AGENTES DE IA · FUNDACIONES OPERACIONALES

El Problema del Sustrato

Por Qué los Agentes de IA Fallan en la Industria Pesada — y Qué Deben Arreglar Primero los CIOs

Un argumento de fundación-antes-que-agentes para líderes de minería, cemento, celulosa y agroindustria en LATAM

AUDIENCIA
CIOs, CDOs, VPs de Operaciones

FOCO
Preparación pre-agentes

REGIÓN
LATAM · Minería · Cemento · Celulosa · Agro

NTT DATA · Natural Resources Practice · 2026 · biztalksnttdata.com

Contenido

01

RESUMEN EJECUTIVO

Las industrias de recursos naturales en América Latina están en los primeros capítulos de una ola de despliegue de agentes de IA: sistemas autónomos o semiautónomos que leen datos operacionales, razonan sobre ellos, y toman o recomiendan acciones. Los pitches de los proveedores son confiados, la presión de los directorios es intensa, y las pruebas de concepto se multiplican. Por debajo, se está formando una contradicción. Los presupuestos de agentes suben. El ROI de los agentes, no.

Hemos observado cuatro agentes de IA desplegados en producción en operaciones mineras LATAM hasta 2025. Los dos que contaban con fundaciones de telemetría limpias entregaron retornos medibles dentro de un trimestre fiscal. Los dos que no, se convirtieron en pasivos de cumplimiento. Uno fue desactivado en silencio después de recomendar un setpoint de recuperación basado en un registro de calibración que llevaba siete años vencido. El otro generó un año de artefactos de decisión cuyo audit trail no pudo reconstruirse cuando un evento de seguridad gatilló una revisión regulatoria.

La diferencia no fue el modelo. Fue el sustrato.

Este paper argumenta una posición simple, impopular entre los proveedores e incómoda para los directorios: **la mayoría de las operaciones de recursos naturales no está lista para desplegar agentes autónomos, porque la fundación digital de la que esos agentes dependen es estructuralmente deficiente.** La calibración de sensores no se mantiene. Los datos maestros están en disputa entre IT[1] y OT[2]. La telemetría de series de tiempo tiene gaps silenciosos. La metadata de proceso vive en las bitácoras de los operadores y en las cabezas de los supervisores, no en sistemas consultables. Y la confianza institucional en el "tenemos los datos" es, en la mayoría de los casos, una herencia de las eras anteriores de dashboards y reportes, donde los datos malos todavía producían salidas de aspecto útil porque había un humano leyéndolas.

Los agentes son distintos. Leen a velocidad de máquina, actúan con confianza de máquina, y documentan sus decisiones en formatos que ningún investigador puede reconstruir después sin el sustrato de origen. No arreglan los datos malos. Los amplifican.

La recomendación no es abandonar los agentes de IA. La recomendación es secuenciar: **primero la fundación, después los agentes.** Este paper describe en qué consiste realmente esa fundación, por qué recursos naturales es especialmente vulnerable a saltarse el paso, dónde están realistamente la mayoría de las operaciones en el espectro de madurez, y qué deberían preguntar los ejecutivos antes de firmar el próximo piloto de agentes.

[1] Information Technology — los sistemas de computación corporativos (ERP, correo, analytics, nube) tradicionalmente propiedad de la organización del CIO.

[2] Operational Technology — los sistemas de control que corren los procesos físicos (DCS, PLC, SCADA, historian, analizadores en línea) tradicionalmente propiedad de ingeniería de planta o mantenimiento.

02

QUÉ SIGNIFICA “MALA CALIDAD DIGITAL” EN LA PRÁCTICA

La frase “calidad de datos” quedó desgastada por una década de consultoría de analytics. Para la preparación de agentes, hay que desempacarla en las superficies de falla específicas que importan.

Deuda de calibración de sensores. Toda medición física depende de un sensor que deriva: poco por día, mucho por año. La calibración se actualiza casi siempre de forma reactiva, cuando un sensor se rompe de manera evidente, no según calendario. Un agente que lee una sonda de pH calibrada verificablemente por última vez en 2019 produce una recomendación tan confiada como uno que lee un sensor calibrado esta mañana. El sustrato no le da ninguna forma de descontar la diferencia.

Fragmentación de datos maestros. El mismo activo físico carga rutinariamente identificadores distintos en el ERP[3], el MES[4], el CMMS[5], el historian y el LIMS[6], especialmente en operaciones que han absorbido adquisiciones y cambios de proveedor a lo largo de veinte años. Un agente que cruza datos entre estos sistemas va a fallar en silencio en los cruces que no están alineados, y va a acertar en la mayoría, produciendo reportes que se ven completos.

Integridad de series de tiempo. Los historian comprimen agresivamente. Los gaps por cortes, fallas y mantenimiento son comunes; algunos quedan marcados, muchos no. Un agente que calcula un promedio móvil sobre una ventana que contiene un gap de varias horas devuelve un valor sin relación con la realidad física. El número es preciso. No es exacto.

