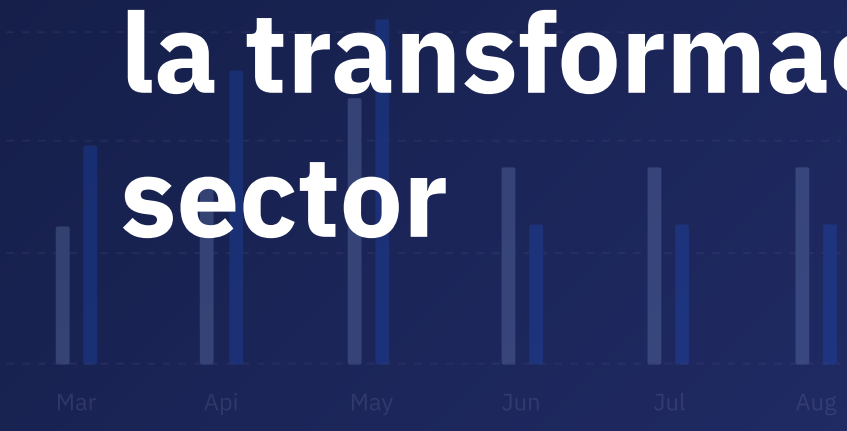
fAI 5.0 nance

GenAI y Agentic AI en la transformación del sector



ARTEFACT

QUIÉNES SOMOS

Autoría & Equipo

El equipo de Autoría está compuesto por profesionales multidisciplinares con experiencia en datos, tecnología, negocios e innovación. Todo el equipo de Artefact trabaja de forma colaborativa para ofrecer soluciones de alto impacto, siempre alineadas con las necesidades de nuestros clientes.

Victor Pontello
Data & AI Consultant Director | LatAm Financial Services Lead

Felipe Longo
Senior Data Consultant

Gustavo Lourenço
Senior Data Consultant

Ilton Chaves
Data Consultant

Pedro Lauand
LatAm Partner

Larissa Ferrari
Senior Data Consultant

Pedro Bertasso
Senior Data Consultant

Vincenzo Spatuzza
Senior Data Consultant

Paolo Gozdzink
Marketing Specialist Brazil & LatAm

Lay d'Arc
Graphic Designer

ACERCA DE NOSOTROS

Artefact

Artefact acelera la adopción de los datos y la inteligencia artificial para generar un impacto positivo en las personas y las organizaciones. Ofrecemos una amplia gama de servicios, desde la estrategia hasta las operaciones, con la implementación de soluciones de IA por sectores industriales, que ayudan a las empresas a aprovechar la ventaja competitiva de la transformación de los datos y la IA.

Índice

INTRODUCCIÓN

El sector financiero es uno de los más dinámicos y transformadores de la economía global.

> El sector financiero se encuentra entre los más dinámicos y transformadores de la economía global. En un escenario en constante evolución, las instituciones bancarias, las corredoras de bolsa y las fintech se enfrentan al desafío de adaptarse rápidamente a las nuevas demandas del mercado, a los avances tecnológicos y a las expectativas de los consumidores, mientras se mantienen competitivas en un entorno cada vez más complejo.

En este ebook, nuestro objetivo es demostrar cómo la **Inteligencia Artificial (IA)** y sus diversas aplicaciones pueden convertirse en aliados estratégicos para las instituciones financieras. Exploraremos desde las aplicaciones más básicas, como la automatización de procesos y el análisis de datos, hasta las tendencias e innovaciones más recientes impulsadas por la **IA Agente**, como la monitorización en tiempo real de transacciones financieras y la detección proactiva de fraudes, redefiniendo la forma en que las empresas interactúan con los datos y toman decisiones.

Mientras que la **GenAI** revolucionó el mundo de la tecnología, posibilitando que los sistemas pudieran crear contenidos e interactuar con los seres humanos, incluyendo el mundo del habla, la escritura y el arte, la IA Agente añade una capa de autonomía y toma de decisiones, permitiendo que estos sistemas actúen de forma proactiva, aprendan de la experiencia y ejecuten tareas complejas con mínima intervención, elevando aún más el impacto de la tecnología y la IA en la vida y en las relaciones cotidianas. En el sector de servicios financieros, esto significa asistentes que no solo generan conocimientos, sino que también los aplican, optimizando las operaciones, detectando fraudes y mejorando la gestión de riesgos en tiempo real, entre muchas otras oportunidades.

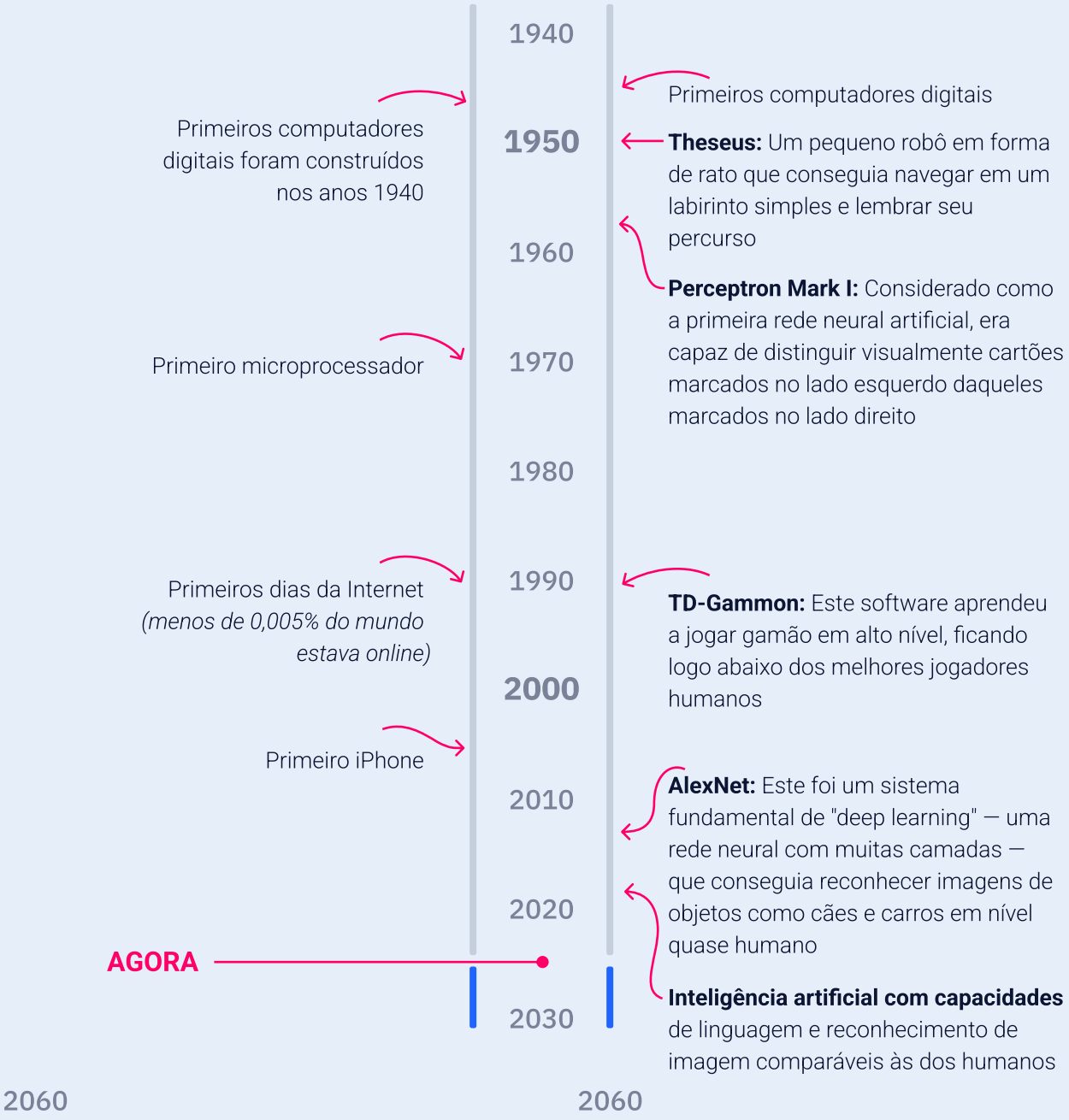
Con años de experiencia trabajando junto a los mayores actores del sector, Artefact ofrece soluciones estratégicas e implementa casos de estudio relevantes que integran de forma integral e innovadora estas tecnologías. Además, esta experiencia previa, junto con el espíritu innovador y emprendedor que impregna a la empresa, nos permite ver una amplia gama de oportunidades para casos de uso que parecían imposibles, pero que, con la evolución tecnológica, no solo son posibles, sino que serán la realidad del sector de servicios financieros en poco tiempo. Estas iniciativas están diseñadas para impulsar el crecimiento sostenible, promover la innovación continua y maximizar resultados consistentes y duraderos.

Esperamos que este material lo ayude a comprender el **valor estratégico de la Inteligencia Artificial** y cómo su empresa puede poner la tecnología al servicio de los negocios e impulsar los resultados a través de las aplicaciones de esta tecnología para navegar con éxito en el **mercado financiero**

¡Buena lectura!

LÍNEA DE TIEMPO

Sistemas destacados de inteligencia artificial [1]



[1] Fuente: The brief history of artificial intelligence: the world has changed fast — what might be next? (https://ourworldindata.org/brief-history-of-ai)

CAPÍTULO 1

El Escenario de la IA en el Sector Financiero

La IA en números: impactos y adopción en el sector financiero

Las instituciones financieras siempre han liderado el uso de datos, impulsando innovaciones tecnológicas en el sector. Desde las computadoras mainframe en la década de 1980 hasta las actuales soluciones avanzadas como la Inteligencia Artificial Generativa (GenAI) y la Inteligencia Artificial Agente (IA Agente), el sector financiero ha sido pionero en la aplicación de estas tecnologías.

EVOLUCIÓN DE LA TECNOLOGÍA - LÍNEA DE TIEMPO

1980's	Lenguaje ensamblador
1998	Lenguaje ensamblador
2007	Lenguaje de instrucciones
2011	Lenguajes orientados a objetos
2022	GenAI
2023	Agentes de IA

Los resultados obtenidos por las instituciones financieras con el uso de GenAI en 2024 son notables y comprueban el potencial de esta tecnología para generar valor en el sector. Casos concretos muestran retornos expresivos: optimización de trading y portafolio con un 25% de ROI, mejora en la experiencia y en el compromiso del cliente alcanzando el 21%, y ganancia de eficiencia en áreas operativas críticas, como el procesamiento de documentos y la generación de informes, llegando al 11%. Estos números dejan claro que la GenAI ya es una realidad transformadora. Las empresas que sepan aprovechar esta oportunidad ahora estarán por delante de la competencia. [2]

¿Cuáles son los principales casos de uso de GenAI con mayor retorno de la inversión (ROI)?

25% Negociación y optimización de carteras

21% Experiencia y compromiso del cliente

11% Procesamiento de documentos

11% Generación de informes

[2] Fuente: State of AI in Financial Services: 2025 Trends - NVIDIA

Además de las ganancias observadas con GenAI, la evolución hacia soluciones basadas en agentes inteligentes está abriendo nuevas y poderosas oportunidades. Estos agentes remodelan el trabajo en tres frentes principales: reducción de costos, al simplificar operaciones y eliminar pasos innecesarios; disminución de lead time, eliminando cuellos de botella humanos y automatizando procesos ; y mejora en la calidad del servicio, a través del uso optimizado de herramientas y ambientes de trabajo. Ejemplos concretos refuerzan este impacto: reducción del 25% en el tiempo de los ciclos de I+D (Investigación y Desarrollo), ahorro de hasta 10 veces en operaciones de call center, aumento de hasta 50 veces en la velocidad de producción de contenido de marketing e impulsos de hasta el 40% en la productividad de equipos de TI. Estos resultados comprueban que la adopción estratégica de agentes es un paso decisivo para generar valor sostenible y medible en el sector financiero.

¿CUÁLES (MOTORES DE VALOR) SE PUEDEN ESPERAR?

Los agentes están transformando la forma en que concebimos el trabajo y desbloqueando 3 oportunidades de valor:

Reducción de Costos
Racionalización y reducción

Reducción del Tiempo de Ejecución
Optimización de procesos y eliminación de cuellos de botella humanos

Calidad del Servicio
Uso optimizado de diferentes herramientas y entornos

ALGUNOS EJEMPLOS [3]

25%
reducción en el tiempo de los ciclos de I+D

X10
reducción de costes en centros de llamadas con agentes

X50
aumento en la velocidad de creación de entradas de blog de marketing

40%
de aumento de la productividad en el departamento de TI

[3] Fuente: Conocimiento interno - Artefact

Históricamente, el sector financiero se construyó sobre un pilar esencial: la confianza. En un ambiente altamente regulado y globalmente interconectado, esta confianza siempre ha estado vinculada a la gestión estratégica de los datos de los clientes. Sin embargo, la dinámica ha cambiado. La lealtad, antes garantizada por un mercado menos centrado en el consumidor, ahora necesita ser conquistada diariamente.

En este nuevo escenario, la IA Generativa, aliada a la IA Agente, surge como una herramienta fundamental para fortalecer la confianza, optimizar operaciones y reinventar la experiencia del cliente. Ya sea en la personalización de servicios, en la mejora de la seguridad o en la eficiencia del back office, la adopción estratégica de estas tecnologías está redefiniendo el futuro del sector financiero.

