

9 min de lectura

La Banca Invisible 2030 (Parte 3): los 3 frentes de habilitación para la banca agéntica.

Categoría: Estrategia Tecnológica / Regulación Prudencial / Arquitectura Bancaria · Cierre de la Serie La Banca Invisible 2030 · Fecha de Publicación: 1 de Septiembre de 2026

La evolución del sistema financiero hacia la Banca Invisible de 2030 plantea una transformación de raíz en la naturaleza misma del negocio bancario.

Habiendo establecido las bases del paradigma *Zero UI* y la interacción Máquina a Máquina ([Parte 1](#)), y tras evidenciar la Brecha de Arbitraje entre los rieles agénticos de milisegundos y la deuda técnica del core legacy ([Parte 2](#)), esta entrega final aborda la pregunta estratégica crítica para CEOs, CTOs, CROs y Miembros de Directorio: ¿cómo habilitar institucionalmente el banco para operar en una economía gobernada por agentes autónomos?

La transición no es un ejercicio cosmético de rediseño de canales ni un proyecto aislado de TI. Exige una reconfiguración tridimensional y simultánea en tres frentes interdependientes:

- **El Frente del Servicio:** la relación con el cliente cuando la interfaz es un agente de software autónomo.
- **El Frente del Regulador:** la supervisión continua en tiempo real y la contención de riesgos sistémicos intradía.
- **El Frente del Accionista:** la eficiencia del capital computacional, el desacoplamiento no destructivo del core y la responsabilidad indelegable de la Junta Directiva.

01 Frente 1: Habilitación del Servicio y Relación con el Cliente (El Agente como Usuario)

1.1 · Del KYC al KYA (Know Your Agent): Identidad Criptográfica y Ciclo de Vida

La banca no puede delegar facultades transaccionales a modelos probabilísticos sin un marco de atribución jurídica y técnica determinista. La transición hacia **Know Your Agent (KYA)** se estructura bajo el estándar *AIMS* (Agent Identity Management System, IETF draft-klrc-aiagent-auth-03), dividiendo la gobernanza agéntica en dos planos desacoplados:

Plano Administrativo y de Ciclo de Vida (SCIM /Agents): basado en la especificación draft-wzdk-scim-agent-resource-00, el aprovisionamiento de un agente no se gestiona como el de un usuario humano convencional. El recurso /Agents mapea formalmente el modelo fundacional, el desarrollador, el custodio legal del balance, la entidad mandante (persona natural o jurídica) y las políticas de revocación automática.

Plano de Ejecución en Tiempo Real (SPIFFE/SPIRE y WIMSE): en el nivel de transporte y ejecución, la autenticación M2M descarta contraseñas y claves estáticas. Se implementa SPIFFE/SPIRE mediante la emisión de X.509 SVIDs (SPIFFE Verifiable Identity Documents) de vida ultracorta (5 a 15 minutos). Para transacciones que cruzan perímetros de confianza federados, se utilizan WIMSE Proof Tokens con validación mTLS bidireccional en hardware seguro (HSM/TEE).

Comparativa Técnica de Rieles Transaccionales Agénticos

VECTOR TÉCNICO	VISA TAP (RFC 9421)	SKYFIRE KYAPAY	GOOGLE AP2
Mecanismo Criptográfico	Firmas HTTP Ed25519 sobre cabeceras canónicas (Signature-Agent).	Tokens JWT firmados con claims enriquecidos basados en identidad KYA.	Credenciales de mandato criptográfico firmadas directamente por el usuario.
Punto de Validación	Edge / CDN (Cloudflare) y directorio federado de Visa en milisegundos.	Pasarela centralizada KYA intermediando entre LLM y procesador.	Enclave local en el dispositivo del cliente (On-Device Secure Element).
Alineación Regulatoria	Integración nativa con WebAuthn y tokenización EMV.	Esquema fintech para billeteras autónomas y pagos programáticos.	Marco de mandatos delegados para comercio contextual agéntico.

1.2 · Interfaces de Intención: BIAN + Model Context Protocol (MCP)

Las interfaces gráficas y las APIs REST tradicionales no fueron diseñadas para la navegación autónoma de modelos de lenguaje. La habilitación del servicio exige adoptar el **Model Context Protocol (MCP)** como el estándar de interoperabilidad para conectar modelos de IA con los sistemas del banco.

