Management Solutions
Making things happen

GenMS Swarm

Inteligencia agéntica



© ManagementSolutions 2026. Todos los derechos reservados

Índice

01. ¿Qué es GenMS Swarm?

02. Casos de uso

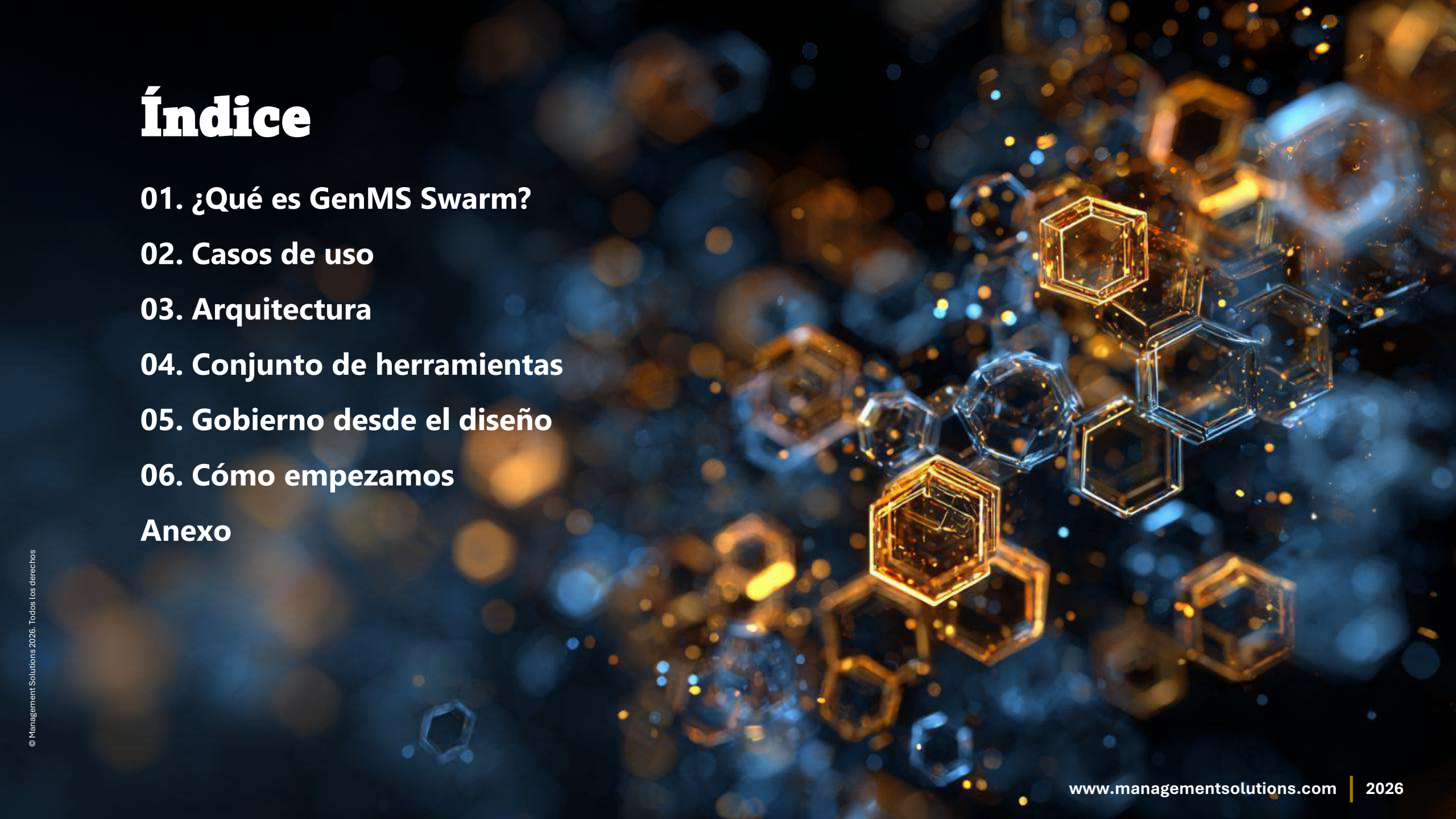
03. Arquitectura

04. Conjunto de herramientas

05. Gobierno desde el diseño

06. Cómo empezamos

Anexo



© Management Solutions 2026. Todos los derechos

01 | ¿Qué es GenMS Swarm?