Metadata de proceso como conocimiento tácito. La razón por la que un parámetro está en su valor actual está registrada, la mayoría de las veces, solo en la cabeza de un ingeniero de procesos senior. El significado de las alarmas, las respuestas estándar, la historia de los umbrales viven en bitácoras de cambio de turno y en conversaciones de relevo. Los agentes no pueden leer eso. Tratan el valor configurado como verdad, incluso cuando el operador que lo fijó lo llamaría un parche para un problema de sensor que nadie tuvo tiempo de arreglar.

Ausencia de señal human-in-the-loop. Las correcciones implícitas que hace un operador con experiencia (una válvula manual ajustada para compensar un sesgo conocido, una alarma ignorada porque lleva dos años mal calibrada) son invisibles para el sustrato. El agente observa el resultado de portada (el proceso está estable) sin observar el trabajo

[3] Enterprise Resource Planning — el sistema de registro de materiales, equipos, órdenes de trabajo, finanzas y abastecimiento (p. ej. SAP, Oracle EBS).
[4] Manufacturing Execution System — la capa entre el ERP y la OT que programa y registra corridas de producción, lotes y turnos en la planta.
[5] Computerized Maintenance Management System — el sistema de registro de equipos, órdenes de trabajo, repuestos y planes de mantenimiento preventivo.
[6] Laboratory Information Management System — el sistema de registro de trazabilidad de muestras y resultados de ensayos del laboratorio de la faena.

humano que lo mantiene estable. Cuando se le pide hacerse cargo, hereda la portada sin la mantención.

El agregado de estas condiciones no es “datos desordenados”. Es un sustrato cuya superficie parece inteligible pero cuya integridad semántica es localmente frágil, de maneras que los humanos con experiencia aprendieron a esquivar. Los agentes no aprendieron a esquivarlas. No pueden. El conocimiento de cómo esquivar no está en el sustrato.

03

HISTORIA CENTRAL — EL ERROR CONFIADO DE LA CONCENTRADORA

El relato que sigue es compuesto, anonimizado y operacionalmente plausible. El patrón que describe se ha observado en más de una faena.

Una concentradora de cobre de gran escala desplegó un agente de IA a mediados de 2024 para asistir la optimización de dosificación de reactivos en el circuito de flotación. El agente ingería datos del historian de planta: sondas de pH en tres puntos de los bancos rougher y cleaner, conductividad, ley de mineral desde el analizador en línea, y flujos de reactivos desde bombas de desplazamiento positivo. Recomendaba ajustes de setpoint a los operadores en ciclos de quince minutos.

Las recomendaciones del agente eran en general buenas. La recuperación mejoró una fracción medible de punto en las primeras seis semanas. El superintendente de metalurgia respaldó el piloto. El equipo de operaciones se acostumbró a seguir las sugerencias del agente sin cuestionarlas. Las recomendaciones eran pequeñas, frecuentes, y casi siempre en la dirección en la que el operador se habría movido de todos modos.

Entonces, a lo largo de unos cuatro meses, la recuperación se degradó en silencio 1,2 puntos. Las recomendaciones del agente no habían cambiado de carácter; seguían siendo pequeñas, frecuentes y confiadas. El gerente de planta atribuyó inicialmente el declive a un cambio mineralógico conocido en la alimentación. El analizador en línea se envió a verificación. Se consultó al proveedor de reactivos por variabilidad de lotes. Se revisó el depósito de relaves aguas abajo. Ninguna de esas investigaciones identificó una causa raíz.

La causa real la descubrió cuatro semanas después un técnico de mantenimiento en una ronda rutinaria de limpieza. La sonda de pH a la cabeza del banco rougher tenía un certificado de calibración fechado en 2017. El certificado se había arrastrado por cada auditoría anual mediante una convención administrativa que nadie había cuestionado. La sonda había derivado aproximadamente 0,4 unidades de pH en los años intermedios, y el agente, operando bajo el supuesto de que la lectura de pH del historian era una representación fiel del pH físico, había estado sobre-recomendando sistemáticamente adiciones de cal, que suprimían correctamente la flotación de pirita pero también suprimían, como efecto secundario, la recuperación de calcopirita.

El costo estimado de los cuatro meses de degradación estuvo en el rango bajo de un dígito de millones de dólares. El costo no fue catastrófico. El patrón, en cambio, es el artefacto

relevante. **El sustrato le dijo una mentira al agente, y el agente les dijo a los operadores una mentira confiada, pequeña y frecuente a su vez, y los operadores siguieron las mentiras pequeñas y frecuentes porque venían vestidas con el lenguaje de la optimización y no con el lenguaje del error.**

Cuando el agente se pausó para investigación, los operadores volvieron a sus heurísticas de dosificación previas al agente. La recuperación se recuperó. La sonda se recalibró. El agente se reactivó. La recuperación mejoró.

La faena no ha vuelto a desplegar el agente en ninguna capacidad de toma de decisiones desde entonces. Hoy corre en modo observador pasivo mientras un programa paralelo reconstruye la disciplina de calibración en todo el circuito húmedo.