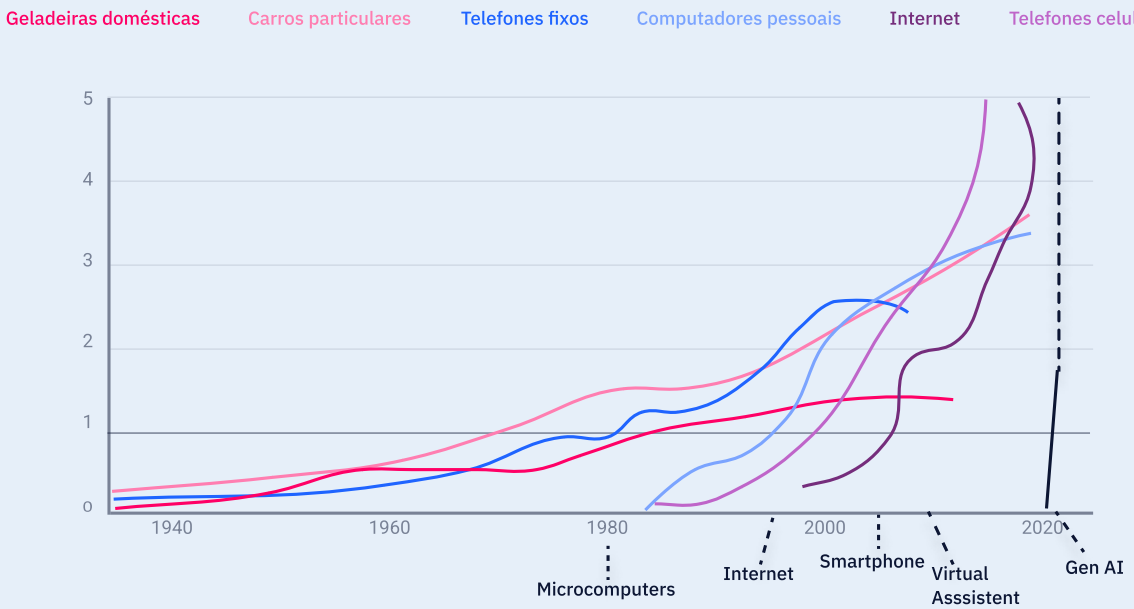
El estado del arte de la tecnología en el sector financiero

Las instituciones financieras apenas están comenzando a explorar el verdadero potencial de la IA Generativa (GenAI), una tecnología que va más allá del procesamiento de información, permitiendo la creación de contenido genuinamente nuevo a partir de comandos simples. A diferencia de las revoluciones tecnológicas anteriores, donde el ser humano siempre fue el agente de transformación –ya sea operando máquinas, dirigiendo cosechadoras o programando sistemas–, ahora la tecnología asume el papel de creador. Impulsada por Modelos de Lenguaje de Gran Escala (LLMs), la GenAI puede generar conocimientos, elaborar estrategias, producir informes e incluso desarrollar código sin la necesidad de intervención humana activa. Este es el verdadero punto de inflexión: la automatización ya no se limita a tareas repetitivas, sino que expande las fronteras de la creatividad y la toma de decisiones, inaugurando un nuevo paradigma para el sector financiero. . [4]

[4] GENERATIVE AI SURVEY - The Technology, the Rewards & the Risks - Artefact (Artefact-GenAI-Survey.pdf)

Los LLMs son modelos avanzados de inteligencia artificial, entrenados con grandes volúmenes de datos textuales, capaces de comprender y generar lenguaje natural con alta precisión. Después de adaptaciones específicas, estos modelos pueden ser utilizados para la interacción, creación de contenido y automatización de procesos textuales. En el ecosistema de la GenAI, también existen modelos orientados a otros formatos de datos, como imágenes, videos y audios, además de los Modelos Multimodales, que integran diferentes tipos de entradas simultáneamente, ampliando aún más las posibilidades de aplicación.

Así como la microcomputadora, Internet y el smartphone transformaron nuestra relación con la tecnología, la GenAI está redefiniendo la interfaz entre humanos y máquinas, haciendo la interacción más accesible, intuitiva y productiva. Las instituciones que adoptan esta tecnología de forma estratégica ya están cosechando beneficios, como ganancias de productividad, reducción de costos y mayor agilidad en la innovación. Los actores que avanzan rápidamente en la implementación de la GenAI están conquistando una ventaja competitiva significativa, mientras que otros, que aún no han comprendido la urgencia de esta transformación, se quedan atrás. [3]



[3] *Fonte: Conhecimento interno - Artefact*

Pero la sofisticación de esta tecnología va más allá de la simple generación de contenido.
El siguiente paso de la IA en el sector financiero implica sistemas más autónomos e inteligentes, capaces de actuar de manera estructurada y estratégica.

Este avance se refleja en el concepto de IA Agente, que amplía las capacidades de la GenAI al permitir que los modelos no solo generen respuestas, sino que tomen decisiones informadas, realicen tareas de forma continua y se adapten dinámicamente al contexto. Ya sea interactuando con el mundo real, accediendo a información reciente o utilizando APIs y herramientas más allá de los datos de entrenamiento, la IA Agente inaugura una nueva era de autonomía y eficiencia.

La siguiente imagen ilustra la transformación significativa que el uso de la IA Agente puede traer al proceso de selección de un ganador de una RFP (Request for Proposal), comparando el modelo tradicional con el modelo optimizado por agentes inteligentes. [3]

El análisis y la validación de RFPs (Requests for Proposal) sigue siendo, en muchos bancos, un proceso operacionalmente intenso, descentralizado y fuertemente dependiente de tareas manuales. Tradicionalmente, este proceso puede tardar hasta cuatro semanas, involucrando múltiples interfaces, varios profesionales y un flujo de trabajo fragmentado. Desde la descarga y lectura de documentos en diversos formatos hasta la verificación de la conformidad contractual, la agrupación por taxonomía y, finalmente, la decisión sobre el proveedor ganador, cada etapa exige un esfuerzo humano considerable — especialmente cuando se realiza a gran escala. Este modelo tradicional se caracteriza por una baja eficiencia, alta carga operativa y poca integración entre las etapas.

Diferencia entre el análisis y la validación de una solicitud de propuesta con o sin agentes

PROCESO ACTUAL

4 SEMANAS

Una interfaz por tarea

Estación de trabajo 1: Administrador

Recopilar y leer respuestas

Descargar documentos;

Leerlos todos

Extraer contenido (xls, pdf)

Copiar y pegar manualmente;
Buscar contenido en todos los PDF.

Estación de trabajo 2: Comprador

Verificar el cumplimiento del Acuerdo Marco.

Verificación manual (Ctrl + F) en todas las reglas

Asignar la oferta de productos en la taxonomía del grupo para la evaluación comparativa.

Encontrar similitudes navegando por la taxonomía

Analizar y validar al ganador de la solicitud de propuesta (RFP).

Redactar y leer conclusiones;
Varias reuniones

HOY CON AGENTIC AI

1 DIA

Una interfaz por proceso

Puesto de trabajo: Comprador

Análisis automatizado con Agente de Flujo de Trabajo

Recopilación y almacenamiento de documentos con una herramienta de ingesta automatizada de datos.

Extracción de contenido por agente de tareas con herramientas de lectura de PDF y OCR

Verificación del cumplimiento por parte del agente de tareas según el acuerdo marco + validación humana.

Correspondencia con elementos similares de la taxonomía y creación de informes comparativos realizados por un subagente de flujo de trabajo.

Análisis completo y propuesta de recomendaciones finales con agente de tareas con razonamiento.

Analizar y validar al ganador de la solicitud de propuesta (RFP).

Validación de la recomendación

Legenda:

Tarefa Manual

Tarefa Automatizada

Interface

Con la introducción de agentes autónomos de IA, este escenario se transforma radicalmente. El uso de la IA Agente permite consolidar el proceso en una única estación de trabajo, centralizada en el comprador, y reducir el tiempo total de ejecución a solo un día. Agentes especializados automatizan la ingesta y extracción de datos a partir de documentos en PDF o Excel, realizan análisis de conformidad con contratos maestros, comparan respuestas con puntos de referencia preestablecidos y proponen recomendaciones finales con una base lógica.

En lugar de múltiples profesionales ejecutando tareas repetitivas y susceptibles de error, la actuación humana pasa a concentrarse únicamente en la etapa final de validación —con un enfoque en el análisis crítico y la toma de decisiones. La ganancia de eficiencia operativa es evidente, pero el impacto estratégico va más allá: al liberar tiempo y capacidad cognitiva de los equipos, los bancos pasan a tener más agilidad para responder a oportunidades de mercado, reducir riesgos y tomar decisiones más informadas.

> **Esta transformación ilustra con claridad el valor de la IA Agente: más que automatizar tareas, reposiciona el papel del humano en las operaciones financieras, elevando la calidad del proceso de toma de decisiones y acelerando el ritmo de la innovación.**

CAPÍTULO 2

IA Agente & Oportunidades

¿Qué es la IA Agente y por qué representa una oportunidad tan grande?

La transformación que vimos en el proceso de RFP es solo una entre cientos de aplicaciones posibles de una nueva generación de inteligencia artificial: la IA Agente. Si la IA tradicional automatizó tareas específicas y la GenAI trajo creatividad y generación de contenido bajo demanda, la IA Agente lleva esta revolución un paso más allá —con agentes capaces de tomar decisiones, actuar con autonomía y adaptarse en tiempo real al entorno.

Ya no se trata de responder a preguntas o generar textos con base en comandos: se trata de ejecutar misiones complejas de principio a fin, conectando sistemas, accediendo a datos dinámicos, operando herramientas y aprendiendo de cada interacción.

Es esta capacidad de operar con propósito, contexto y coordinación lo que hace de la IA Agente una oportunidad estratégica gigantesca para el sector financiero.

Y aunque el concepto pueda parecer nuevo, su evolución ya está en curso —con hitos técnicos que avanzan año tras año y nos acercan rápidamente a un escenario en el que los agentes inteligentes podrán actuar como verdaderos colaboradores digitales. La siguiente línea de tiempo muestra cómo esta transformación está ocurriendo a un ritmo acelerado.

Los «agentes» han evolucionado rápidamente y sus nuevas capacidades han abierto el camino a oportunidades transformadoras de automatización de procesos.

	2023: Asistir	2024: Actuar	2025: Automatizar
Tipo de agentes	Agentes de recuperación	Agentes de acción	Agentes de orquestación	Agentes de percepción
Capacidades tecnológicas	RAG	Uso de herramientas (APIs)	Colaboración entre múltiplos agentes / humanos	Uso de interfaz gráfica (GUI)
Observabilidad	Monitoreamiento	Evaluación	Supervisión
Diseñar soluciones con GenAI	El agente recibe el contexto para responder.	El agente puede planificar y ejecutar acciones con herramientas según sea necesario.	Agentes especializados trabajan juntos y escalan a humanos cuando es necesario.	Los agentes pueden interactuar con interfaces gráficas sin necesidad de API.

LangChain LangGraph crewai Runner H microsoft/JARVIS MetaGPT Claude

Este viaje evolutivo se puede entender a partir de tres grandes hitos:

Asistir Actuar Automatizar

Cada fase representa una ampliación concreta de las capacidades de los agentes de IA y de las oportunidades que estas desbloquean para el sector financiero:

2023: Asistir

En esta primera etapa, surgen los retrieval agents, especializados en recuperar información relevante a partir de contextos proporcionados. Con el soporte de tecnologías como RAG (retrieval-augmented generation) y herramientas como LangChain, estos agentes son capaces de buscar datos, consolidar información y responder con precisión. La observabilidad en este momento todavía está restringida al monitoreo básico del rendimiento de los modelos.

2024: Actuar

Con los avances en el uso de APIs, emergen los acting agents, capaces de ejecutar acciones concretas en herramientas externas. Esta evolución marca un giro operativo: los agentes no solo responden, sino que también actúan con base en lo que interpretan, integrándose a sistemas como CRMs, ERPs y plataformas analíticas. En este mismo año, surgen los orchestration agents, que orquestan múltiples agentes o colaboran con humanos.

2025: Automatizar

- En 2025, los agentes evolucionan hacia interacciones más complejas con sistemas, dando origen a los perception agents. La principal diferencia radica en la capacidad de estos agentes para navegar por interfaces gráficas (GUIs), sin depender exclusivamente de las APIs para ejecutar acciones. Esto permite automatizar flujos que involucran múltiples aplicaciones o sistemas legados —algo especialmente relevante en instituciones financieras con arquitecturas de TI más heterogéneas. Tecnologías como Runner H y Claude viabilizan este avance, mientras que los mecanismos de supervisión activa garantizan el control y la seguridad operativa.

Este roadmap tecnológico demuestra con claridad que la IA Agente no es una tendencia lejana. Ya está moldeando nuevas formas de operar, tomar decisiones y generar valor. Para los líderes del sector financiero, entender estas fases es esencial para estructurar estrategias robustas, identificar casos de uso prioritarios y garantizar una adopción escalable y segura de esta tecnología transformadora.