Una capa de investigación agéntica, desplegada en su entorno de datos y código, que coordina agentes de IA especializados para descubrir y explicar problemas que escapan a los controles predefinidos.

01 CONECTA CON SUS DATOS
Desplegado en su perímetro

- Bases de datos
- Almacenes de datos
- Lagos de datos
- Data marts
- Repositorios de código

Se mantienen los datos y controles de calidad existentes

02 INVESTIGA DE FORMA AUTÓNOMA
Agentes especializados trabajan en paralelo



03 APOYA LAS DECISIONES HUMANAS
Evidencias, no respuestas opacas

- Hallazgo trazable
- Evidencias de respaldo
- Grado de confianza
- Corrección propuesta
- Registro de auditoría

Control humano

¿POR QUÉ AGÉNTICO?	Controles tradicionales	GenMS Swarm	Controles tradicionales	GenMS Swarm	Controles tradicionales	GenMS Swarm
	Errores conocidos codificados en reglas	**Problemas desconocidos explorados como hipótesis**	Comprobaciones predefinidas	**Análisis ad hoc entre tablas, temporal y estadístico**	Estáticos hasta su revisión	**Mejora con hallazgos validados**

02 | Casos de uso

Swarm investiga donde terminan los controles predefinidos:
más allá de la calidad del dato, en distintos sectores, entornos de datos y bases de código.

01 Detección de problemas de calidad del dato

EL PROBLEMA

Las reglas de calidad solo detectan errores conocidos. Los problemas más costosos son aquellos para los que nadie ha escrito una regla.

QUÉ INVESTIGA SWARM

- Anomalías entre campos, tablas y periodos
- Incoherencias entre tablas y fallos en la lógica de negocio
- Causa raíz y registros afectados

RESULTADO

Incidencias de calidad trazables, con evidencias, confianza y correcciones propuestas

02 Conciliación entre sistemas

EL PROBLEMA

Las diferencias entre sistemas de origen, almacén e informes se detectan tarde y se explican manualmente.

QUÉ INVESTIGA SWARM

- Diferencias entre sistemas, etapas y fechas
- Errores de mapeo y transformación
- Desfases temporales frente a discrepancias reales

RESULTADO Ver detalle en el anexo ›

Diferencias explicadas y clasificadas por causa raíz, con soluciones propuestas

03 Revisión de seguridad del código

EL PROBLEMA

Las revisiones muestrean el código con criterios desiguales. La IA avanzada puede convertir vulnerabilidades latentes en exposiciones reales.

QUÉ INVESTIGA SWARM

- Todos los repositorios incluidos, en paralelo y con un criterio común
- Vulnerabilidades, defectos y desviaciones de los estándares de programación
- Causa raíz de incidencias abiertas

RESULTADO

Hallazgos priorizados con pruebas, gravedad y plan de corrección

04 Aseguramiento de datos para modelos

EL PROBLEMA

Los modelos se validan por sus resultados. Los problemas de los datos subyacentes aparecen tarde, a menudo en auditoría.

QUÉ INVESTIGA SWARM

- Coherencia de los datos de desarrollo y entrada
- Incumplimientos de reglas de negocio en factores clave de riesgo, p. ej., datos de pérdida dado el impago
- Evidencias para validadores y auditores

RESULTADO Ver detalle en el anexo ›

Problemas de datos documentados para su validación y revisión de auditoría

03 | Arquitectura

Un ecosistema multiagente jerárquico que crea Detectives especializados,
que a su vez son sistemas multiagente y perfeccionan continuamente sus capacidades.

Entradas

Una base de datos
(p. ej., 200 tablas, 5.000 campos)

Su diccionario de datos

Su modelo de datos

Opcional: histórico de calidad y reglas vigentes

Ejecución

Orquestador

Prospector

Hipótesis n.º 1:
«Las cuentas marcadas como cerradas siguen registrando transacciones después de su fecha de cierre».