La lección que sacó el superintendente de metalurgia, cuando el caso se revisó en una reunión trimestral de seguridad, no fue "el agente se equivocó". Fue: "el agente hizo exactamente lo que le pedimos con los inputs que le dimos. No le dimos los inputs que creíamos estarle dando."

04

CUATRO MODOS DE FALLA DE LOS AGENTES DE IA SOBRE SUSTRATO DEFICIENTE

La historia de la concentradora es una instancia de un patrón que se repite con regularidad mecánica: el sustrato le dice una mentira al agente, y el agente traduce esa mentira en recomendaciones que parecen optimización. Hay cuatro modos de falla característicos por los que esa traducción sale mal, y se repiten en los despliegues donde la calidad digital subyacente no fue auditada.

Amplificación de alucinaciones. Es el más discutido y posiblemente el menos peligroso, porque produce salidas visiblemente erróneas. El agente lee un input obsoleto o inconsistente y genera una recomendación que un operador con experiencia reconoce de inmediato como fuera de rango. La recomendación se rechaza. No hay más daño que una pérdida de confianza del operador. Este modo de falla es recuperable.

Confianza mal depositada. Es el modo de falla que experimentó la concentradora descrita en la Sección 03. Las recomendaciones del agente son *plausibles* (dentro del rango de acciones que un operador con experiencia podría tomar) pero están sistemáticamente sesgadas por un input del que el agente no tiene forma de desconfiar. El efecto del sesgo es pequeño por ciclo y se acumula a lo largo de semanas o meses. Los operadores no rechazan las recomendaciones porque, individualmente, se ven razonables. El sesgo es invisible hasta que la evidencia acumulada (una curva de recuperación, una tendencia de rendimiento, una deriva de intensidad energética) obliga a investigar. Es el modo de falla más caro porque es el que permanece más tiempo sin detectarse.

Lavado de decisiones. Cuando un agente emite una recomendación y un operador la sigue, el registro institucional de la decisión lista al operador como el actor. El rol del agente queda registrado en un campo de metadata que nadie revisa después. Cuando la decisión resulta

haber sido errónea, la responsabilidad persigue al operador. Cuando el razonamiento del agente se reconstruye más tarde, el analista encuentra que el operador estaba, en la práctica, ejecutando una recomendación de máquina. La cadena de responsabilidad se evapora. Las investigaciones de seguridad, las auditorías ESG[7] y las consultas regulatorias se vuelven difíciles de concluir porque el actor de registro no es la entidad que decide. Es un problema de gobernanza, no un problema técnico, pero lo crea la arquitectura técnica.

Colapso del audit trail. Los agentes de IA que razonan sobre datos y producen recomendaciones rara vez persisten el estado completo sobre el que razonaron. Registran la recomendación, a veces registran los inputs, casi nunca registran los datos de contexto que el modelo usó para interpretar los inputs. Cuando un regulador o una aseguradora pregunta, seis meses después de un incidente, qué "sabía" el agente al momento de una decisión específica, la respuesta es casi siempre: no podemos reconstruirlo. El historian ya rotó. La memoria de trabajo del agente era efímera. El razonamiento intermedio ya no existe. La institución descubre, en el peor momento posible, que ha estado operando sin una historia de decisiones auditable.

Estos cuatro modos de falla no son exóticos. Cada uno de ellos ha ocurrido en despliegues productivos observables a través de nuestra práctica de consultoría en los últimos dieciocho meses. Tampoco son específicos de la minería; los patrones se transfieren a operaciones de cemento, celulosa y agroindustria cambiando solo los nombres de los procesos unitarios.

Figura 1 Los cuatro modos de falla característicos de los agentes de IA desplegados sobre sustrato no auditado, mapeados contra los dos ejes que más determinan su costo organizacional: visibilidad de la falla y consecuencias de la decisión subyacente.

[7] Environmental, Social and Governance — el marco de reportería que reguladores, inversionistas y financistas exigen crecientemente para sustentar declaraciones de sostenibilidad.

05

LAS CAPAS DE FUNDACIÓN DE LAS QUE DEPENDEN LOS AGENTES

Si “mala calidad digital” es el problema y “primero la fundación, después los agentes” es la recomendación, ¿de qué está hecha realmente la fundación? Las capas que siguen no son un modelo de madurez de consultora. Son un checklist operacional de lo que un agente autónomo necesita para producir decisiones que una organización pueda confiar y defender.

Capa 1 — Datos maestros. Cada entidad sobre la que un agente razona (equipo, material, ubicación, etapa de proceso, rol de operador) debe tener un identificador canónico único, resoluble en todos los sistemas que el agente lee. El trabajo de reconciliar identificadores entre ERP, MES, CMMS, historian y LIMS es ingrato y caro. También es innegociable. Un agente operando sobre datos maestros fragmentados producirá salidas individualmente correctas y agregadamente sin sentido.

Capa 2 — Calidad de telemetría. Cada sensor cuya lectura entra al razonamiento de un agente debe tener un registro de calibración vigente, un rango de incertidumbre documentado, un régimen de detección de fallas, y una política clara de cómo el agente maneja condiciones de falla o fuera de rango. “Confiar en el sensor” no es una política. “Desconfiar de lecturas de más de N horas de un sensor cuyo certificado de calibración tiene más de M meses” es una política.