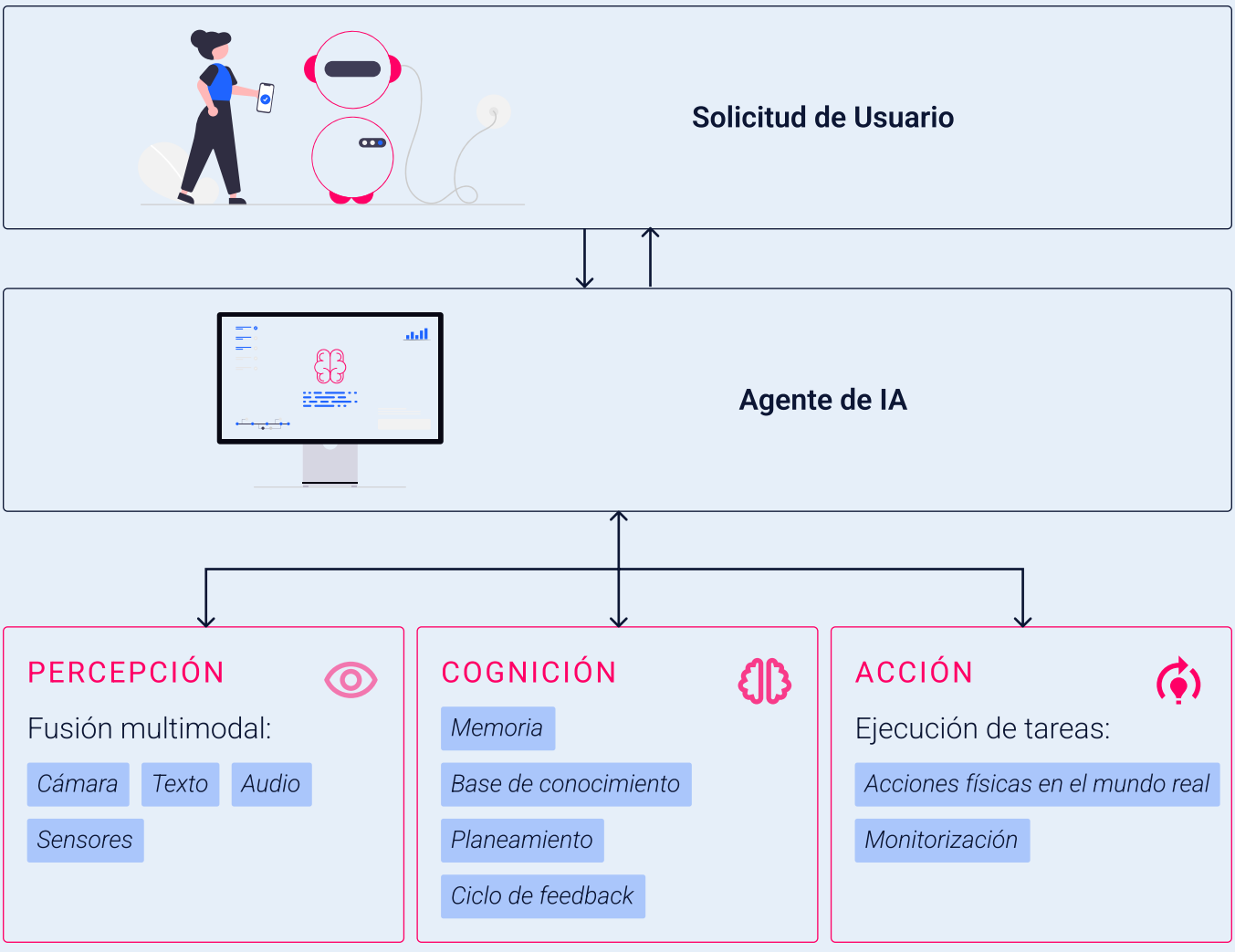
Para comprender el potencial real de la IA Agente, es fundamental entender cómo estos agentes operan en la práctica. A diferencia de las soluciones basadas únicamente en grandes modelos de lenguaje (LLMs), los agentes son entidades autónomas compuestas por tres componentes principales:

☑ **Percepción:** ☑ **Cognición:** ☑ **Acción:**

Cada uno de estos bloques tiene un papel esencial en la forma en que el agente interactúa con el entorno, toma decisiones y ejecuta tareas de forma continua y adaptable.

AGENTE, NOVA DEFINICIÓN

Una entidad autónoma que percibe, razona, actúa y se adapta a los cambios.



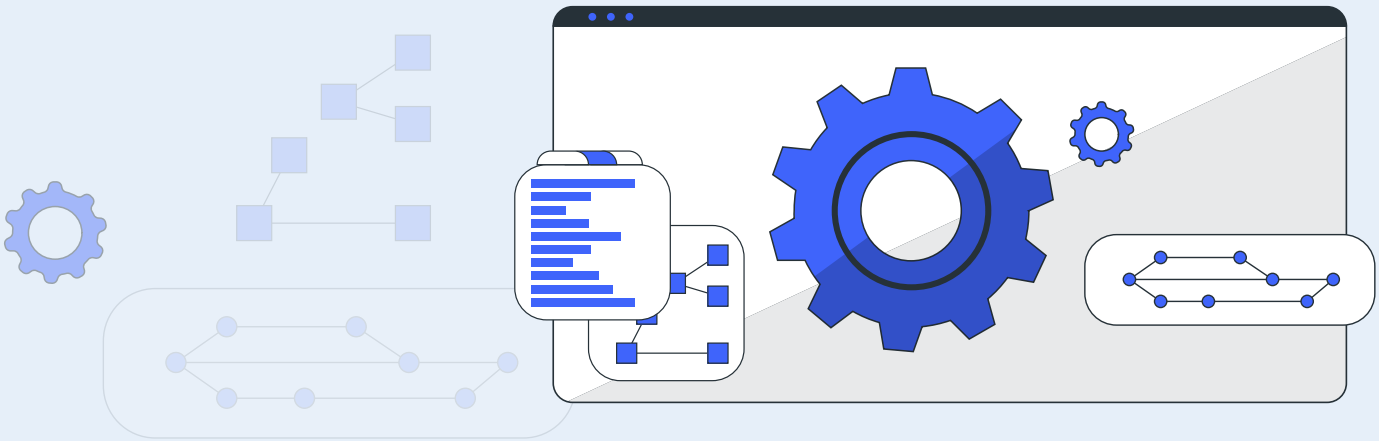
La percepción permite que el agente recoja información de diferentes fuentes, a través de la llamada fusión multimodal. Esto incluye datos de texto, audio, imágenes (cámaras) y sensores diversos. En contextos empresariales, por ejemplo, esto significa que un agente puede cruzar información estructurada de sistemas (como ERPs y CRMs) con datos no estructurados (como correos electrónicos, grabaciones de atención y documentos), creando una visión rica y contextualizada de la situación.

> **Con base en los datos percibidos, el componente de cognición entra en acción. Él combina memoria, bases de conocimiento y capacidad de planificación para interpretar las señales recibidas, tomar decisiones y adaptarse. Un aspecto fundamental aquí es el ciclo de retroalimentación (feedback loop), que permite al agente aprender de sus propias acciones y refinar su comportamiento a lo largo del tiempo —esencial para lidiar con ambientes complejos, como el mercado financiero.**

Por último, el agente es capaz de ejecutar tareas con base en el razonamiento desarrollado. Esto puede incluir desde interacciones con sistemas a través de API, navegación en interfaces gráficas, hasta acciones físicas en ambientes conectados. Además, el agente mantiene la capacidad de monitorizar los resultados de sus acciones, garantizando una operación más segura, eficiente y auditable.

Esta arquitectura permite que los agentes actúen de forma proactiva, respondan a los cambios en tiempo real y evolucionen con base en la experiencia —lo que los convierte en una de las innovaciones más prometedoras de la IA moderna.

La IA Agente permite automatizar flujos de trabajo complejos de principio a fin a través de la coordinación entre múltiples agentes. Un ejemplo claro y aplicable al sector financiero es el procesamiento automatizado de solicitudes de crédito.



En este flujo, los agentes actúan de forma orquestada, combinando razonamiento (Reason) y ejecución (Act) en una secuencia lógica de etapas, como muestra el diagrama a continuación:

Al recibir una solicitud de crédito, se inicia una cadena de agentes, cada uno responsable de una parte del proceso:

Al recibir una solicitud de crédito, se inicia una cadena de agentes, cada uno responsable de una parte del proceso:

1. Extracción de la información

Un agente extrae automáticamente datos relevantes de la solicitud utilizando OCR, traducción e interpretación de interfaces. Él transforma las entradas brutas en información estructurada.

2. Enriquecimento de datos (CRM)

A continuación, otro agente valida y enriquece estos datos con información del CRM, añadiendo contexto al perfil del cliente.

3. Simulación de tasas y condiciones

Con los datos listos, un tercer agente ejecuta una simulación de crédito basada en modelos de machine learning para prever la tasa ideal.

4. Optimización y personalización de la oferta

En caso de que la tasa no sea ideal, el sistema identifica alternativas (como la exigencia de un fiador o una mayor entrada), simula nuevamente y optimiza la propuesta de forma personalizada.

5. Respuesta al cliente

Por último, un agente elabora y envía automáticamente la respuesta al cliente, conteniendo la simulación, el análisis y la sugerencia recomendada.

Este ejemplo muestra cómo las cadenas de agentes pueden replicar procesos cognitivos, pero con una velocidad, precisión y escalabilidad superiores. Cada agente entiende su papel, consulta sistemas, toma decisiones con base en reglas o modelos, aprende de los resultados y colabora con los demás para cumplir el objetivo final.

Sin embargo, para que esta sofisticación sea implementada de manera segura y eficaz, algunas consideraciones son cruciales. Es indispensable establecer frameworks de gobernanza robustos para asegurar que los agentes actúen dentro de límites éticos y reglamentarios, previniendo errores y violaciones de datos, así como la construcción de "barreras de seguridad" que imponen límites a lo que se les permite a los agentes y previenen la alucinación. Además, la integración de agentes aumenta la superficie de ataque cibernético, exigiendo soluciones como la autenticación contextual y la monitorización continua. A pesar de su autonomía, los agentes también requieren supervisión humana en decisiones de alto impacto, garantizando que sus acciones estén alineadas con los objetivos organizacionales.

> **Al equilibrar autonomía, seguridad y gobernanza, el uso de agentes en el sector financiero promete inaugurar una nueva era de eficiencia e innovación, transformando profundamente la forma en que los servicios financieros son ofrecidos y consumidos.**

Según el Citi GPS, la IA Agente marca un cambio hacia la economía del “hazlo por mí”, en la que la tecnología asume la toma de decisiones y la ejecución de tareas. La mayor autonomía de [6] IA Agente, en relación con la GenAI, le permite lidiar con procesos repetitivos y con un uso intensivo de datos, optimizando los flujos de trabajo, mejorando la conformidad y perfeccionando la toma de decisiones. En la siguiente imagen, se hace más explícito lo que significa la transición entre la GenAI y la IA Agente y sus posibles aplicaciones en el sector de servicios financieros. [5]

---

[5] Extraído de: Citi GPS Agentic AI Finance & the ‘Do It For Me’ Economy - January 2025

[6] Adaptado del artículo:: Agentic AI will transform financial services with autonomy, efficiency, and inclusion. - Medium (Agentic AI will transform financial services with autonomy, efficiency, and inclusion. | by Sahaj Godhani | InsiderFinance Wire)

GENAI

Habilita

Agentic AI

Más **autónoma**.

Mayor necesidad de **capacidad computacional**.

Mayor impacto en la **productividad.**

Mayor **eficacia en la ejecución de tareas.**

Mayor nivel de **complejidad de las tareas.**

Interacción con el cliente

Compromiso con el cliente: automatiza herramientas personalizadas de planificación financiera e interfaces de aplicaciones personalizadas.

Gestión de relaciones: optimiza las comunicaciones con los clientes y adapta las estrategias de compromiso.

Asesoramiento financiero personal: proporciona asesoramiento financiero en tiempo real, ajustándose en función de los hábitos de consumo.

Producto y precios (Product & Pricing)

Evaluación crediticia y concesión de préstamos: evalúa el historial crediticio, personaliza las ofertas de préstamos y gestiona las cuentas de alto riesgo de forma autónoma.

Precios dinámicos: ajusta los precios y las ofertas de retención en función del comportamiento del cliente en tiempo real.

Cumplimiento normativo y prevención del fraude

Supervisión de transacciones: supervisa los riesgos de blanqueo de capitales (AML), identifica e interviene de forma dinámica en transacciones de alto riesgo.

Siniestros y suscripción: automatiza la clasificación de siniestros y perfecciona las evaluaciones de riesgo.

Monitorización del riesgo financiero: rastrea las amenazas en tiempo real y recomienda medidas de mitigación.

Prueba de conformidad de software: identifica errores, garantiza la conformidad e implementa actualizaciones.

Automatización de procesos y calidad: automatiza la clasificación de siniestros y detecta anomalías operativas.

Inteligencia de Mercado

Análisis del mercado competitivo: realiza un seguimiento de las estrategias de la competencia y proporciona información para realizar ajustes estratégicos.

Seguimiento de las tendencias del mercado: supervisa los cambios en el mercado y alerta a los analistas sobre los riesgos u oportunidades emergentes.

En el sector financiero, Agentic AI significa, por ejemplo, asistentes inteligentes que identifican riesgos, ajustan carteras, automatizan procesos complejos e incluso gestionan interacciones con clientes sin necesidad de intervención humana.