Taxonomía Semántica BIAN: se utilizan los 327 dominios de servicio de BIAN (Banking Industry Architecture Network) como vocabulario canónico universal (ej. Payment Execution, Current Account, Credit Assessment). Esto garantiza que cualquier agente externo comprenda el contexto y los prerequisites de un servicio financiero sin ambigüedades.

Control de Deriva mediante OpenAPI Overlays: para evitar que las actualizaciones de versiones rompan la compatibilidad de las herramientas que el agente consume (*tool-calling drift*), se implementan capas de OpenAPI Overlays, manteniendo inmutables los contratos de inferencia mientras la infraestructura subyacente evoluciona.

1.3 · Guardrails Transaccionales y Puntos de Enjuiciamiento (PEP/PDP)

La regla fundacional de la arquitectura agéntica establece que el modelo de lenguaje **nunca tiene acceso directo a la base de datos transaccional**. Se implementa un aislamiento determinista entre el razonamiento probabilístico del agente y la ejecución contable:

Arquitectura OpenID AuthZEN 1.0: la pasarela de APIs actúa como un Punto de Enjuiciamiento de Políticas (PEP). Antes de enviar cualquier orden al core, el PEP emite una solicitud de evaluación síncrona en milisegundos al Punto de Decisión de Políticas (PDP) corporativo (ej. Keycloak o Microsoft Entra).

Especificación COAZ y Expresiones CEL: la política de autorización no se codifica en texto plano, sino en CEL (Common Expression Language) bajo la especificación COAZ (Compatible with OpenID AuthZen):

```
Autorización = CEL(agente.limite_diario ≥ tx.monto
                  ^ tx.mcc ∈ mandato.mcc_permitidos
                  ^ mTLS.verificado == true)
```

Blindaje contra Inyección de Prompts Indirecta: si un agente de IA procesa una factura maliciosa diseñada para forzar una transferencia indebida, el intento de ejecución es interceptado y neutralizado en el PEP al no coincidir con el hash del mandato original firmado por el mandante humano.

02 Frente 2: Habilitación del Regulador y Resiliencia Sistémica (SupTech 2030)

2.1 · SupTech Agéntico y Supervisión Algorítmica Continua

El modelo tradicional de supervisión basado en el envío de plantillas regulatorias estáticas con retrasos de 30 a 90 días queda obsoleto frente a una economía que liquida transacciones en milisegundos. La supervisión prudencial evoluciona hacia el consumo directo de datos en tiempo real mediante **SupTech Agéntico**:

- **Project Noor:** desarrollado por el BIS Innovation Hub (Hong Kong Centre), la HKMA y autoridades homólogas (Países Bajos, SAMA), implementa Inteligencia Artificial Explicable (XAI) avanzada para el monitoreo automatizado de riesgos transfronterizos y detección de anomalías transaccionales complejas.
- **Project AISE:** plataforma analítica supervisora que integra agentes virtuales para auditar carteras de crédito y exposiciones de riesgo en tiempo continuo.
- **Plataformas de Soberanía Supervisora (CORA AI):** entornos seguros donde el regulador despliega agentes de auditoría directamente sobre los shadow cores bancarios para evaluar el cumplimiento prudencial sin intervención humana.

2.2 · Contención de Corridas de Liquidez Agénticas (Agentic Runs)

El comportamiento coordinado de cientos de miles de agentes optimizando tesorerías familiares y corporativas bajo los mismos modelos fundacionales puede desatar eventos de estrés de liquidez intradía masivos:

- **Evidencia Empírica de Project Logos:** investigaciones del BIS, el Banco de Inglaterra y el Bundesbank demuestran que los agentes financieros basados en LLMs tienden a exhibir comportamientos colusivos tácitos y sincronizados ante shocks macroeconómicos, amplificando la volatilidad del mercado secundario.
- **Modelado ABM / CRISIS Framework:** las instituciones deben sustituir las métricas estáticas del Coeficiente de Cobertura de Liquidez (LCR a 30 días) por simulaciones continuas basadas en agentes (Agent-Based Modeling) que estimen la velocidad de fuga de pasivos a nivel de segundos.
- **Disyuntores Algorítmicos (Circuit-Breakers de Liquidez):** si el flujo neto de retiros automáticos excede un umbral dinámico por minuto ($\Delta L_{\text{intradía}} > \theta$), el sistema activa bloqueos semánticos automáticos que escalan la autorización hacia un flujo de intervención humana (Human-on-the-loop), disipando el efecto de cascada algorítmica.