Capa 3 — Integridad de series de tiempo. El historian debe distinguir, para cualquier ventana de tiempo, entre *no hay datos* y *el dato es cero*. Los gaps deben marcarse en la ingesta, no interpolarse invisiblemente. Las consultas del agente contra series de tiempo deben propagar la conciencia del gap hacia la salida del agente. Si el 18 % de la ventana faltaba, la recomendación del agente debe reflejar esa incertidumbre, no absorberla en silencio.

Capa 4 — Metadata de proceso. El conocimiento tácito de por qué los parámetros operacionales están en sus valores actuales debe externalizarse progresivamente. Es un programa multianual en la mayoría de las operaciones y no puede secuenciarse como precondición de cada despliegue de agentes. Pero sí debe secuenciarse como precondición del despliegue de agentes *de alto riesgo*, dondequiera que una recomendación pueda mover un setpoint que afecte la seguridad, el cumplimiento ambiental, o ingresos a escala material.

Capa 5 — Telemetría human-in-the-loop. El agente debe observar y registrar las correcciones humanas que se aplican a sus recomendaciones. Cuando un operador hace override de un setpoint, el override debe capturarse con códigos de razón y devolverse al loop de entrenamiento o calibración del agente. Sin esto, el agente nunca aprende en qué se está equivocando; acumula un modelo privado de la operación que diverge de la práctica real de la operación.

Estas cinco capas se componen. Una falla en cualquiera de ellas socava la confiabilidad del agente. Una organización que invirtió fuerte en las capas 2 y 3 dejando la capa 1 intacta seguirá produciendo agentes cuyas salidas están sistemáticamente rotas en las fronteras de los cruces. Las capas no son pasos opcionales de un rollout por fases; son condiciones simultáneamente necesarias.

Figura 2 Las cinco capas operacionales de fundación de las que depende un agente autónomo, apiladas desde la más fundacional (datos maestros) hasta la más adyacente al agente (telemetría human-in-the-loop). Las dos capas de la base, datos maestros y calidad de telemetría, son las que más comúnmente están rotas en despliegues productivos.

06

POR QUÉ RECURSOS NATURALES ES ESPECIALMENTE VULNERABLE

Varias industrias despliegan agentes de IA sobre sustratos imperfectos. Las operaciones de recursos naturales están especialmente expuestas por cuatro razones que se componen entre sí.

Densidad de activos físicos. Una concentradora moderna, una planta de cemento, una planta de celulosa o una explotación agrícola de escala industrial tiene miles de sensores desplegados a lo largo de kilómetros de equipos de proceso. El costo de una disciplina de calibración exhaustiva escala con la cantidad de sensores, y el costo operacional de sacar un sensor de línea para verificarlo es más alto que en industrias más livianas. La deuda de calibración se acumula más rápido que en, digamos, banca o telecomunicaciones.

Baja madurez IT/OT. Las integraciones que le permiten a un agente de IA leer a través de ambos dominios son recientes, frágiles, y casi siempre construidas por integradores en vez de ser propiedad de los equipos de ingeniería de IT o de OT. Cuando una integración se rompe, el sustrato del agente se rompe con ella. El tiempo hasta la detección es largo porque

ninguno de los dos lados considera la integración su responsabilidad primaria. Un agente que depende de un pipeline que no es de nadie eventualmente va a equivocarse en algo que nadie puede explicar.

Consecuencias de seguridad y medioambiente. Una recomendación errónea de un agente en un contexto de atención al cliente produce un reembolso. Una recomendación errónea de un agente en un depósito de relaves, en un horno de cemento, en un circuito de flotación o en un calendario de riego puede producir un evento regulatorio, una respuesta sindical, o un incidente ambiental con consecuencias de varios años. Las consecuencias mueven el cálculo desde "lanza el piloto e itera" hacia "verifica el sustrato antes de exponer el agente a los actuadores".

Presión regulatoria y ESG. Los mismos directorios que empujan el despliegue de agentes están también respondiendo a la presión de los inversionistas por reportería ESG verificable. Las dos exigencias entran en tensión cuando las decisiones del agente no pueden reconstruirse. Una operación que no puede demostrar, seis meses después del hecho, por qué se usó un patrón particular de dosificación de reactivos se encontrará sin poder sustentar sus declaraciones de sostenibilidad ante auditores que, cada vez más, leen dentro del sistema en vez de leer el reporte.

Estos cuatro factores no hacen que los agentes de IA sean inadecuados para recursos naturales. Hacen que el argumento de secuenciamiento sea mucho más importante de lo que sería en industrias más livianas. Primero la fundación no es una preferencia. Es una precondición para no crear, en la búsqueda de la excelencia operacional, una clase de pasivos de cumplimiento y seguridad que no existía antes de desplegar los agentes.