Constructor

Detective n.º 1

Subagentes y herramientas
- Planificar
- Ejecutar SQL, Python…
- Evaluar y replantear

…

Detective n.º n

Subagentes y herramientas
- Planificar
- Ejecutar SQL, Python…
- Evaluar y replantear

Base de conocimiento («la Colmena»)

Incidencia DQ 1
Incidencia DQ 2
Incidencia DQ 3
Incidencia DQ n

Para cada incidencia de calidad:
- *Descripción*
- *Evidencias*
- *Solución propuesta*
- *Grado de certeza*
- *Registro de auditoría*
- *Comentarios humanos*

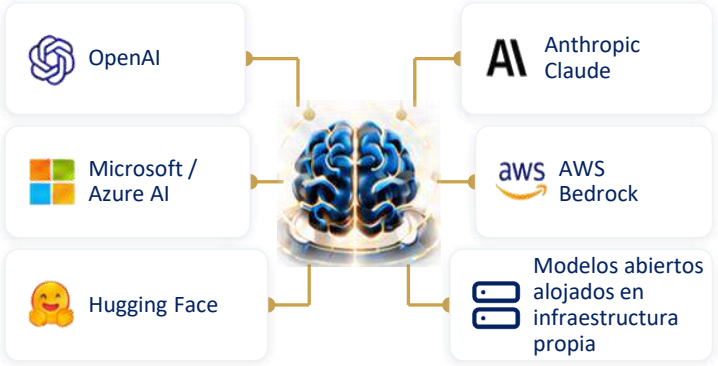
04 | Conjunto de herramientas

Cualquier LLM. Analítica y RAG nativos. Cualquier herramienta compatible con MCP.

Cada Detective utiliza el modelo y la capacidad adecuados en cada paso de la investigación.

01 CAPA COGNITIVA INDEPENDIENTE DEL MODELO

Planifica, razona y elige el modelo adecuado para cada tarea



Cualquier modelo, mediante una pasarela independiente del proveedor

CICLO DE RAZONAMIENTO DEL DETECTIVE

Comprender hipótesis › Dividir en tareas verificables › Ejecutar › Evaluar evidencias › Replantear

Nube · API · infraestructura propia

02 HERRAMIENTAS NATIVAS DE INVESTIGACIÓN

Genera, ejecuta y recupera información bajo demanda



- **SQL:** consultar, cruzar y validar modelos de datos complejos
- **Python:** estadística, aprendizaje automático y explicabilidad
- **Código a medida:** crear funciones cuando ninguna herramienta existente sirve
- **RAG:** recuperar contexto de documentos estructurados y no estructurados
  Word · PDF · Excel · TXT · JSON

Recuperación y ejecución bajo demanda

03 ECOSISTEMA AMPLIABLE DE HERRAMIENTAS MCP

Conecta capacidades externas sin reconstruir Swarm



MCP permite descubrir capacidades especializadas y reutilizarlas en todos los Detectives

Descubrible · reutilizable · gobernado

05 | Gobierno desde el diseño

Autonomía acotada desde el diseño: límites configurables, supervisión humana y evidencias, con controles adaptados a cada despliegue, caso de uso, herramienta y acción.

01 ANTES — Definir los límites



- Delimitar las fuentes y herramientas aprobadas
- Configurar accesos, presupuestos y condiciones de parada
- Definir puntos de aprobación por caso de uso

02 DURANTE — Controlar la investigación



- Aplicar límites de consulta y ejecución
- Derivar a revisión los pasos con baja confianza
- Pausar o detener una investigación

03 DESPUÉS — Documentar cada resultado



- Conservar hipótesis, consultas, código y resultados
- Validar los hallazgos antes del uso operativo
- Incorporar resultados aprobados al conocimiento compartido

Autonomía configurable

Elegir el modo de operación adecuado

SUGERIR
Solo hallazgos y acciones propuestas

REVISAR
Aprobación humana antes de ejecutar

EJECUTAR
Acciones preaprobadas dentro de los límites

Se configura por caso de uso, herramienta y acción

PRINCIPIOS DE CONTROL | Mínimo privilegio | Supervisión humana | Trazabilidad completa | Ejecución segura ante fallos

06 | Cómo empezamos

Un itinerario desde el inicio hasta un swarm operativo con sus datos y gestionado por su equipo.