Con esta sofisticación tecnológica, las instituciones pueden ampliar sus operaciones, ofrecer personalización masiva y aumentar significativamente su eficiencia, manteniendo la confianza del cliente en un mercado cada vez más dinámico.

El futuro del sector no lo definirá solo quien adopte la IA, sino quien sepa explotar al máximo su sofisticación e inteligencia estratégica.

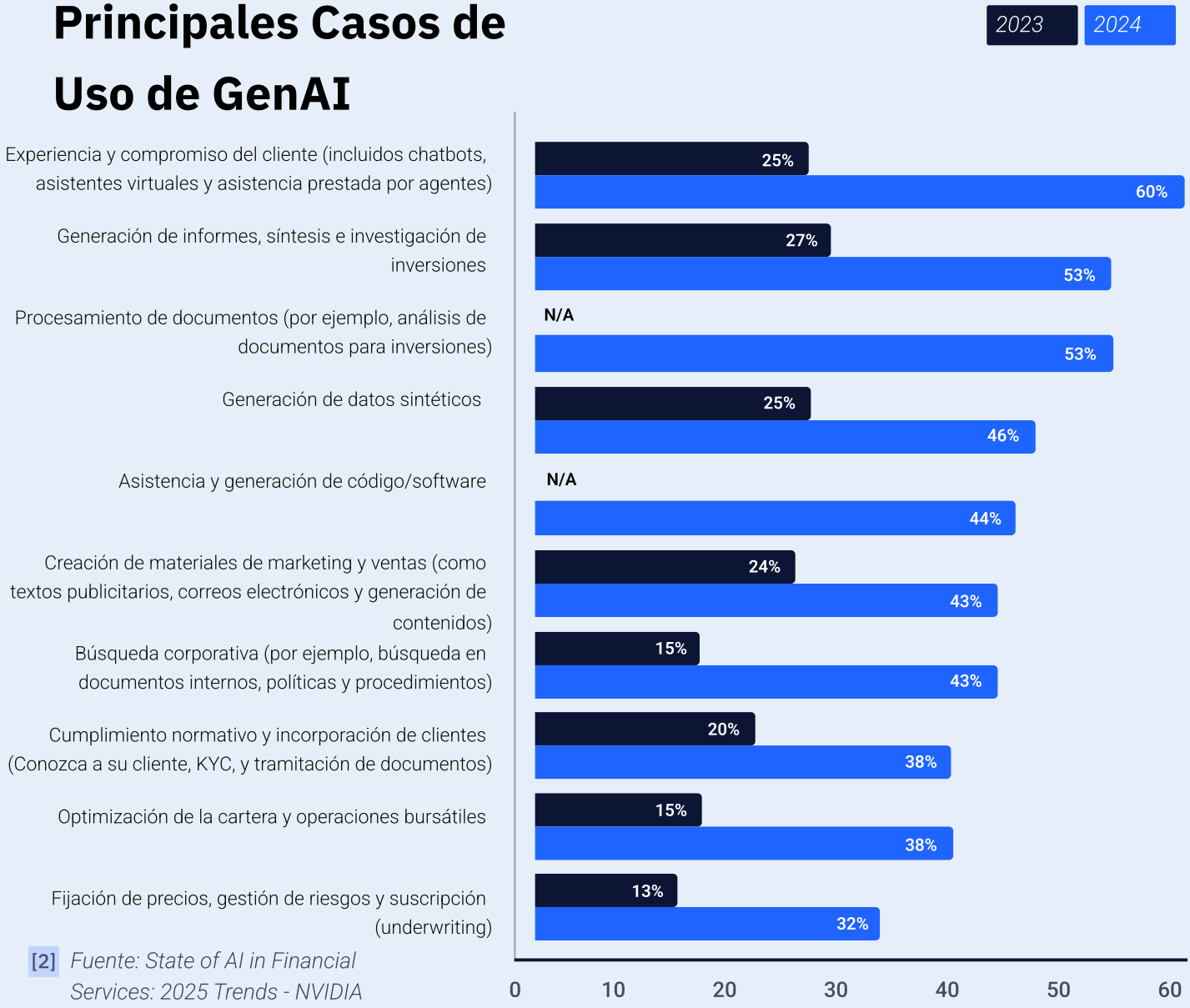
CAPÍTULO 3

GenAI y Agentic AI y sus Aplicaciones

GenAI e IA Agente: aplicaciones y el Arte de lo Posible

En el sector financiero, la adopción de soluciones basadas en IA generativa ha ganado una tracción acelerada, con empresas que buscan explorar nuevos casos de uso que impulsen la eficiencia, la personalización y la innovación. Conforme se evidencia en la imagen [2], hubo un salto expresivo de 2023 a 2024 en diversos frentes. Se destacan áreas como la Experiencia del Cliente, que evolucionó del 25% al 60%, y la Generación de Informes y la Investigación, que saltó del 27% al 53%. Otros casos, como el Procesamiento de Documentos, la Asistencia en la construcción de códigos, y la Búsqueda Empresarial, también avanzaron rápidamente, reflejando la creciente apuesta del mercado financiero por el potencial transformador de la GenAI.

Principales Casos de Uso de GenAI

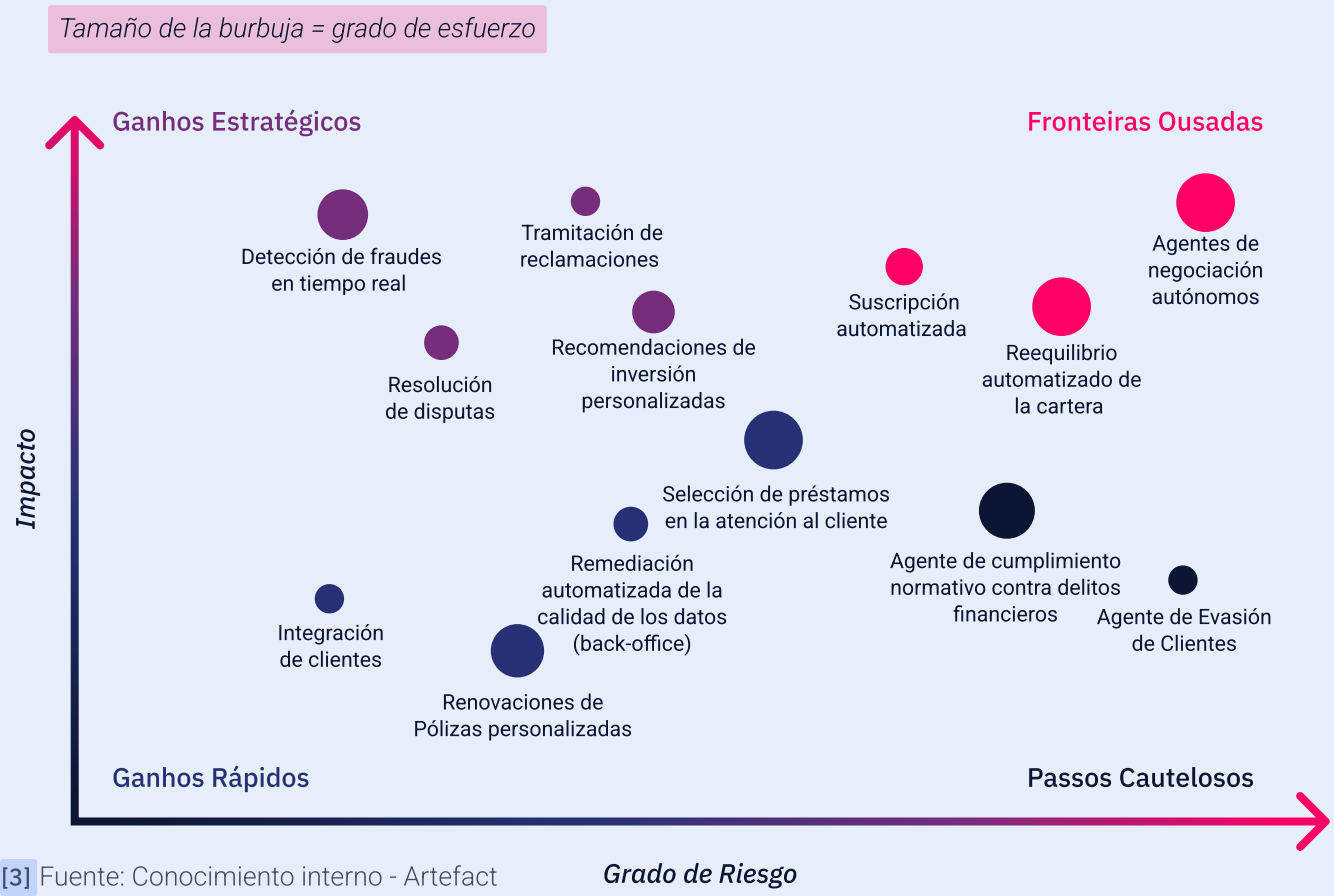


[2] *Fuente: State of AI in Financial Services: 2025 Trends - NVIDIA*

Las aplicaciones de GenAI en el sector financiero ya son una realidad, como muestra el estudio. Además, esta tecnología crea las bases para que una serie de casos de uso aún más transformacionales puedan ser desarrollados, a través de la IA Agente.

Para guiar la adopción estratégica de soluciones con IA Agente en este sector, es esencial considerar tres dimensiones centrales: riesgo, impacto y complejidad. La imagen a continuación ilustra esta lógica, en la que el eje vertical representa el impacto potencial de la solución, mientras que el eje horizontal indica el grado de riesgo involucrado. Por último, el tamaño de la burbuja refleja el esfuerzo necesario para construir el agente.

Su roadmap de Agentic AI se perfeccionará mediante una combinación de riesgo, impacto y complejidad.



[3] Fuente: Conocimiento interno - Artefact

En este escenario, dos casos se destacan como quick wins. El Client Onboarding permite automatizar la jornada inicial del cliente con agentes que recogen documentos, validan información, integran sistemas (como CRMs y antifraude) y notifican al equipo interno o al propio cliente. Por otro lado, el Personalized Policy Renewals hace que los agentes revisen historiales de pólizas, para sugerir renovaciones personalizadas con base en perfiles y enviar comunicaciones proactivas al cliente.

En la capa de Strategic Wins, se destacan aplicaciones que exigen un esfuerzo moderado, pero que entregan un alto impacto para el negocio. Casos como Real Time Fraud Detection, Disputes Resolution y Claims Processing optimizan procesos críticos con ganancias significativas en velocidad y precisión, sin el nivel de riesgo de las fronteras más avanzadas.

En las Bold Frontiers, están los casos de uso que combinan alto impacto con mayor riesgo y complejidad —como Autonomous Trading Agents, Auto Portfolio Rebalancing y Automated Underwriting. Estas aplicaciones representan el futuro de la inteligencia autónoma en el sector, exigiendo arquitecturas robustas y una gobernanza sofisticada.

Para facilitar la priorización y estructuración de iniciativas con IA Agente, es útil organizar los casos de uso en grandes frentes de valor.

Abajo, presentamos una visión estructurada que agrupa los principales casos en cuatro categorías clave:

- **Gestión de Riesgos**
- **Automatización de Procesos**
- **Atención al Cliente y Personalización**
- **Toma de Decisiones e Insights**

Para cada caso, se indica dónde genera más valor, con marcas de adhesión total o parcial. Esto permite visualizar rápidamente cómo cada iniciativa se conecta a objetivos estratégicos y operacionales, ayudando en la priorización y aplicación de soluciones basadas en IA Agente.

Los Agentes de Gestión de Riesgos están orientados a la mitigación de riesgos financieros, operacionales y regulatorios, con capacidad de actuación continua y aprendizaje constante. La Automatización de Procesos engloba a los agentes que optimizan tareas rutinarias u operacionales, reduciendo errores y aumentando la eficiencia con flujos más responsivos y fluidos. La Atención al Cliente y Personalización incluye a los agentes que interactúan con canales y personas para ofrecer experiencias personalizadas e integradas en tiempo real. Por último, la Toma de Decisiones e Insights cubre a los agentes enfocados en el análisis avanzado y el soporte a la decisión, permitiendo que la inteligencia artificial actúe de forma activa en la construcción de estrategias y en la operación del negocio.

Casos de Uso	Gestión de Riesgos	Automatización de Procesos	Atención al Cliente y Personalización	Toma de Decisiones e Insights
Detección de fraudes en tiempo real	●	○	○	○
Agente de cumplimiento normativo contra delitos financieros	●	○	○	○
Triagem de Empréstimos no Atendimento ao Cliente	●	○	○	○
Tramitación de reclamaciones (o siniestros)	○	●	○	● *Con bucles de retroalimentación)*
Suscripción automatizada	● *[parcial]*	●	○	○
Corrección automatizada de la calidad de los datos (back-office)	○	●	○	○
Resolución de disputas	● *[parcial]*	●	○	○
Integración de clientes	○	● *[parcial]*	●	○
Renovaciones de pólizas personalizadas	○	● *[parcial]*	●	●
Recomendaciones de inversión personalizadas	○	○	●	○
Agente de Evasión de Clientes	● *[parcial]*	○	○	●
Reequilibrio automático de la cartera	● *[parcial]*	○	○	●
Agentes de negociación autónomos	● *[alta complejidad]]*	○	○	● *[decisiones autónomas]*

La categorización anterior permite visualizar de forma clara dónde cada caso de uso de IA Agente puede generar más valor dentro de las operaciones financieras. Para profundizar en el análisis, a continuación, detallamos cada uno de los cuatro frentes principales, aportando ejemplos concretos de agentes, sus funciones y el impacto esperado. Esta visión ayuda a tangibilizar cómo los agentes operan en la práctica, evidenciando las ganancias de eficiencia, mitigación de riesgos y personalización que pueden traer a diferentes áreas del sector.