07

DÓNDE ESTÁN REALMENTE LAS OPERACIONES LATAM

A través de las operaciones que hemos recorrido en Chile, Perú, Brasil, Argentina, Colombia y México en los últimos tres años, la distribución es incómoda pero consistente. Una minoría pequeña de operaciones invirtió en las capas de fundación antes de la ola de agentes — normalmente programas digitales greenfield en proyectos de capital recientes, u operaciones cuya matriz corrió una racionalización de datos maestros de varios años. Esas pueden desplegar agentes de IA con credibilidad y con los resguardos apropiados. Son la excepción.

La mayoría tiene gaps de fundación que todavía no ha dimensionado, y está desplegando pilotos de agentes en paralelo de todos modos, porque los pitches de los proveedores y la presión del directorio operan en un calendario que no espera a que el trabajo de fundación termine. Esas operaciones están acumulando los modos de falla descritos en la Sección 04 a un ritmo que se va a hacer visible en 2026 y 2027.

> "Nos dijeron durante dos años que teníamos los datos. El primer mes del agente fue la primera vez que alguien intentó usar los datos de la manera en que decíamos que podíamos. Llevamos once meses reconstruyendo las partes que no existían."
>
> — VP de Operaciones, minera diversificada grande, LATAM

Lo específico de LATAM es la brecha entre la narrativa pública (transformación digital acelerada, liderazgo en IA minera) y la realidad operacional (deuda de calibración, fragmentación de datos maestros, fragilidad de integraciones). Las operaciones que se comprometen públicamente con el liderazgo en agentes de IA sin auditar primero su sustrato no son los ganadores del próximo ciclo. Son los casos de estudio del próximo ciclo.

08

CINCO PREGUNTAS ANTES DE DESPLEGAR

El diagnóstico que sigue no es un modelo de madurez. Son cinco preguntas que un CIO o un CDO puede hacer, en una conversación con su contraparte de operaciones y su integrador, que van a hacer visibles los gaps de sustrato más comunes antes de comprometer el piloto de agentes.

PREGUNTA DIAGNÓSTICA *Para los sensores cuyas lecturas van a entrar al loop de decisión del agente: ¿cuál es la edad mediana de sus certificados de calibración vigentes, y cuál es el procedimiento documentado para manejar lecturas de sensores cuya calibración expiró?*

Si la respuesta a cualquiera de las dos partes de esta pregunta es "tendría que averiguarlo", la fundación no está lista y el agente todavía no debería estar tomando decisiones.

PREGUNTA DIAGNÓSTICA *Para los datos maestros que el agente va a cruzar entre sistemas: ¿quién es el dueño del identificador canónico de cada clase de entidad, y cuándo fue la última auditoría de reconciliación entre los sistemas que el agente va a leer?*

Si la respuesta es "estamos corriendo un proyecto de reconciliación", el agente debería correr en modo observador hasta que ese proyecto entregue.

PREGUNTA DIAGNÓSTICA *Si un regulador o una aseguradora nos preguntara, en dieciocho meses más, qué "sabía" el agente al momento de una decisión específica, ¿podemos reconstruir el estado completo de los inputs, la versión del modelo y la traza de razonamiento?*

Si la respuesta es no, el agente no puede desplegarse en ninguna capacidad que afecte seguridad, medioambiente o reportería regulada.

PREGUNTA DIAGNÓSTICA *Cuando un operador hace override de una recomendación del agente, ¿el override se captura con códigos de razón, se persiste, y se devuelve al loop de evaluación del agente?*

Si no, el agente va a divergir de la realidad operacional en silencio y no vamos a saber que divergió hasta que el daño acumulado obligue a investigar.

PREGUNTA DIAGNÓSTICA *Cuando el agente se equivoca, no catastróficamente sino con errores pequeños y frecuentes, ¿el KPI[a] de quién lo atrapa?*

[a] Key Performance Indicator — indicador clave de desempeño; métrica medible para evaluar un proceso, equipo o sistema contra un objetivo definido.

Si la respuesta es "el mismo KPI que el agente está optimizando", el agente se corrigió su propia prueba y la institución no tiene ningún control independiente sobre su exactitud.

Estas cinco preguntas no garantizan un despliegue exitoso. Hacen visibles los gaps de fundación más comunes antes de que el compromiso se firme. El costo de hacerlas es una

conversación. El costo de no hacerlas es el costo de los modos de falla de la Sección 04 multiplicado por la duración de la operación sin detección.

09

EL ARGUMENTO DE SECUENCIAMIENTO — PRIMERO LA FUNDACIÓN, DESPUÉS LOS AGENTES

El pushback más común al argumento de fundación-primero es que el trabajo de fundación toma años y el despliegue de agentes está pasando ahora. Que los dos no pueden secuenciarse; tienen que ir en paralelo.

Ese pushback es parcialmente correcto y mayormente erróneo.

Es parcialmente correcto en que el trabajo de fundación *puede* y *debe* correr en paralelo con despliegues de agentes de alcance acotado: específicamente, agentes en modo observador que producen recomendaciones sin tomar acciones, y agentes cuyo alcance de acción está acotado a dominios reversibles y de bajo riesgo. Hay valor de aprendizaje genuino en correr agentes contra datos imperfectos mientras la fundación se reconstruye.