La fase 0 adapta el plan a su entorno, datos y prioridades.

0 — FASE 0 · 2 semanas

Análisis y diseño

Acordamos sus datos, prioridades y alcance del piloto antes de empezar el desarrollo.

1 — FASE 1 · Implantación · 12-14 semanas

01 — Preparación técnica

Despliegue en su perímetro y conexión con la base de datos piloto y su LLM.

02 — Adaptación

Carga del diccionario, modelo de datos e histórico de calidad; ajuste de hipótesis a su ámbito.

03 — Piloto funcional

El swarm opera con datos reales; los responsables de datos revisan los hallazgos.

04 — Transferencia y cierre

Entregamos a su equipo el conocimiento, el código y la documentación.

2 — FASE 2 · Continuidad

Escalado según su hoja de ruta

Más bases de datos, ámbitos y misiones para el mismo motor.

CAPACIDAD, NO LICENCIA

Su perímetro

Todo funciona dentro de su entorno, incluido el LLM.

Su LLM

Su modelo corporativo, proporcionado y gestionado por ustedes.

Su código

Ustedes son propietarios de los entregables generados. Nosotros conservamos únicamente nuestro conocimiento previo.

Una base de datos para empezar

Primero, un piloto acotado; después, escalado por ámbito.

Anexo | Detalle del caso: conciliación entre sistemas (1/2)

Contratos huérfanos detectados, cuantificados y vinculados a dos causas raíz distintas, al comprobar si los contratos de Riesgos y FINREP llegan a la tabla de conciliación.

542k	contratos huérfanos o descuadrados
13.5B	exposición de riesgo en contratos huérfanos
7	acciones SQL dirigidas y totalmente trazables
$0.10	coste del modelo para toda la investigación

SECUENCIA DE INVESTIGACIÓN

A1–A2 Detección de huérfanos › A3 Cuantificación de diferencias › A4 Descuadres de asignación › A5 Análisis de patrones › A6 Validación de muestra › **Criticidad alta · certeza 3/3**

F1 Contratos ausentes de la tabla de conciliación

HALLAZGO

Tanto la tabla de Riesgos como FINREP contienen contratos que no llegaron a la tabla de conciliación: una diferencia pequeña en Riesgos y otra grande en FINREP.

- 33.797 en Riesgos
- 532.304 en FINREP

SIGUIENTE PASO PROPUESTO

Cuantificar el impacto de la exposición sin conciliar en la información regulatoria.

F2 Dos fallos distintos explican los contratos huérfanos

HALLAZGO

Los huérfanos de Riesgos carecen de metadatos de reporting: un fallo de calidad en origen. El 80,6 % de los huérfanos de FINREP se concentra en un segmento (vivienda, Stage 1): un fallo específico de ETL.

- Metadatos: 100 % NULL
- 80,6 % en un segmento

SIGUIENTE PASO PROPUESTO

Auditar primero la lógica ETL de ese segmento y restablecer la clasificación en origen.

F3 Discrepancias en registros conciliados

HALLAZGO

También hay diferencias en registros presentes en ambos sistemas. Los contratos conciliados muestran descuadres en la asignación de saldos de provisiones y riesgos.

- ≥10.000 descuadres
- Saldos de provisiones y riesgos

SIGUIENTE PASO PROPUESTO

Revisar por qué los controles de conciliación no detectaron estos descuadres.

‹ Volver a casos de uso (caso 02)

Anexo | Detalle del caso: conciliación entre sistemas (2/2)

Investigation Report

Final report generated from detective outputs.

Execution **exec_ccc720b1f45d** · Status **done** · Total cost **$0.1040** · Tokens **75,487** (70,729 in / 4,758 out)

Orphan Contracts in FinRep vs [redacted] Reconciliation Table

Report: hyp-8f06cb7bbdce · Criticality: High · Certainty: 3/3

Reconciliation breaks between source systems indicate failed data quality controls and risk of regulatory reporting errors. Missing contracts in cruce table may hide risk exposure or financial statement misstatements.