Gestión de Riesgos

Casos-clave

Detección de fraudes en tiempo real

Agente que monitoriza transacciones financieras en tiempo real, identificando patrones sospechosos de fraude (como comportamiento fuera del perfil del cliente) y activando inmediatamente bloqueos o investigaciones automatizadas.

Agente de cumplimiento normativo contra delitos financieros

Agente especializado en compliance que verifica continuamente la conformidad de las transacciones con normas como AML (anti-lavado de dinero) y KYC (conozca a su cliente), cruzando datos con listas de sanciones y perfiles de riesgo.

Triagem de Empréstimos no Atendimento ao Cliente

Agente que realiza el análisis preliminar de crédito o préstamo en el front office, combinando datos estructurados y no estructurados para calcular puntuaciones, validar documentación y sugerir decisiones iniciales.

Estos casos tienen un alto impacto e implican un riesgo considerable, siendo áreas clásicas de uso de la inteligencia artificial para la mitigación de riesgos financieros y operacionales. Con la IA Agente, es posible:

- Crear agentes que monitorizan transacciones en tiempo real, detectando patrones anómalos y activando flujos de trabajo de investigación automática o semiautomática.
- Habilitar un compliance continuo con agentes que verifican la adhesión a políticas y regulaciones.
- Automatizar la pre-selección de clientes o préstamos con base en puntuaciones de riesgo, historial y comportamiento.

Automatización de Procesos

Casos-clave

Tramitación de reclamaciones

Agente responsable de recibir, clasificar, validar y tramitar siniestros automáticamente, utilizando reconocimiento de documentos, reglas de negocio y datos históricos.

Suscripción automatizada:

Agente que analiza automáticamente el perfil del proponente, consulta bases externas, aplica reglas actuariales y proporciona una decisión inicial de aceptación o rechazo del riesgo.

Corrección automatizada de la calidad de los datos

(backoffice)

Agente enfocado en la identificación y corrección de problemas de calidad de datos (DQ), como datos inconsistentes o ausentes, manteniendo la integridad de las bases y alimentando correctamente los sistemas analíticos.

Resolución de disputas

Agente que actúa en la mediación y resolución de disputas (por ejemplo, en pagos o transacciones bancarias), organizando evidencias, evaluando la situación y sugiriendo soluciones o encaminamientos.

Estos casos se benefician directamente de la automatización inteligente de etapas repetitivas, optimizando tiempo y costos. La propuesta con la IA Agente es:

- Desarrollar agentes especialistas que ejecutan tareas como validación de documentos, cruce de datos y aplicación de reglas de negocio.
- Incorporar lógica adaptativa, permitiendo que los agentes aprendan excepciones y mejoren continuamente.
- Integrar estos agentes con sistemas BPM (Business Process Management) o ERPs para la orquestación de principio a fin del proceso.

Atención al Cliente y Personalización

Casos-clave

Agente de Evasión de Clientes
Agente que monitoriza comportamientos de los clientes (como la caída en el uso de productos, quejas o cambios de patrón) para prever el riesgo de churn y sugerir acciones de retención personalizadas.

Integración de Çlientes
Agente que conduce el proceso de entrada de nuevos clientes de forma adaptativa, solicitando documentos conforme al perfil, explicando etapas e integrando datos automáticamente en sistemas internos.

Renovaciones de Pólizas Personalizadas
Agente que evalúa el historial del asegurado, cambios de perfil y contexto actual para ofrecer renovaciones personalizadas y proactivas, ajustando coberturas y precios según la necesidad.

Recomendaciones de Inversión Personalizadas
Agente que sugiere inversiones personalizadas con base en el perfil de riesgo, objetivos, preferencias y eventos de mercado, pudiendo incluso interactuar en lenguaje natural con el usuario.

Estos casos representan iniciativas de personalización y compromiso proactivo del cliente. Usando la IA Agente:

- Es posible crear agentes que identifican señales tempranas de churn e inician acciones de mitigación, como ofertas y contacto personalizado.
- Los agentes de onboarding hacen que el proceso sea más fluido, adaptando el flujo a las necesidades del cliente con base en perfiles y comportamiento
- Los agentes de recomendación construyen jornadas personalizadas con base en datos históricos, preferencias y contexto actual del usuario.

Toma de Decisiones e Insights

Casos-clave

Rebalanceamento Automático de Portfólio

Agente que evalúa carteras de inversión o productos y, con base en criterios predefinidos (como perfil de riesgo o eventos externos), propone y ejecuta reubicaciones automáticas.

Recomendaciones de Inversión Personalizadas

(También presente en Personalización) — aquí el foco es en la explicabilidad y generación de conocimientos en tiempo real para auxiliar al inversor en la toma de decisiones autónoma.

Agentes de Negociación Autónomos

En este contexto, el agente de siniestros recoge la retroalimentación continua de las decisiones y resultados, aprendiendo de errores y aciertos para refinar sus futuras decisiones.

Estos casos se enfocan más en mejorar la capacidad de tomar decisiones más fundamentadas, asertivas y eficientes a través de la construcción de agentes que se organizan en un flujo de trabajo para levantar todos los puntos necesarios. Siendo así, con la IA Agente:

- Los agentes financieros pueden reevaluar portafolios automáticamente, con base en eventos de mercado y perfiles de riesgo.
- En sinergia con herramientas analíticas, los agentes ofrecen recomendaciones con explicaciones (XAI), aumentando la confianza de los usuarios.
- Los agentes pueden actuar como copilotos de decisión, preparando conocimientos, simulaciones e incluso sugiriendo acciones dentro de un flujo de toma de decisiones.

A medida que los casos de uso de la IA Agente se multiplican, una pregunta estratégica cobra fuerza: ¿cómo transformar estas iniciativas en soluciones escalables e integradas en el día a día de la organización? La respuesta está en la construcción de productos digitales inteligentes, en los que los agentes son diseñados para operar dentro de flujos de trabajo reales, interactuando con sistemas, datos y personas. Este enfoque garantiza que la IA deje de ser solo una "capa de recomendación" para convertirse en parte activa de la operación y de la experiencia ofrecida al cliente.

INTEGRACIÓN CON FLUJOS DE TRABAJO DE NEGOCIO A TRAVÉS DE LA IA AGENTE

La clave para extraer valor de estos casos es tratarlos como productos digitales compuestos por agentes inteligentes, cada uno integrado a los procesos principales de la empresa. Esto implica:

- Definir roles específicos para los agentes en cada jornada (ejemplo: evaluador de riesgo, analista de siniestros, recomendador de inversiones).
- Proyectar interacciones entre agentes, humanos y sistemas, con gobernanza y supervisión.
- Monitorizar KPIs de rendimiento y aprendizaje continuo de los agentes, como parte de un ciclo ágil de evolución del producto.

De esta forma, en la siguiente tabla tenemos ejemplos de productos digitales que combinan los casos de uso presentados anteriormente con un enfoque en los desafíos más específicos de cuatro segmentos dentro de los servicios financieros: Banca Minorista, Banca de Inversión, Seguros y Pagos. De esta forma, podemos ver con más claridad cómo sería la aplicación de la IA Agente en la construcción de productos digitales que atienden a los desafíos específicos de cada negocio.

Categorías clave para la generación de valor

Categoría	1. Automatización de procesos	2. Gestión de riesgos	3. Atención al cliente, compromiso y personalización	4. Toma de decisiones e insights basados en IA
Minorista	Analista de mesa de crédito virtual capaz de orquestar y automatizar todo el proceso de concesión de crédito de principio a fin.	Gestor de riesgo operativo virtual, que supervisa transacciones, investiga perfiles y responde a eventos en tiempo real.	Conserje financiero digital que entiende segmentos y perfiles, ejecuta y recomienda productos basándose en el contexto individual.	Agente autónomo de apoyo a la toma de decisiones, que recopila datos, genera análisis, explica tendencias, simula escenarios y sugiere acciones.
Banco de inversión	Deal Book Automation Agent que automatiza la recopilación de datos, el análisis y la generación de materiales para respaldar las transacciones.	Risk Intelligence Agent, capaz de supervisar el mercado en tiempo real, identificar exposiciones críticas y recomendar acciones.	Gestor de relaciones virtual que conoce al cliente, se anticipa a sus necesidades, responde de forma consultiva y ejecuta tareas.	Deal Hunter Virtual, que analiza tendencias, identifica oportunidades y sugiere ideas de originación con explicaciones estructuradas.
Seguros	Asesor de reclamaciones capaz de recibir, analizar, clasificar, decidir y comunicar el estado de un proceso de siniestro.	Risk Sentinel, capaz de monitorear señales de riesgo en tiempo real, identificar patrones anómalos, generar alertas, simular impactos y sugerir acciones mitigadoras.	Conserje digital 24/7, capaz de anticipar necesidades, trazar perfiles, comprender estilos de vida y personalizar recomendaciones de seguros.	Agente autónomo que utiliza datos históricos, tendencias de mercado y comportamiento de los clientes para generar conocimientos y optimizar la cartera de seguros y productos.
Pagos	Agente autónomo de conciliación financiera, que coordina, automatiza, analiza y actúa sobre el proceso de conciliación y liquidación de pagos.	Risk Guardian, que actúa como un centinela digital, supervisa constantemente las transacciones, detecta patrones de riesgo y actúa en tiempo real.	PayBot Concierge combina asistencia técnica, asesoramiento financiero y ofertas personalizadas basadas en los patrones de transacción.	Estratega de pagos capaz de analizar grandes volúmenes de datos, generar recomendaciones, simular escenarios e interactuar con ejecutivos en lenguaje natural.

> **La IA Agente eleva la construcción de productos digitales en el sector financiero al permitir que los agentes autónomos ejecuten tareas complejas —como monitorizar riesgos, personalizar ofertas o automatizar procesos— siempre alineados con los objetivos de negocio. A diferencia de la IA tradicional, que entrega análisis o previsiones, la IA Agente transforma estos modelos en motores de acción continua, ampliando su impacto y acelerando la generación de valor en áreas como crédito, seguros e inversiones.**

En los próximos años, se espera que los agentes inteligentes actúen de forma cada vez más colaborativa e integrada, redefiniendo cómo las instituciones financieras orquestan datos, decisiones y experiencias para clientes y operaciones.

CAPÍTULO 4

Casos de éxito en el sector financiero

Casos de Éxito: cómo la IA está redefiniendo el Sector Financiero

Entre todas las aplicaciones posibilitadas por la evolución de las herramientas de GenAI e IA Agente, Artefact ya ha desarrollado implementaciones prácticas para diversos clientes del sector financiero a nivel global. Estos casos de uso no son solo proyecciones o tendencias futuras — son aplicaciones reales, ya implementadas por grandes actores del sector. Las instituciones líderes están utilizando la GenAI para transformar sus operaciones e interacciones con los clientes, generando resultados concretos en eficiencia, personalización e innovación.

OPTIMIZANDO LA EXPERIENCIA DEL CLIENTE: UN NUEVO PARADIGMA EN ATENCIÓN AL CONSUMIDOR

La transformación digital ha impuesto a las empresas financieras desafíos crecientes en la gestión de sus interacciones con los clientes. En un mercado altamente competitivo, la capacidad de responder a las dudas de los consumidores con velocidad, precisión y en un lenguaje que se adapte a cada contexto de los consumidores se ha convertido no solo en un diferencial estratégico, sino en un factor determinante para la fidelización. En este contexto, Artefact identificó una oportunidad significativa: la construcción de un chatbot que sustituyera la necesidad de interacción humana constante, garantizando respuestas precisas, seguras y a tiempo en un gran banco de inversiones europeo.

Desafios y Obstáculos

A pesar de ser prometedor, el proyecto enfrenta barreras bien definidas en su implementación, independientemente de la estructura de la compañía. La primera de ellas es atender a la creciente expectativa de los usuarios de respuestas accesibles, rápidas y de calidad, un elemento crítico para la satisfacción del cliente. Paralelamente, la solución necesita estar en conformidad con estándares rigurosos de seguridad, que incluyen no solo la protección de datos sensibles, sino también la adhesión a las regulaciones específicas del sector financiero.

Desarrollando la Solución

La solución propuesta se basa en un modelo de GenAI diseñado para operar con alta eficiencia y seguridad. El modelo se apoya en una base de conocimiento de preguntas y respuestas (Q&A), estructurada a partir de datos relevantes de la organización. La solución también utiliza algoritmos de filtrado avanzados para interpretar la pregunta del usuario, extraer la información más pertinente de la base de conocimiento y generar respuestas contextualizadas y precisas. Complementaria a las estructuras necesarias para el modelo, se destaca la capacidad de adaptar la solución propuesta tanto en ambientes de nube (Azure, AWS y GCP) como on-premise, preservando la integridad de los datos y alineándose con las exigencias de seguridad.