Es mayormente erróneo en que el argumento de secuenciamiento no se trata de diferir todo el trabajo de agentes hasta que todo el trabajo de fundación termine. Se trata de no desplegar agentes en roles de toma de decisiones para procesos cuyo sustrato no fue auditado. El secuenciamiento es por proceso, no por organización.

Una formulación práctica: **la auditoría de sustrato es precondición de todo despliegue de agentes cuyo alcance de acción incluya un setpoint que afecte seguridad, medioambiente o ingresos materiales.** Para los despliegues de agentes fuera de ese alcance (recomendar calendarios de turnos, redactar órdenes de trabajo de mantenimiento, resumir reportes operacionales), la auditoría de sustrato es recomendada pero no bloqueante.

Esta formulación les da a las operaciones un camino productivo. No exige que un programa de fundación de años termine antes de que cualquier agente salga. Sí exige que el sponsor ejecutivo de cada proyecto de agentes sepa si el alcance de acción propuesto cruza el umbral de seguridad/medioambiente/ingresos, y si lo cruza, que exija la auditoría de sustrato antes de comprometer.

El incentivo del proveedor es empujar el umbral hacia arriba, caracterizar cada piloto como de bajo riesgo y reversible, porque la auditoría de fundación es fricción de la que el proveedor no se beneficia. El trabajo del ejecutivo es empujar el umbral hacia abajo, caracterizar cada piloto como potencialmente consecuente, porque el costo de equivocarse sobre el umbral es asimétrico. Equivocarse hacia abajo produce excelencia operacional. Equivocarse hacia arriba produce pasivos de cumplimiento.

En nuestra experiencia de consultoría, los ejecutivos que parten asumiendo riesgo alto atrapan los problemas temprano y entregan agentes funcionando dentro de doce meses. Los ejecutivos que parten asumiendo riesgo bajo lanzan pilotos de agentes rápido y se pasan los dos años siguientes reconstruyendo la fundación mientras le explican al directorio

por qué los pilotos se pausaron. El primer grupo termina el ciclo con agentes. El segundo lo termina con explicaciones.

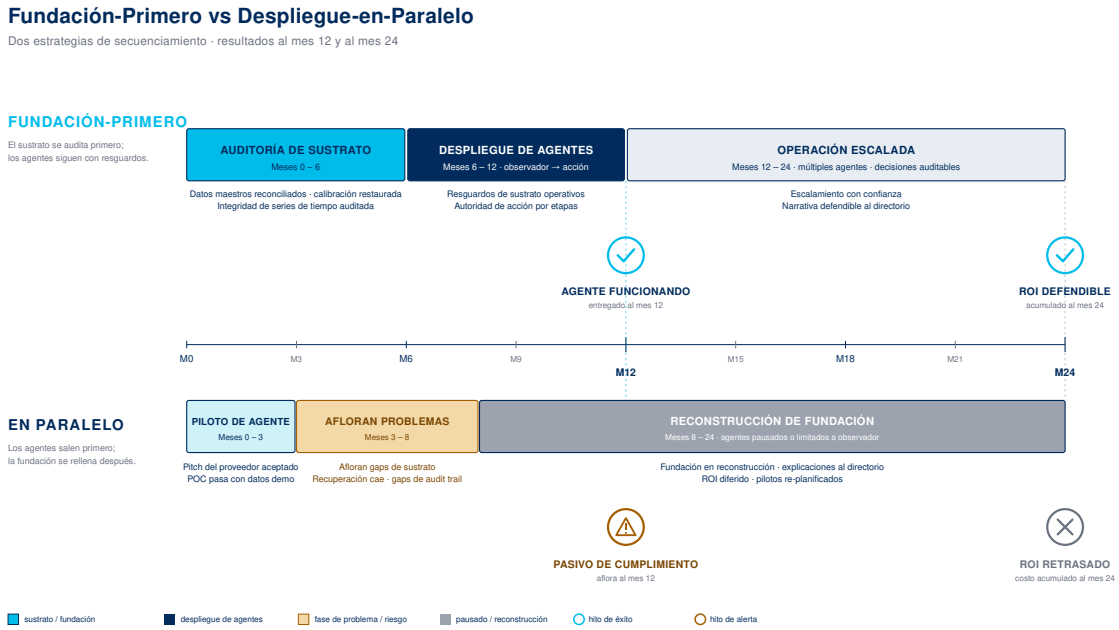

Figura 3 Dos estrategias de secuenciamiento comparadas sobre un horizonte de veinticuatro meses. El camino fundación-primero entrega un agente funcionando y defendible al mes 12 y escala desde ahí. El camino de despliegue-en-paralelo llega al mes 12 con un pasivo de cumplimiento en acumulación y se pasa el año siguiente reconstruyendo la fundación que debió construirse primero.

10

ANTI-PATRONES QUE HEMOS VISTO REPETIDAMENTE

Los anti-patrones que siguen no son teóricos. Cada uno se ha observado en más de una operación en los últimos dieciocho meses.