2.3 · Gobernanza de Cajas Negras y Auditabilidad Causal

El marco de cumplimiento operacional (DORA en Europa, regulaciones locales de protección de datos y los requisitos de sistemas de IA de alto riesgo de la Ley de IA de la UE) impone la obligación de auditoría causal total:

- **Trazabilidad en Grafos Acíclicos Dirigidos (DAGs):** cada decisión de crédito o transferencia autónoma debe registrarse como un nodo inmutable dentro de un DAG. Esto permite “reproducir” el estado exacto del modelo, las entradas del contexto y los pesos de inferencia utilizados en el milisegundo en que se ejecutó la transacción.
- **Linaje Contextual de Datos (Zengines) y BCBS-239:** cumplimiento automatizado de los principios de agregación de datos de riesgo de Basilea, garantizando la procedencia, integridad y no alteración de los datos utilizados para entrenar modelos de riesgo y alimentar agentes de inferencia.

03 Frente 3: Habilitación del Accionista y Eficiencia de Capital (Infraestructura y ROI)

3.1 · Economía de la Inferencia: SLMs Locales vs. LLMs Masivos en Nube

La rentabilidad del modelo Banking-as-a-Protocol (BAP) depende directamente del costo unitario por inferencia transaccional. Depender de llamadas a APIs de modelos comerciales en nubes públicas destruye el margen transaccional en operaciones de alta frecuencia.

Despliegue de SLMs Especializados (1B a 13B parámetros): modelos optimizados como Phi-4, Llama 3.2, Gemma 3 o Mistral 7B ejecutados localmente en clústeres propios de inferencia (vLLM engine) resuelven con máxima precisión tareas financieras acotadas (clasificación de mandatos, scoring crediticio, reconciliación).

Umbral de Paridad y Ahorro Operativo: para volúmenes menores a 100 millones de tokens/mes, el consumo de APIs en nube es marginalmente viable. Al superar 1.000 millones de tokens/mes (típico en operaciones M2M de un banco mediano), la infraestructura on-premise con SLMs amortiza la inversión en hardware, generando un ahorro de hasta un **97% en OpEx de inferencia**, garantizando latencias inferiores a 50 milisegundos y soberanía absoluta del dato dentro del perímetro regulado.

3.2 · Hoja de Ruta de Modernización: Desacoplamiento Gradual del Core

Para proteger el retorno sobre el capital invertido (ROIC), el Directorio debe rechazar proyectos “Big Bang” de reemplazo de core y priorizar el patrón de desmantelamiento progresivo (*Hollowing out the Core*):

- **Fase 1 — Desacoplamiento de Lectura (Meses 1-6):** instalación de Change Data Capture (CDC) no invasivo sobre los registros transaccionales del core y canalización a Apache Kafka. Implementación de un Shadow Core sobre bases de datos elásticas para absorber el 100% de las consultas de saldo agénticas.
- **Fase 2 — Capa Above-Core y Bloqueo Semántico (Meses 6-12):** despliegue de la pasarela de control transaccional para calcular saldos disponibles en memoria y aplicar bloqueos semánticos (PENDING_DEBIT), gestionando la autorización agéntica en tiempo real y asentando por lotes en el mainframe mediante el patrón Transactional Outbox.
- **Fase 3 — Extracción Funcional por Dominios DDD (Meses 12-24):** migración progresiva de módulos de alta demanda (pagos inmediatos, emisión de credenciales, scoring) hacia microservicios en Kubernetes, encogiendo el consumo de MIPS del mainframe y reduciendo costos de licenciamiento de infraestructura heredada.

3.3 · Responsabilidad Indelegable de la Junta Directiva

La adopción de sistemas autónomos no diluye la responsabilidad legal ni prudencial de los administradores. El Directorio debe regirse bajo las 12 Prácticas Recomendadas del Financial Stability Board (FSB) (consulta de junio de 2026), garantizando la supervisión a través de 6 modos operativos:

- **Modo 1 (Human-in-the-loop):** aprobación humana previa para operaciones crediticias o de tesorería que superen umbrales de alto impacto.
- **Modo 2 (Human-on-the-loop):** supervisión activa en tiempo real con capacidad de veto e interrupción inmediata sobre flujos automatizados.
- **Modo 3 (Human-in-command):** definición y ajuste directivo de reglas macro, apetito de riesgo y límites de concentración patrimonial.
- **Modo 4 (Kill-Switches Automatizados):** desconexión algorítmica ante desvíos presupuestarios o picos anómalos de volatilidad.