Executive Summary

Hypothesis supported

The reconciliation process between [redacted] and FinRep systems has failed significa
33,797 [redacted] contracts and 532,304 FinRep contracts are missing from the recor
financial balance allocations. [redacted] orphans lack basic staging metadata entirely, v
early processing stages, representing 13.5 billion in unreconciled financial risk. These
quality controls that should prevent such mismatches. Urgent action is required to inv
processing, and to validate whether current financial statements are complete and acc

Final Detective Plan

Identify Orphan Contracts in [redacted] Not Present in Cruce Reconciliation

INSTRUCTIONS

Query t_[redacted]_contract to find all contract_id values that do NOT exist in t_cr
provision_finrep_bal, and risk_finrep_bal for orphan records. Use NOT EXISTS anti
contracts that failed to reconcile.

EXPECTED OUTCOME

A list of contract_id records absent from the cruce table. Zero rows indicates com

Conclusion Evidence

HIGH · Supported

The hypothesis of orphan contracts in FinRep vs. [redacted] reconciliation is CLEARLY SUPPORTED by the evidence. The investigation detected 10,000+ orphan [redacted] contracts and 10,000+ orphan FinRep contracts absent from the reconciliation table (t_cruce.[redacted]_finrep), with the reconciliation table containing 33,797 fewer [redacted] contracts and 532,304 fewer FinRep contracts than expected. All orphan [redacted] contracts exhibit NULL values for financial reporting fields (sector_finrep, max_stage_finrep, provision_finrep_bal, risk_finrep_bal), while FinRep orphans concentrate 80.6% in the HOG sector at S1 stage, representing 13.5 billion in orphaned risk exposure. Additionally, 10,000+ reconciled contracts exhibit allocation mismatches between [redacted] and Cruce, demonstrating that reconciliation failures span both missing and discrepant records. This represents a systemic breakdown in data quality controls with high regulatory reporting risk.

The investigation executed seven targeted actions: orphan detection queries (A1-A2) confirmed non-zero orphan populations in both source systems; aggregate count validation (A3) quantified reconciliation gaps; allocation mismatch detection (A4) identified discrepancies in financial balances across stage-specific and overall provision/risk fields; pattern analysis (A6) revealed highly uneven distribution indicating targeted ETL failures rather than random gaps; and sample validation (A7) confirmed orphans lack FinRep metadata. The concentrated distribution (particularly S1/HOG) suggests specific transformation pipeline failure points, while the uniform NULL values across [redacted] orphans indicate upstream data quality issues. A data gap exists due to unavailable raw source tables (t_[redacted]_raw, t_finrep_raw), which would provide definitive root cause attribution, but the aggregate table analysis is sufficient to confirm reconciliation failure.

Immediate corrective actions are required: (1) Investigate [redacted] data quality procedures to restore stage/sector classification for 33,797 orphaned contracts; (2) Audit ETL logic for S1/HOG segment processing to identify why 432,333 FinRep contracts fail reconciliation in this segment; (3) Validate reconciliation control procedures to detect why allocation mismatches persist in 10,000+ matched records; (4) Quantify regulatory reporting impact of orphaned risk exposure (13.5B+) and determine if financial statements require restatement; (5) Execute remediation prioritizing HOG/S1 segment given concentration of orphans and risk exposure.

‹ Volver a casos de uso

Anexo | Detalle del caso: aseguramiento de datos para modelos (1/2)

Un conjunto de datos de LGD, un día: tres hipótesis comprobadas y tres problemas detectados, con todos los hallazgos remitidos al responsable del modelo para su validación.

<1 Dia	desde la entrega de datos hasta los hallazgos documentados
3	hipótesis investigadas de forma autónoma
2	hipótesis respaldadas con evidencias cuantificadas
+1	problema detectado al descartar una hipótesis

H1 HIPÓTESIS NO CONFIRMADA

Impagos separados más de nueve meses agrupados incorrectamente

HALLAZGO

Ningún identificador de impago abarca periodos distantes: la lógica de separación funciona. La investigación reveló el problema contrario: impagos separados por pocos días con identificadores distintos, que podrían inflar el recuento.