La trampa de velocidad del POC. Un proveedor ofrece una prueba de concepto de seis semanas. La operación acepta porque el calendario es corto. El POC funciona contra datos de demostración. El “despliegue productivo” se trata como un ejercicio de configuración y no como un proyecto de integración. Los gaps de sustrato afloran en la semana dos de producción, y el agente o se desactiva o corre contra datos no confiables, porque el trabajo de integración necesario para remediar empujaría el calendario más allá de lo que el compromiso original del POC permitía.

La selección de proveedor agente-primero. La operación emite un RFP[8] por una capacidad de agentes de IA. El RFP es técnicamente excelente. El proveedor ganador tiene un modelo fuerte y una ingeniería de datos débil. El contrato se firma. El trabajo de ingeniería de datos que debió ser precondición se convierte en una orden de cambio a mitad del proyecto, a la tarifa diaria del consultor, después de que el proyecto perdió el respaldo ejecutivo.

El atajo de linaje de datos. El agente se despliega contra datos que pasaron por pipelines de ETL[9] o de ingeniería de features cuyo linaje está documentado solo en código. Cuando el agente produce una salida inesperada, el equipo no puede trazar qué input movió qué output a través de qué transformación. La investigación toma semanas por incidente. El equipo de operaciones pierde la confianza más rápido de lo que el equipo de ingeniería puede restaurarla.

El argumento de "la gobernanza la agregamos después". La operación despliega el agente sin estructura de gobernanza: sin dueño responsable de la calidad de las salidas, sin protocolo de override, sin gatillo de desmantelamiento. La gobernanza se trata como un entregable de fase 2. La fase 2 nunca parte porque la fase 1 ya está en producción y el equipo se movió al siguiente piloto. El agente envejece en su lugar, acumulando gradualmente desalineación con la práctica real de la operación, hasta que un incidente fuerza la pregunta de quién ha sido responsable de él.

El cumplimiento como idea tardía. El agente se despliega sin pensar explícitamente en la auditoría ESG, el caso de seguridad, o los requisitos de reportería regulatoria. Cuando llega una auditoría, la operación descubre que las salidas del agente están entrelazadas con decisiones reguladas de maneras que el despliegue original no anticipó. Desenredarlas exige o reconstruir el audit trail a posteriori (caro y parcial) o pausar el agente hasta que el caso quede documentado (operacionalmente disruptivo).

Cada uno de estos anti-patrones tiene una raíz común: tratar al agente como una capacidad discreta en vez de como un participante de un sistema operacional y de gobernanza existente. El arreglo es tratar el despliegue del agente como el proyecto de integración que realmente es, con la auditoría de sustrato, el diseño de gobernanza y el gatillo de desmantelamiento como entregables de primera clase.

11

CÓMO SE VEN REALMENTE LOS DATOS LISTOS PARA AGENTES

Esta sección es el checklist operacional que un CIO puede entregarle a un equipo de operaciones y a un integrador para usar como gate de precondición de cualquier despliegue de agentes cuyo alcance de acción cruce el umbral de seguridad/medioambiente/ingresos descrito en la Sección 09.

[8] Request for Proposal — el documento formal de licitación que un comprador emite para invitar propuestas competitivas contra un alcance definido.

[9] Extract, Transform, Load — el patrón de ingeniería de datos que mueve datos desde sistemas de origen, a través de transformaciones, hacia un almacén de destino.

Sobre datos maestros: cada entidad sobre la que el agente va a razonar tiene un identificador canónico documentado, un dueño documentado responsable de su vigencia, y un procedimiento de reconciliación documentado hacia los otros sistemas que el agente va a leer. La reconciliación fue auditada dentro de los últimos doce meses.

Sobre telemetría: cada sensor cuya lectura entra al razonamiento del agente tiene un certificado de calibración vigente, un rango de incertidumbre documentado, un régimen automatizado de detección de fallas, y una política documentada de cómo el pipeline del agente maneja condiciones de falla o de calibración expirada. La política especifica qué hace el agente, no qué debería hacer. No hay manejo implícito.

Sobre series de tiempo: el historian distingue *no hay datos* de *el dato es cero* en cada ruta de consulta que el agente usa. El marcado de gaps es en la ingesta, no en el consumo. Las consultas del agente propagan la conciencia del gap hacia sus salidas.

Sobre metadata de proceso: la documentación del agente incluye una lista de los parámetros de proceso cuyos valores configurados reflejan problemas de sensores conocidos, sesgos de instrumentación o parches operacionales. El agente tiene prohibido razonar sobre esos parámetros como si fueran mediciones directas de la realidad física subyacente.

Sobre telemetría de overrides: cada override de un operador sobre una recomendación del agente se captura con ID del operador, timestamp, la recomendación, la acción de override, y un código de razón de una lista controlada. Los overrides los revisa semanalmente un dueño con nombre y se devuelven al loop de evaluación del agente mensualmente.