INTERFAZ DE USUARIO

El usuario hace una pregunta.

Aquí están sus respuestas y las fuentes.

Filtro de entrada

Verifica que el usuario cumpla con las políticas de la aplicación en cuanto a preguntas, relevancia y seguridad de la entrada.

Validador de Salida

Verifica la salida de los flujos y garantiza que sea adecuada.

Orquestrador

Decide qué agente será necesario para procesar la entrada del usuario.

Modelo de Conversación

Genera la salida final (texto) que se mostrará al usuario.

Texto - para - SQL

Transforma la pregunta en una consulta SQL.

Ejecución de la Consulta

Recupera datos relevantes de la base de datos utilizando la consulta anterior.

Adaptado de Artefact One Pager – Agente de inteligencia empresarial aumentada aplicado al sector farmacéutico.

Benefícios Alcanzados

Como resultado de la implementación del bot de conversación en diferentes escenarios, nuestros clientes alcanzaron una reducción del 90% en el tiempo promedio de respuesta a los usuarios, promoviendo una atención más ágil y satisfactoria. En términos financieros, el proyecto logró generar un ahorro superior a los 5 millones de euros en el costo operacional del equipo de atención al cliente, reafirmando el valor estratégico de la GenAI para optimizar procesos y asignar recursos de forma más eficiente.

Destacados

La implementación de esta solución evidencia cómo las tecnologías de GenAI pueden redefinir paradigmas en el sector financiero, combinando eficiencia operacional y excelencia en la atención al cliente. Superando desafíos técnicos y operacionales, nuestros clientes no solo modernizaron su interacción con los consumidores, sino que también establecieron un estándar de innovación que puede servir como referencia para toda la industria.

Además, al integrar el chatbot a APIs de sistemas internos —como consultas de saldo, estado de procesos o datos de registro— la solución ya incorpora un agente operacional dentro de un ecosistema de IA Agente, capaz de interactuar de forma autónoma con diferentes sistemas para ejecutar tareas y entregar respuestas aún más personalizadas y resolutivas.

Este caso ilustra cómo la GenAI actúa como un habilitador técnico en las soluciones ; viabilizando una interfaz que va más allá de la simple generación de respuestas y se posiciona como un motor estratégico de soluciones inteligentes. En un escenario donde los datos, sistemas e interacciones necesitan ser orquestados con inteligencia, agilidad y seguridad, iniciativas como esta demuestran el potencial de la inteligencia artificial para transformar operaciones, mejorar la experiencia del cliente y generar ganancias tangibles de eficiencia y rendimiento de las aplicaciones para el sector financiero.

Compilado y adaptado a partir de demostraciones funcionales y estudios de MVP aplicados en bancos digitales y tradicionales, centrados en chatbots, API conversacionales y asistentes cognitivos.

Eficiencia operacional: optimizando procesos de Middle y Backoffice con GenAI

El sector financiero, caracterizado por su complejidad operacional y alta regulación, enfrenta desafíos significativos en la gestión de datos no estructurados. Archivos de transacciones bancarias, documentos de compliance e informes de due diligence representan una cantidad masiva de información que demanda una enorme cantidad de tiempo para ser analizada manualmente, y que, cuando no se gestiona adecuadamente, puede limitar la eficiencia organizacional en procesos automatizados. Para superar esta barrera, se desarrolló una nueva solución basada en inteligencia artificial generativa (GenAI), enfocada en transformar la manera en que los datos son procesados y utilizados en el middle y backoffice.

DESAFÍOS Y LIMITACIONES

El proyecto enfrentó desafíos específicos al lidiar con datos no estructurados, que pueden ser divididos en dos factores principales:

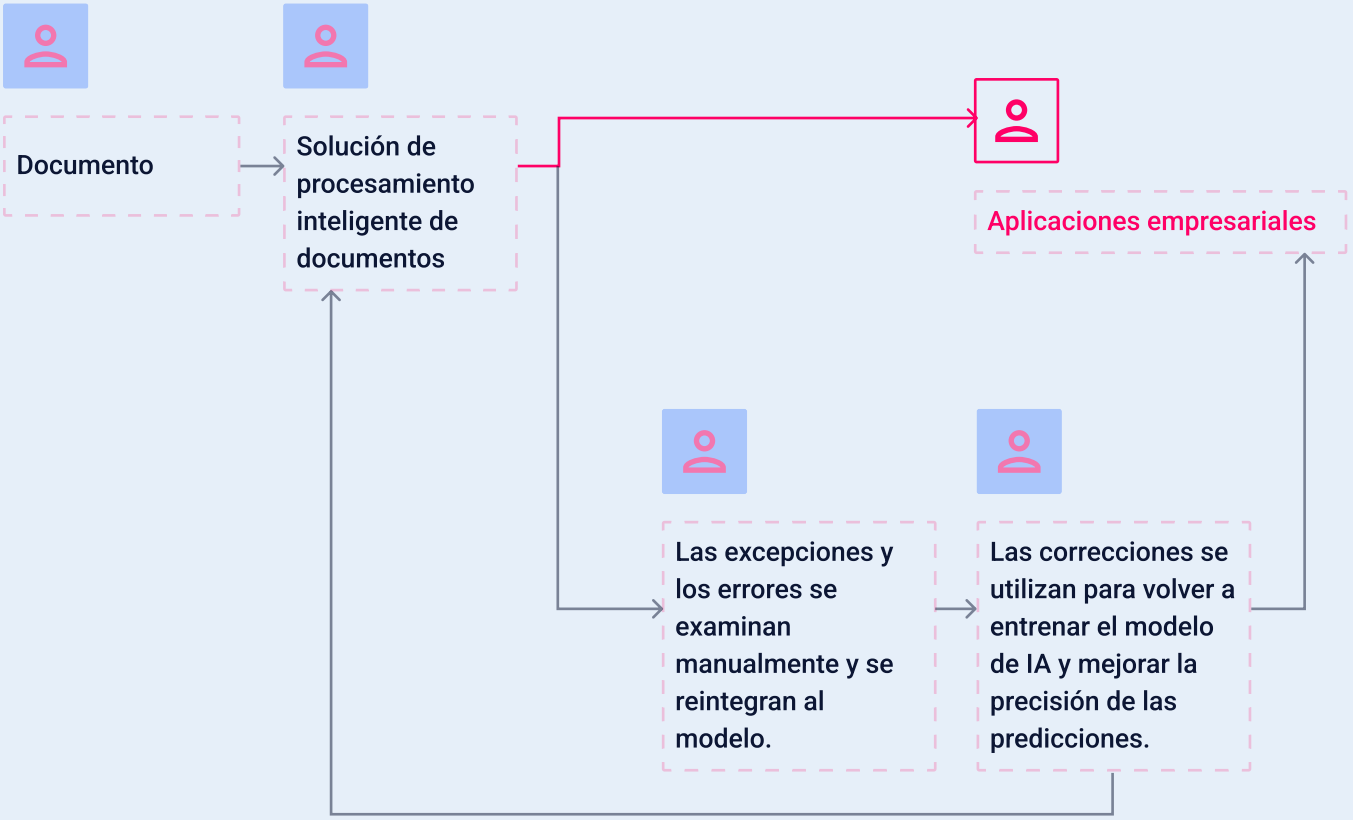
1. La propensión de los modelos de IA generativa a alucinaciones –respuestas que se alejan de la realidad de los datos– exigiendo el desarrollo de mecanismos rigurosos para garantizar la calidad de la información extraída.

2. El alineamiento de la solución a las regulaciones y políticas de compliance específicas del sector financiero, principalmente en este campo donde se tratan datos extremadamente sensibles, garantizando la seguridad y la conformidad en todas las etapas del procesamiento.

Solución Desarrollada

El enfoque combinó la capacidad de la GenAI para interpretar datos no estructurados con técnicas avanzadas de procesamiento y extracción de información. La solución se centralizó en un modelo entrenado para transformar datos no estructurados en formatos estructurados y dirigirlos a tratamientos que tengan en cuenta los objetivos específicos de cada área.

El impacto práctico de esta aplicación fue significativo. Las empresas pudieron obtener matrices de riesgo con base en documentos de compliance, resúmenes consolidados de transacciones bancarias y análisis detallados de due diligence. Además, la solución categorizó las quejas de los clientes con precisión, proporcionando recomendaciones automatizadas para la resolución de problemas en chats y llamadas, reduciendo así la necesidad de interacción humana.



Contenido adaptado a partir de materiales técnicos y resúmenes ejecutivos sobre el uso de la IA aplicada al procesamiento inteligente de documentos en el sector financiero.

Beneficios Obtenidos

Los resultados del proyecto incluyen:

- Aumento sustancial en la **productividad** debido a la facilidad de análisis y síntesis de grandes volúmenes de documentos;
- **Reducción de 1/3 en las llamadas al call center,** con tratamiento automatizado de quejas y categorización de problemas;
- Estructuración de matrices de riesgo altamente detalladas, mejorando la **gestión de compliance;**
- **90% de precisión en el análisis de transacciones bancarias,** elevando el nivel de confianza y eficiencia operacional.

Estas soluciones no solo resolvieron los desafíos inmediatos, sino que también trajeron ganancias de eficiencia para diversos sectores, como la banca abierta, el capital privado (PE) y las aseguradoras. La adaptabilidad del modelo permitió su aplicación en diferentes contextos empresariales, estableciendo un nuevo estándar para el uso de la GenAI en la gestión de datos críticos.

Reflexiones

Este proyecto demuestra cómo las tecnologías de IA generativa pueden transformar los procesos internos en el sector financiero, ofreciendo análisis más rápidos, precisos y alineados con las exigencias regulatorias. Con la automatización inteligente de tareas antes manuales y complejas, las empresas lograron optimizar sus operaciones y redirigir esfuerzos a áreas estratégicas, reafirmando el papel de la GenAI como un motor de innovación y eficiencia.

Considerando la arquitectura ya implementada, si la solución pasara a no solo interpretar los documentos recibidos, sino también a buscar de forma autónoma nuevos datos en sistemas internos o bases externas a través de APIs —como consultas automáticas para actualizar informes de compliance o validar información de transacciones bancarias— ya incorporaría las características y el potencial relacionado con un agente. Así, el modelo evolucionaría de una solución de análisis pasiva a un agente dinámico, que no solo procesa datos, sino que también actúa continuamente para enriquecer y validar sus análisis en tiempo real.

Adaptado a partir de iniciativas prácticas y estudios de casos sobre la aplicación de GenAI y NLP en la automatización de documentos, el análisis de riesgos y la atención al cliente en el sector financiero.

CAPÍTULO 5

Principales Desafíos en el Sector Financiero y cómo superarlos

Principales Desafíos en el Sector Financiero y cómo superarlos

Antes de mostrar los principales desafíos en la implementación de GenAI y la IA Agente en el sector financiero, es importante resaltar algunas convicciones obtenidas a partir de la implementación de casos de uso reales y que son determinantes para superar los principales desafíos y garantizar el éxito de la aplicación de GenAI y la IA Agente en la práctica.

Convicciones de Artefact sobre el uso de GenAI e IA Agente en el Sector Financiero [3]

Una adopción exitosa de GenAI y la IA Agente requiere una planificación cuidadosa y la superación de desafíos significativos, como la integración de sistemas legados y la gestión eficaz de los datos en un ambiente regulado. Para guiar en el proceso de adopción de GenAI, Artefact posee **siete grandes convicciones** para industrias del sector de servicios financieros:

Pruebas de concepto rápidas con visión estratégica a largo plazo

La calidad de los datos como base para el éxito

Propiedad intelectual como diferencial competitivo

Compliance y seguridad como pilar fundamental

Framework de orquestación para integración fluida

Gestión de cambios para garantizar la adhesión organizacional

Retroalimentación continua como motor de evolución

Pruebas de concepto rápidas con visión estratégica a largo plazo

La creación de pruebas de concepto (proof-of-concepts) rápidas es esencial para ganar tracción inicial, pero debe estar integrada en una visión estratégica a largo plazo. Este alineamiento garantiza que la ejecución inmediata no se limite a iniciativas aisladas y permita que las empresas capturen valor sostenible de la GenAI a lo largo del tiempo.

La calidad de los datos como base para el éxito

El verdadero valor de la GenAI se desbloquea cuando la tecnología es alimentada con datos de alta calidad, preferentemente de primera mano. Estos datos proporcionan una base sólida para la generación de conocimientos confiables y resultados precisos, fundamentales para atender las demandas de un mercado exigente

Propiedad intelectual como diferencial competitivo

La creación de propiedad intelectual propietaria es crucial para garantizar la escalabilidad futura. Esto incluye el desarrollo de ingeniería de prompt avanzada y el registro sistemático de salidas para construir una base de conocimiento que pueda ser usada en el refinamiento y la mejora continuos de la tecnología.

Compliance y seguridad como pilar fundamental

La industria de servicios financieros opera en un ambiente altamente regulado, y la adopción de GenAI debe garantizar la conformidad con las regulaciones locales e internacionales. Además de gestionar riesgos relacionados con la seguridad de los datos y la propiedad intelectual, es esencial que las soluciones de GenAI sean diseñadas para atender las exigencias regulatorias, preservando la confianza de las partes interesadas (stakeholders).