- ~1.200 casos divididos en exceso
- Detectado el riesgo opuesto

POR QUÉ IMPORTA

Incluso una hipótesis descartada produjo un hallazgo.

H2 HIPÓTESIS CONFIRMADA

LGD realizada sin evidencia de resolución del impago

HALLAZGO

Se calcularon pérdidas finales para impagos sin cierre, subsanación, reestructuración ni fallido. Los casos son recientes y se concentran en una entidad y un producto: indican un problema sistemático.

- 44 operaciones
- 3,9 M€ de exposición
- Tasa media de pérdida >100 %

POR QUÉ IMPORTA

Se identifican entradas no válidas antes de que lleguen al reporting regulatorio.

H3 HIPÓTESIS CONFIRMADA

Margen de conservadurismo final inferior a su referencia por componentes

HALLAZGO

En el 99,3 % de los registros, el margen de conservadurismo final queda por debajo de la referencia por componentes, una media de 4,2 pp. El patrón apunta a la metodología de calibración.

- 99,3 % de los registros
- >22.700 exposiciones
- Brecha de −4,2 pp

POR QUÉ IMPORTA

Posible infravaloración de LGD (art. 179 del CRR) elevada a validación de modelos.

‹ Volver a casos de uso

Anexo | Detalle del caso: aseguramiento de datos para modelos (2/2)

Executive Summary

HYPOTHESIS NOT SUPPORTED

Our investigation found no evidence that defaults occurring more than nine months apart were improperly combined into single LGD observations. Every default event in the dataset has its own unique identifier, and no identifier spans distant time periods. The default separation logic is functioning correctly. However, the analysis identified a related concern in the opposite direction: approximately 1,182 cases where defaults occurring within days of each other were assigned separate identifiers, potentially inflating the count of default events. This over-splitting warrants further review.

Executive Summary

HYPOTHESIS SUPPORTED

We found 44 defaulted facility records (totalling €3.9 million in exposure) where a final loss rate (realised LGD) has been calculated despite the default having no recorded resolution—no closure date, no cure, no restructuring, no write-off, or any other event that would justify finalising the loss estimate. Most affected records are recent (2024–2025) and belong to a single entity and product type, suggesting a systematic process issue rather than isolated errors. The average recorded loss rate exceeds 100%, further indicating premature or erroneous calculation. This requires immediate remediation t

Executive Summary

HYPOTHESIS SUPPORTED

Our analysis confirms a critical and systematic data quality issue: in virtually all records (99.3%) within the Spanish Residential Mortgage LGD portfolio, the final margin of conservatism estimate (LGD_MoC) is lower than its component-level conservatism benchmark (LGD_MoC_C) by an average of 4.2 percentage points. This affects over 22,700 exposures with significant EAD, indicating a calibration methodology error rather than isolated data issues. The finding poses a high risk of regulatory non-compliance under CRR Art. 179 due to potential under-estimation of LGD, and requires immediate model validation review and recalibration.

‹ Volver a casos de uso

Anexo | Biblioteca ModelCraft™

Una biblioteca para todo el ciclo de vida del modelo, desde los datos originales hasta la supervisión en producción: nativa en PySpark, determinista y validada por expertos con criterios regulatorios.

Datos

- Ingesta unificada desde cualquier origen
- Controles de calidad: nulos, duplicados y valores atípicos
- Ingeniería de variables y análisis exploratorio

Modelización

- Supervisado: RF, GBM, XGBoost, SVM
- Agrupación: K-means, GMM
- Selección de variables, ajuste y validación cruzada

Validación

- Pruebas estadísticas: normalidad, varianza y correlación
- Indicadores regulatorios: WoE, IV, PSI, HHI
- Métricas binarias y multiclase, estabilidad

Despliegue

- Empaquetado en procesos reutilizables
- Exportación ONNX y pickle
- Inferencia distribuida (Spark) y local (pandas)

Seguimiento

- Métricas de evaluación con datos nuevos
- Indicadores de deriva (PSI) y estabilidad
- Equidad entre grupos en producción

GARANTÍAS INTEGRADAS

Distribución desde el diseño
El mismo código de prueba a miles de millones de filas.