Sobre el audit trail: por cada decisión del agente, el sistema persiste el estado completo de los inputs, la versión del modelo, la traza de razonamiento y la recomendación. La persistencia es por la duración que exija el régimen regulatorio más exigente al que la operación esté sujeta, más un año. No hay borrado de historia de decisiones por ventana rodante.

Sobre gobernanza y evaluación: el agente tiene un dueño con nombre en la cadena de mando de operaciones, no en IT, con autoridad para pausarlo, un protocolo documentado de revisión y escalamiento, y un gatillo de desmantelamiento que se dispara automáticamente bajo condiciones definidas. El desempeño se mide con un KPI que pertenece a una organización distinta del equipo que despliega e independiente del objetivo de optimización del agente. La deriva entre los dos se revisa mensualmente.

Una operación que puede responder “sí, por escrito” a cada punto de arriba está lista para desplegar agentes autónomos en dominios de alto riesgo. Una operación que no puede, no está lista. El gap de auditoría de sustrato no es un problema de madurez. Es un gate de preparación para el despliegue.

> “La forma más barata de hacer más inteligente a un agente de IA es hacer confiable su sustrato de datos. La forma más cara es saltarse ese paso.”
>
> — NTT DATA · Natural Resources Practice · 2026

12

IDEAS CLAVE

Las afirmaciones de este paper — cada una capaz de sostenerse sola — se presentan como un bloque por idea: ayuda de memoria.

1 Los agentes no arreglan los datos malos. Los amplifican — leyendo a velocidad de máquina, actuando con confianza de máquina, y documentando decisiones en formatos que ningún investigador puede reconstruir después sin el sustrato de origen.

2 "Tenemos los datos" es una herencia de la era de los dashboards, cuando los datos malos todavía producían salidas de aspecto útil porque había un humano leyéndolas. El agente hereda la portada sin la mantención humana que la sostenía.

3 El modo de falla más discutido — la alucinación — es el menos peligroso, porque es visiblemente erróneo y se rechaza. El más caro es la confianza mal depositada: recomendaciones lo bastante plausibles para seguirlas, sistemáticamente sesgadas por un input del que el agente no puede desconfiar, sin detección hasta que el daño acumulado obliga a investigar.

4 El lavado de decisiones es un problema de gobernanza creado por la arquitectura técnica: el operador es el actor de registro, el agente es la entidad que decide, y cuando la decisión sale mal la cadena de responsabilidad se evapora — frente a un investigador de seguridad, un auditor ESG o un regulador.

5 Si no puedes reconstruir lo que el agente sabía al momento de una decisión específica — estado completo de inputs, versión del modelo, traza de razonamiento — estás operando sin una historia de decisiones auditable. La mayoría de las instituciones lo descubre en el peor momento posible.

6 Cinco capas de fundación son el gate del despliegue de alto riesgo: datos maestros, calidad de telemetría, integridad de series de tiempo, metadata de proceso, telemetría human-in-the-loop. Primero-la-fundación-después-los-agentes no es cautela. Es secuenciamiento — la misma inversión que harías de todas formas, ordenada para que el agente se encuentre con datos en los que puede confiar.

7 El test de preparación es binario y documental: una operación que puede responder las cinco preguntas diagnósticas "sí, por escrito" está lista para agentes de alto riesgo. Una que no puede, no está lista. La auditoría de sustrato es un gate de despliegue, no una aspiración de madurez.

13

LECTURAS RELACIONADAS Y PRÓXIMOS PASOS

Contenido Relacionado de Esta Serie

Esta es la primera pieza publicada de la serie Foundations Before Agents. Las piezas siguientes examinarán la metodología de auditoría de sustrato, los patrones de integración IT/OT para la preparación de agentes, y la arquitectura de gobernanza necesaria para que los agentes sean auditables.

Todo el contenido disponible en biztalksnttdata.com

Próximos Pasos

1. **Corre el diagnóstico de cinco preguntas de la Sección 08** con tu contraparte de operaciones y tu integrador actual o prospectivo. La conversación es corta. Las respuestas son reveladoras.
2. **Para despliegues de agentes ya en curso: identifica el umbral de alcance de acción de la Sección 09.** Si las acciones de tu agente cruzan el umbral de seguridad/medioambiente/ingresos y la auditoría de sustrato no se completó, pausa la autoridad de acción mientras corre la auditoría. El modo observador no es un paso atrás; es un paso hacia un despliegue que puedes defender.
3. **Para despliegues de agentes aún no comprometidos:** exige la auditoría de sustrato como precondición en el procurement. Trata la ingeniería de fundación como parte del alcance del proyecto, no como un programa aparte. Los proveedores que responden bien a ese framing son los que tienen despliegues que funcionan.
4. **Para las conversaciones a nivel de directorio:** la métrica que importa no es la cantidad de agentes desplegados. Es el porcentaje de despliegues de agentes cuyas decisiones pueden auditarse y defenderse seis meses después del hecho. Esa es la métrica que va a distinguir el liderazgo operacional del pasivo acumulado en los próximos dos ciclos.

NTT DATA · Natural Resources Practice · 2026 · biztalksnttdata.com