Framework de orquestación para integración fluida

Además de la preparación técnica, un framework de orquestación es indispensable para integrar la GenAI de manera coherente en los flujos de trabajo y sistemas existentes. Esta estructura garantiza que las soluciones de GenAI sean aplicadas de forma eficaz, sin interrumpir procesos establecidos.

Gestión de cambios para garantizar la adhesión organizacional

La implementación de la GenAI exige más que tecnología; demanda una transformación cultural y organizacional. La gestión de cambios es crucial para involucrar a las partes interesadas, capacitar a los equipos y asegurar que la adopción de la tecnología sea integrada a los procesos de forma eficaz. Sin un plan estructurado de change management, el riesgo de resistencia y baja adhesión aumenta significativamente.

Retroalimentación continua como motor de evolución

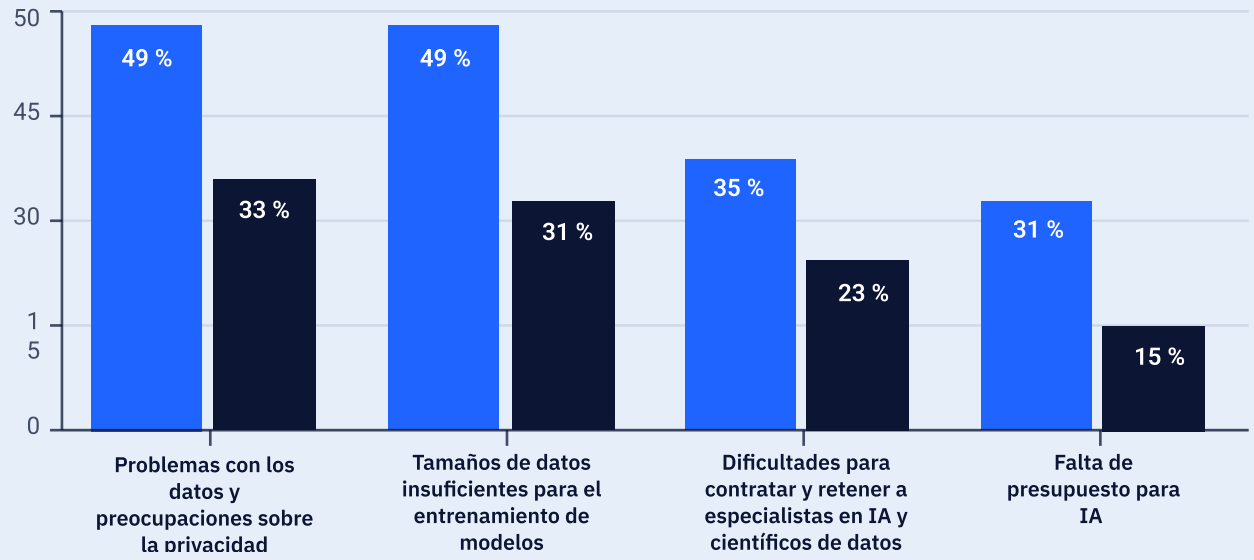
Los mecanismos de retroalimentación continua, con participación humana en el proceso, son esenciales para garantizar mejoras constantes, alineamiento con los objetivos de negocio y la creación de propiedad intelectual. Este ciclo de mejora es un factor determinante para el éxito a largo plazo.

Con estas convicciones, Artefact refuerza su capacidad para guiar a las empresas del sector financiero en una adopción estratégica y eficiente de la GenAI, maximizando el impacto positivo de la tecnología mientras se superan los desafíos asociados a su implementación.

Principales desafíos para la Implementación de IA en el Sector Financiero

> *Los desafíos tradicionales para la implementación de IA en el sector financiero, como las preocupaciones con los datos, las limitaciones de presupuesto y la dificultad para reclutar talentos, vienen presentando una tendencia a la baja.*

Según el informe State of AI in Financial Services: 2025 Trends de NVIDIA, cuestiones como la privacidad de los datos y el tamaño insuficiente de las bases para el entrenamiento, que antes eran barreras expresivas, registraron una reducción significativa de 2023 a 2024. Este movimiento indica una madurez creciente en el mercado: las organizaciones están avanzando de la etapa de experimentación a la consolidación de proyectos a gran escala, reflejando un ambiente más preparado para capturar el valor estratégico de la inteligencia artificial. [2]



[2] *Fuente: El estado de la IA en los servicios financieros: tendencias para 2025 - NVIDIA*

A pesar de la reducción de los obstáculos iniciales, en el sector financiero la atención a los riesgos sigue siendo fundamental. La naturaleza altamente regulada y sensible de la industria exige que las iniciativas de IA sean diseñadas con un rigor adicional para garantizar la confianza, el compliance y la estabilidad del mercado.

La siguiente imagen refuerza esta necesidad, mostrando que la implementación de la IA y la GenAI va mucho más allá de la tecnología —involucra la gestión activa de riesgos estratégicos. Navegar con éxito por estos desafíos exige estructuras de gobernanza robustas, sistemas de IA transparentes y una fuerza de trabajo calificada, capaz de usar y supervisar efectivamente estas tecnologías.

Proteger el valor de la IA/GenAI al tiempo que se aborda el imperativo del riesgo

En los servicios financieros, el riesgo de la IA no es solo una cuestión de tecnología, sino también de confianza, cumplimiento normativo y estabilidad del mercado.

Caso de uso de IA	Tipo de riesgo Inherente	Descripción del Riesgo	Control existente	Puntuación de Riesgo Residual
Detección de fraudes	Sesgo y Justicia	Riesgo de generar predicciones de fraude sesgadas debido a conjuntos de datos de entrenamiento desequilibrados.	Conjuntos de entrenamiento diversos y representativos con auditorías periódicas.	BAJO
Aprobación del préstamo	Transparencia y explicabilidad	La falta de claridad en las decisiones de aprobación de créditos puede provocar la insatisfacción del cliente o el escrutinio regulatorio.	Herramientas de explicabilidad integradas (por ejemplo, valores SHAP); supervisión humana.	MEDIO
Generador de resúmenes de informes financieros GenAI	Alucinación	Riesgo de generar informes financieros incorrectos o engañosos, resúmenes o conclusiones.	Verificación de datos integrada en el flujo de salida; modelo ajustado con datos financieros verificados; revisión humana para informes críticos.	ALTO

El framework presentado organiza los principales riesgos en cuatro grandes categorías: Regulatorio y Compliance, Mercado y Crédito, Reputación y Operacional. Dentro de estas categorías, se mapean riesgos específicos como la gobernanza de modelos, el riesgo de decisiones algorítmicas, la calidad de los datos y la integración de sistemas. Cada tipo de riesgo es ejemplificado con casos de uso: en la detección de fraudes, el principal desafío es mitigar los sesgos de datos; en la aprobación de crédito, garantizar la transparencia y la explicabilidad; y en la generación de informes financieros con GenAI, controlar las alucinaciones —riesgo que, incluso con medidas de mitigación, todavía presenta una alta criticidad residual.

> **Este panorama refuerza que, incluso en un ambiente de maduración tecnológica, el éxito de la IA en el sector financiero depende directamente de la capacidad de las organizaciones para proteger el valor generado, gestionando los riesgos de forma estructurada y continua.**

Además de los riesgos inherentes a los modelos, la necesidad de lidiar con información sensible y atender a las exigencias de los organismos reguladores hace que la implementación de la IA sea aún más compleja. Sectores como la prevención del fraude y el análisis de crédito exigen altos niveles de precisión, transparencia y compliance, ya que los errores pueden generar impactos financieros y reputacionales relevantes. [2]

A pesar de los avances tecnológicos y de la madurez creciente del mercado, la implementación de la IA Generativa (GenAI) en el sector financiero todavía enfrenta desafíos considerables, especialmente en relación a la escalabilidad. Solo cerca del 10% de las instituciones financieras logran escalar sus iniciativas de IA con éxito. Aunque herramientas como ChatGPT o Copilot facilitan la creación de pruebas de concepto (POCs), el gran obstáculo está en la transformación de estas soluciones en productos que generen un impacto real. Para que la GenAI tenga un efecto significativo, necesita ser capaz de responder a cuestiones estratégicas de negocio de manera robusta y escalable, yendo más allá de la simple integración con procesos existentes. Sin esta estructuración, medir el retorno sobre la inversión, reducir costos y aumentar la satisfacción del cliente se convierte en un desafío complejo.

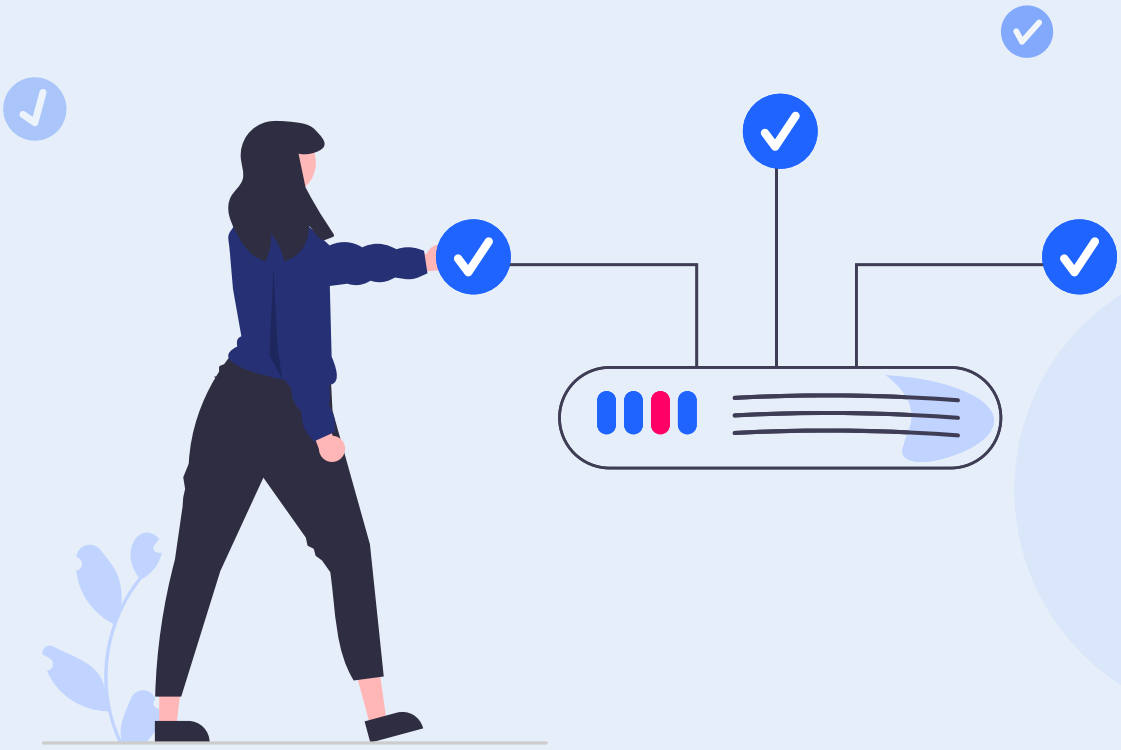
[7] Fuente: AI in Financial Services: Key Market Trends and Insights for 2024 - Joffrey Martinez, Global Financial Services lead, Artefact - AI For Finance Event 2024;Powered by Artefact]

Buenas prácticas para superar desafíos y garantizar una implementación exitosa

Para garantizar una implementación exitosa de la IA en el sector financiero, las instituciones deben adoptar un enfoque estratégico que enfrente los principales desafíos de forma integrada

> La construcción de una gobernanza sólida, con políticas claras y mecanismos de supervisión adecuados, es indispensable para gestionar riesgos y asegurar la conformidad regulatoria.

Además, invertir en el desarrollo de talentos especializados en IA puede mitigar la escasez de profesionales calificados, facilitando la adopción de esta tecnología de manera eficaz. La mejora en la calidad de los datos también es crucial: las políticas robustas de gobernanza de datos, las prácticas mejoradas de recogida y la protección de la privacidad y la seguridad son factores determinantes para el éxito. Por último, la transparencia en los sistemas de IA aumenta la confianza, garantiza una supervisión eficaz y, junto con prácticas éticas bien definidas, reduce los sesgos y promueve decisiones justas.



La gestión estructurada de riesgos y la definición clara de roles y responsabilidades siguen siendo pilares esenciales para el éxito sostenible de las iniciativas de IA. El framework de gobernanza de IA descrito en el capítulo ofrece una visión práctica y jerárquica, mostrando cómo diferentes niveles de responsabilidad —desde los comités de gobernanza hasta los equipos técnicos— colaboran para garantizar la implementación segura y eficaz de estas tecnologías. Esta estructura destaca la importancia de alinear las competencias multidisciplinares, como las áreas legal, de riesgo e ingeniería, a procesos bien definidos, que abarcan desde las validaciones previas a la producción hasta la supervisión continua en producción. Este enfoque no solo mitiga los riesgos, sino que también promueve la transparencia y el compliance, asegurando que las iniciativas de IA estén alineadas con los objetivos estratégicos y regulatorios del sector financiero.