Determinismo
Entrenamiento e inferencia idénticos entre ejecuciones y entornos.

Validación experta
Revisión según criterios regulatorios de MS; sin fuga de datos.

Sin dependencia del proveedor
Solo Python y PySpark, en infraestructura propia o cualquier nube.

QUÉ LO DIFERENCIA

Explicabilidad
SHAP y LIME, local y global, reproducibles entre ejecuciones.
Justificar cada decisión del modelo ante negocio y supervisores.

Equidad entre grupos
Más de 25 métricas en las cuatro dimensiones habituales del sesgo.
Alineado con el Reglamento de IA y NIST AI RMF.

Un entregable
Un solo archivo ONNX o pickle, desde validación hasta producción.
Equivalencia garantizada por diseño.

TÉCNICAS
GLM y modelos de puntuación · Random Forest, XGBoost, Gradient Boosting · Agrupación: K-means, GMM

APLICACIONES
Modelos regulatorios (IRB, IFRS 9) · Modelos de gestión: riesgo, comercial, precios, marketing y recobros

MODALIDADES DE COLABORACIÓN
Acceso a la biblioteca · Acelerador de proyectos · Codesarrollo de una biblioteca corporativa

‹ Volver al conjunto de herramientas

Anexo | Cobertura de riesgos y controles

Los principales modos de fallo se asocian a controles técnicos y humanos explícitos, configurables por despliegue, caso de uso, herramienta y acción.

01

Modificación insegura de datos

Cambios de datos no válidos o destructivos

ENFOQUE DE CONTROL

Plantillas validadas · condiciones WHERE acotadas · límites de filas · validación posterior · copias · reversión

02

Explosión de consultas

Aumento del coste o sobrecarga de la base de datos

ENFOQUE DE CONTROL

Presupuestos · límites de frecuencia · réplicas de lectura · condiciones de parada

03

Conflictos sin resolver

Correcciones incoherentes o contrapuestas

ENFOQUE DE CONTROL

Conciliación · comparación de evidencias · escalado a una persona

04

Sesgo de confirmación

Falsos positivos sistemáticos

ENFOQUE DE CONTROL

Pruebas de refutación · validación independiente

05

Conocimiento desactualizado

Recomendaciones obsoletas

ENFOQUE DE CONTROL

Ciclos de aprendizaje · nuevo análisis del esquema · seguimiento continuo

06

Exposición de datos

Datos enviados fuera del ámbito aprobado

ENFOQUE DE CONTROL

Solo LLM aprobados por el cliente · listas de campos permitidos · enmascaramiento de datos personales · registro de entradas y salidas

07

Inyección de instrucciones

Instrucciones ocultas en datos o documentos

ENFOQUE DE CONTROL

El contenido se trata como datos, nunca como instrucciones · listas de herramientas permitidas · límites de acciones

08

Hallazgos sin respaldo

Conclusiones plausibles sin evidencias

ENFOQUE DE CONTROL

Cada afirmación y cifra se vincula a consultas ejecutadas · comprobación de evidencias · puntuación de confianza

09

Cambio de modelo

Cambios de comportamiento tras sustituir el LLM

ENFOQUE DE CONTROL

Pruebas de regresión antes de cada cambio · versiones fijadas · nueva validación de resultados

Management Solutions
Making things happen

Presencia internacional
Una sola firma

Equipo multidisciplinar

Conocimiento de
las mejores prácticas

Experiencia demostrada

Máximo compromiso

Maitane Astiz
Partner
maitane.astiz@managementsolutions.com

Federico Sánchez
Partner
Federico.Sanchez.Buenadicha@msspain.com

Ana Isabel Quesada
Partner
ana.quesada@msspain.com

Estefanía Espinar
Partner
estefania.espinar@msspain.com

Javier Mendoza
Manager
javier.mendoza.puelles@msspain.com

© Management Solutions 2026. Todos los derechos reservados