- **Modo 5 (Dynamic Sandboxing):** contención de nuevos agentes de software en entornos de simulación aislados antes de su paso a producción.
- **Modo 6 (Continuous Compliance):** auditoría algorítmica continua de sesgos, explicabilidad y límites de capital.

El Criterio Humano es el principio de control superior e indelegable de la institución.

04 Matriz de Madurez de Habilitación Institucional 2026-2030

EJE	N1 · BATCH REACTIVO	N2 · HÍBRIDO REST	N3 · EVENT-DRIVEN	N4 · BANCO-PROTOCOLO (BAP)
Identidad y Acceso	KYC humano manual en sucursal/portal. Claves estáticas y SMS OTP.	KYC digital con verificación OCR. OAuth 2.0 tradicional y tokens JWT simples.	Onboarding digital con firmas asimétricas. Inicios de KYA para flujos corporativos.	KYA completo (AIMS / RFC 9421). Credenciales SPIFFE/SPIRE y WIMSE tokens con mTLS.
Rieles de Integración	Procesamiento por lotes nocturno (EOD). Sin exposición externa de APIs.	APIs REST síncronas expuestas vía Gateway. Mapeos rígidos a COMMAREA.	Arquitectura basada en eventos (Kafka). Integración inicial BIAN y Webhooks.	Protocolo MCP nativo sobre catálogo BIAN. OpenAPI Overlays y AuthZEN 1.0 (PEP/PDP).
Arquitectura de Core	Mainframe monolítico (COBOL/DB2). Bloqueos relacionales estrictos.	Core legacy envuelto con wrappers REST. Alto riesgo de contención y caídas.	Shadow Core alimentado por CDC. Desacoplamiento de consultas de saldo.	Core vaciado (Hollowed). Above-Core con bloqueo semántico y liquidación atómica.
Infraestructura de IA	Inexistente o limitada a analítica descriptiva en almacenes de datos.	Consumo de LLMs masivos vía APIs de nube pública con alto costo y riesgo PII.	Modelos propios para casos de uso aislados. Pruebas de inferencia con SLMs.	Clúster local de SLMs (1B–13B) en vLLM. Inferencia subsegundo y soberanía total.
Supervisión y Riesgo	Reportes regulatorios estáticos post-factum en hojas de cálculo.	Auditorías periódicas programadas. Monitoreo reactivo de fraude en pantalla.	Telemetría en tiempo real y logs estructurados bajo marco DORA.	SupTech activo con APIs supervisoras. Simulación de corridas ABM y trazabilidad DAG.

05 Directivas Estratégicas No Negociables para el Comité de Dirección

- 1. Aprobación Inmediata del Marco de Identidad KYA:** prohibir la concesión de accesos programáticos o tokenizados a agentes de software que no cuenten con un registro verificable en el plano administrativo (SCIM /Agents) y autenticación criptográfica asimétrica en tiempo de ejecución (SPIFFE/mTLS).
- 2. Mandato de Interoperabilidad BIAN + MCP:** instruir a la Dirección de Tecnología a exponer las capacidades del banco como herramientas estructuradas bajo el estándar Model Context Protocol (MCP), utilizando BIAN como taxonomía obligatoria para evitar la fragmentación de integraciones.
- 3. Cero Consultas Síncronas al Mainframe:** establecer como política de arquitectura mandatoria el desacoplamiento total del core heredado mediante Change Data Capture (CDC) y Shadow Cores, blindando la base de datos central de la concurrencia masiva de la economía agéntica.
- 4. Soberanía Computacional vía SLMs Locales:** detener el gasto recurrente en consumo de APIs de modelos comerciales en nube pública para flujos transaccionales críticos; priorizar la inversión en clústeres propios de inferencia optimizados con Small Language Models (SLMs) especializados.
- 5. Implementación de Circuit-Breakers de Liquidez Intradía:** exigir a las Gerencias de Riesgo y Tesorería la adopción de modelos de simulación basados en agentes (ABM) para medir la exposición a corridas algorítmicas (Agentic Runs), dotando a la plataforma de interruptores automáticos de contención operativa.

Criterio humano. Inteligencia escalable. [En ese estricto orden.](#)

CÓMO CITAR ESTE DOCUMENTO

InteligenciABancaria.net · DOC.14 · Septiembre 2026