando un agente que no debe lealtad a la organización y que aprende de personas con incentivos contrarios a entregarle su conocimiento. Ese proceso no existe todavía en la mayoría de las organizaciones. Construirlo es probablemente la tarea más importante que enfrentarán las organizaciones en los próximos años.

Referencias

1. MIT (2025), reporte sobre adopción empresarial de IA generativa, citado en LeadDev, *Separating AI hype from engineering reality*, febrero de 2026.
2. FinTellec AI, análisis sobre proyectos de IA en servicios financieros, recogido en *CIO Magazine*, «Beyond the hype: 4 critical misconceptions derailing enterprise AI adoption», enero de 2026.
3. Bankins, S., Ocampo, A. C., Marrone, M., & Restubog, S. L. D. (2024); Kinowska, H., & Sienkiewicz, Ł. J. (2023); meta-análisis «Algorithmic anxiety», *Frontiers in Psychology*, basado en mil cuatrocientas cincuenta y cuatro narrativas digitales de trabajadores.
4. Drucker, P. F., *The Age of Discontinuity: Guidelines to Our Changing Society*, Harper & Row, 1969; desarrollado posteriormente en *Post-Capitalist Society* (1993) y *Management Challenges for the 21st Century* (1999).
5. Rodríguez, M. L., Universidad de Castilla-La Mancha, intervención en el Mobile World Capital Barcelona, mayo de 2026, recogida por *Infobae*.
6. *Harvard Business Review*, artículo del 6 de mayo de 2026 sobre la metáfora del «AI coworker» y el fracaso de programas empresariales de agentes; recogido posteriormente por Boston Consulting Group y Metaintro.
7. *Reworked*, marzo de 2026, sobre el problema conceptual de atribuir lealtad a sistemas de IA.
8. Stanford CodeX, law.MIT.edu, Stanford HAI Digital Economy Lab y Consumer Reports, workshop sobre AI loyalty, 2025.
9. Aguirre, A., et al., *AI Loyalty: A New Paradigm for Aligning Stakeholder Interests*, arXiv:2003.11157.
10. Boston Consulting Group, *AI at Work 2025*, encuesta a más de diez mil trabajadores en once países.
11. «Assist me or replace me? The role of artificial intelligence awareness in employee psychological contract and counterproductive work behavior», *Frontiers in Public Health*, 2024.
12. «Artificial Intelligence Replacement Dysfunction (AIRD)», *Cureus*, septiembre de 2025.