Gobernar para crecer: definición de funciones y responsabilidades para el éxito sostenible de la IA

Una gobernanza eficaz de la IA debe ser construida sobre pilares multifuncionales que integren el valor del negocio, la escalabilidad técnica y la conformidad ética. Conforme a lo ilustrado en la imagen, tres frentes fundamentales sostienen este enfoque: la gestión de la demanda y la generación de valor, que incluye frameworks claros para el desarrollo de casos de uso, bibliotecas centralizadas de prompts optimizados y métricas para medir la adopción y el ROI ; la escalabilidad y la integración, con un enfoque en la seguridad desde el diseño, la supervisión continua y la gobernanza integrada de datos vectoriales (VDBs) para garantizar la consistencia y la calidad ; y la ética y la reglamentación, que aseguran el alineamiento con legislaciones globales, como el GDPR y la AI Act, y promueven prácticas de mitigación de sesgos para decisiones más justas e inclusivas. Esta estructura integrada es esencial para escalar el uso de la GenAI con eficiencia, seguridad y adhesión a los estándares de gobernanza corporativa y regulatorios.

GOVERNANZA DE LA IA

Gestión de la Demanda & Valor

Marco de Desarrollo

Guía estructurada con buenas prácticas, hitos y etapas de validación para el desarrollo de casos de uso de GenAI.

Biblioteca y Directrices de Prompts

Biblioteca centralizada de prompts optimizados con reglas claras de uso para garantizar la estandarización.

Adopción y Medición del ROI

Promover la adopción mediante la gestión del cambio y medir el éxito con indicadores clave de rendimiento (KPI) de retorno de la inversión y satisfacción del usuario.

Escalabilidad & Integración

Control y Supervisión

Modelos fiables y eficientes con supervisión continua, evaluación de KPI e insights sobre la experiencia del usuario.

Seguro por Diseño

Seguridad de nivel empresarial con autenticación segura y separación de derechos de acceso.

Gobernanza Integrada de Datos para VDB

Las bases de datos vectoriales (VDB) heredan la gobernanza de las fuentes de datos, garantizando la coherencia, sin duplicaciones y con calidad de datos.

Ética & Normativa

Cumplimiento Normativo

Alinear el uso de GenAI con las normativas globales (Ley de IA, RGPD, CCPA) mediante la gestión de riesgos y la trazabilidad.

Mitigación del Sesgo y Promoción de la justicia

Garantizar resultados imparciales detectando sesgos, diseñando para la inclusión y evaluando continuamente la equidad.

Machine Learning: Mejorando la Imparcialidad y la Interpretabilidad

Para superar los desafíos del sesgo algorítmico y de la interpretabilidad del modelo en machine learning, las instituciones financieras deben priorizar algoritmos con reconocimiento de imparcialidad y técnicas de IA explicables (XAI) [8] [9]. La implementación de procedimientos rigurosos de prueba y validación puede ayudar a identificar y mitigar los sesgos en los datos de entrenamiento y en las salidas del modelo [8]. La utilización de técnicas como los valores SHAP y LIME para cuantificar el peso de las características en los resultados de los modelos, puede proporcionar conocimientos sobre los procesos de toma de decisiones de los modelos de machine learning, mejorando la interpretabilidad y la transparencia [8]. Además, el involucramiento con las partes interesadas y los reguladores para establecer directrices claras para la ética y la transparencia de la IA es esencial para la implantación responsable de la IA [8].

IA Generativa: Mitigando Alucinaciones y Garantizando la Seguridad de los Datos

Para enfrentar los desafíos de la GenAI, es necesario enfocarse en la seguridad, la privacidad y la precisión de los datos. La implementación de la Generación Aumentada de Recuperación (RAG) puede mejorar la precisión de la respuesta y el contexto, extrayendo información de fuentes internas de la empresa. Los bancos deben implementar medidas robustas de seguridad de los datos y obtener el consentimiento explícito del cliente para el uso de la IA, con el fin de adherirse a las estrictas regulaciones de privacidad de los datos [8]. Además, el involucramiento activo entre los bancos y los organismos reguladores es necesario para navegar en el escenario regulatorio en evolución y mitigar posibles imprecisiones en las previsiones de la IA, estableciendo estructuras transparentes y eficaces [8].

[8] Datos de: The Alan Turing Institute – The AI Revolution, 2024.

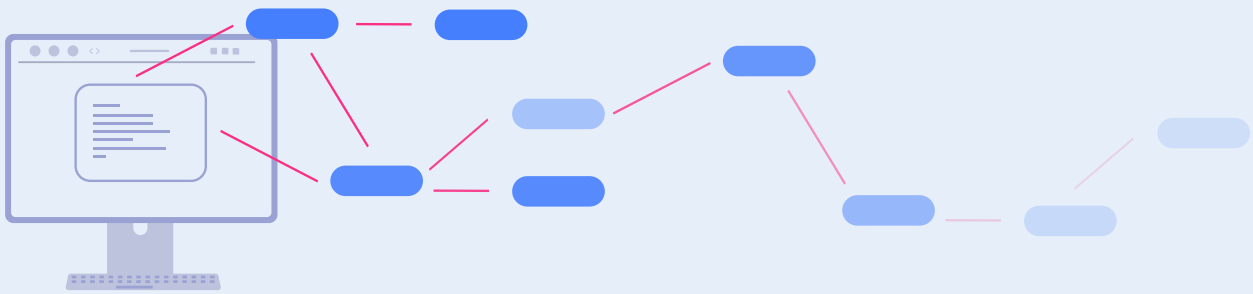
[9] Fuente: WhatNext.Law – Legal challenges of AI in finance.

IA Agente: Fortaleciendo la Gobernanza y los Límites Éticos

Para superar los desafíos de la IA Agente, es preciso fortalecer la gobernanza, los límites éticos y la infraestructura de seguridad. Las instituciones financieras deben establecer directrices éticas claras para la toma de decisiones de la IA, incorporando la supervisión humana y los mecanismos de responsabilización [8] [9]. La implementación de medidas de seguridad robustas y los denominados "límites de protección" es crucial para evitar el uso indebido, como el lavado de dinero o la negociación con información privilegiada [8] El sesgo algorítmico en la IA Agente puede ser abordado a través de la monitorización cuidadosa, estrategias de mitigación y procesos transparentes de toma de decisiones [8].

Por último, la escalabilidad de la IA como un todo involucra cuestiones éticas, ambientales y sociales, además de depender fuertemente de una gestión eficaz de cambios (change management), especialmente debido a la complejidad organizacional de estas instituciones.

La promoción de una cultura de responsabilidad y propiedad sobre los datos es esencial, pero a menudo choca con la falta de alineamiento entre diferentes áreas y la resistencia al cambio. Así, el éxito en la implementación de la IA exige un enfoque holístico que equilibre la innovación tecnológica, la responsabilidad social y ambiental y el fomento y la valorización de una cultura organizacional orientada a los datos.



Basado en AI for Finance Event 2024 – Masterclass por expertos de Artefact.

CAPÍTULO 6

Estrategia en la implementación de IA en el Sector Financiero

Estrategia en la implementación de IA en el Sector Financiero

El éxito en la implementación de casos de uso de IA comienza con una evaluación precisa de la madurez organizacional y del modelo operacional existente. Este diagnóstico es fundamental para identificar las lagunas y los potenciales riesgos que puedan comprometer el proyecto. Con base en este análisis, las instituciones financieras pueden adoptar las siguientes estrategias:

DIAGNÓSTICO Y PLANIFICACIÓN ESTRATÉGICA

Para garantizar una adopción exitosa de cualquier tecnología, es esencial comprender dónde se encuentra hoy la empresa en su jornada de madurez digital. Solo con un diagnóstico preciso es posible estructurar una planificación sólida que dirija a la organización hacia el nivel deseado. Este entendimiento permite que la implementación de inteligencia artificial generativa se haga de manera estratégica, maximizando el valor y minimizando los riesgos.

En Artefact, evaluamos la madurez digital de las empresas a través de tres dimensiones fundamentales:

Estrategia

Los datos son valorizados como un activo esencial para impulsar la estrategia de negocios, garantizando que la toma de decisiones esté orientada por conocimientos robustos.

Entrega

La organización debe ser capaz de transformar los datos en resultados concretos, aplicando inteligencia analítica para optimizar procesos y generar un impacto positivo.

Gestión y Gobernanza de Datos

La gestión adecuada de los datos asegura la calidad, el compliance y la seguridad, factores indispensables para un ambiente altamente regulado como el sector financiero.

Estas dimensiones son evaluadas dentro de 3 aspectos:

Personas

Se evalúa el compromiso del liderazgo con una estrategia de datos bien estructurada y financiada, en la capacitación de los colaboradores para el uso estratégico de los datos y en el patrocinio ejecutivo para la gobernanza y la gestión eficaces, garantizando el alineamiento organizacional y la generación de valor para el negocio.

Procesos

Se evalúa el uso sistemático de los activos de datos para generar valor, mejorar la propuesta al cliente y optimizar operaciones. Además, se analiza la integración de la analítica y los conocimientos confiables en los procesos de negocio, así como la existencia de una gobernanza estructurada, con una definición clara, documentación, ejecución y monitorización de los procesos de gestión de datos, garantizando su confiabilidad.

Tecnología

Se evalúa la existencia de una estrategia técnica clara y un roadmap que define las capacidades necesarias para alcanzar los objetivos de datos de la organización. También considera si la arquitectura tecnológica soporta la entrega de casos de uso prioritarios y si está preparada para el futuro. Además, evalúa la conformidad, la seguridad, la estabilidad y el control de acceso a los datos.

De forma bidimensional podemos ejemplificar las dimensiones y los aspectos de la madurez analizada de acuerdo con la imagen de abajo:

	Estrategia	*Entrega*	*Gestión y Gobernanza de Datos*
	Los datos se valoran como un activo estratégico.	**Los datos impulsan la entrega**	**La gestión y la gobernanza de datos garantizan el cumplimiento normativo y mejoran la calidad de los datos.**
Personas	El liderazgo asume activamente una estrategia de datos ambiciosa y práctica, con el nivel adecuado de financiación y apoyo organizativo.	Todos los miembros de la organización, tanto usuarios empresariales como especialistas en datos, cuentan con los conocimientos necesarios para utilizar los datos y generar resultados empresariales.	La gestión y gobernanza de datos cuentan con patrocinio ejecutivo.
Procesos	Los activos de datos se utilizan sistemáticamente para crear nuevo valor, mejorar la propuesta al cliente e impulsar mejoras operativas.	Los procesos empresariales se basan en análisis e información fiables.	Los procesos de gestión y gobernanza de datos son responsabilidad clara, definida, documentada, entregada y supervisada. Los datos son fiables.
Tecnología	Existe una estrategia técnica clara y una hoja de ruta que define las capacidades necesarias para alcanzar las ambiciones de datos de la organización.	La pila tecnológica permite la entrega de casos de uso prioritarios de forma adecuada y preparada para el futuro.	Los datos son conformes, seguros, estables y con control de acceso.

Con este enfoque estructurado, las instituciones financieras pueden no solo implementar la IA de manera eficiente, sino también garantizar que la tecnología esté alineada con la estrategia de negocios, fortaleciendo la competitividad y la resiliencia organizacional.

Compromiso de las Partes Interesadas

El compromiso de las partes interesadas es un elemento esencial para el éxito en la adopción de la IA en el sector financiero. Involucrar a los líderes de las áreas de negocios, tecnología, compliance y seguridad desde el inicio del proceso permite alinear las expectativas y garantizar el apoyo continuo a lo largo del proyecto. Además, un enfoque colaborativo, que promueva el compromiso y la responsabilidad compartida entre los equipos, es fundamental para integrar diferentes perspectivas y construir una visión unificada de los objetivos estratégicos.

La transformación digital y la adopción de la IA van más allá de la tecnología: exigen una gestión del cambio estructurada dentro de la organización. Esto implica tanto el desarrollo de habilidades sociales, como la adaptación cultural y la mentalidad orientada a los datos, como de habilidades técnicas, como la capacitación necesaria para manejar las nuevas herramientas de inteligencia artificial.

En Artefact, aceleramos este proceso de cambio organizacional para nuestros clientes a través de nuestro marco de trabajo, que cuenta con tres palancas esenciales:

Conocimiento

Desarrollar talentos especializados en datos y análisis avanzados, capacitando a la organización para contratar y retener los perfiles adecuados para impulsar su estrategia de datos.

Formación

Creación de programas de formación para garantizar que los equipos comerciales y los directivos tengan conocimientos y autonomía en el uso de datos, fomentando la toma de decisiones basada en datos.

Aculturación

Promover una base de conocimiento común para todos los colaboradores, estableciendo definiciones estandarizadas para conceptos de datos y KPIs, además de garantizar que toda la organización comprenda el valor generado por el uso inteligente de los datos.

Pilotos e Iteraciones: un enfoque ágil para reducir riesgos y maximizar valor

> Cuando hablamos de soluciones de IA, es esencial reconocer que existen riesgos inherentes, como alucinaciones, sesgos en los modelos y desafíos de interpretabilidad. Para mitigar estos desafíos, la gestión de cambios mencionada anteriormente debe crear un ambiente propicio para la experimentación, permitiendo que los equipos prueben y validen nuevas soluciones de manera controlada.

BUSINESS DISCOVERY

El primer paso en esta jornada es la fase de descubrimiento del negocio (Business Discovery). En esta etapa inicial, es fundamental mapear los posibles casos de uso de la tecnología dentro de la organización, intentando ser lo más exhaustivo posible. Esta etapa debe estar directamente alineada con la visión estratégica de la empresa, garantizando que la adopción de la IA tenga un propósito claro y genere un impacto real.

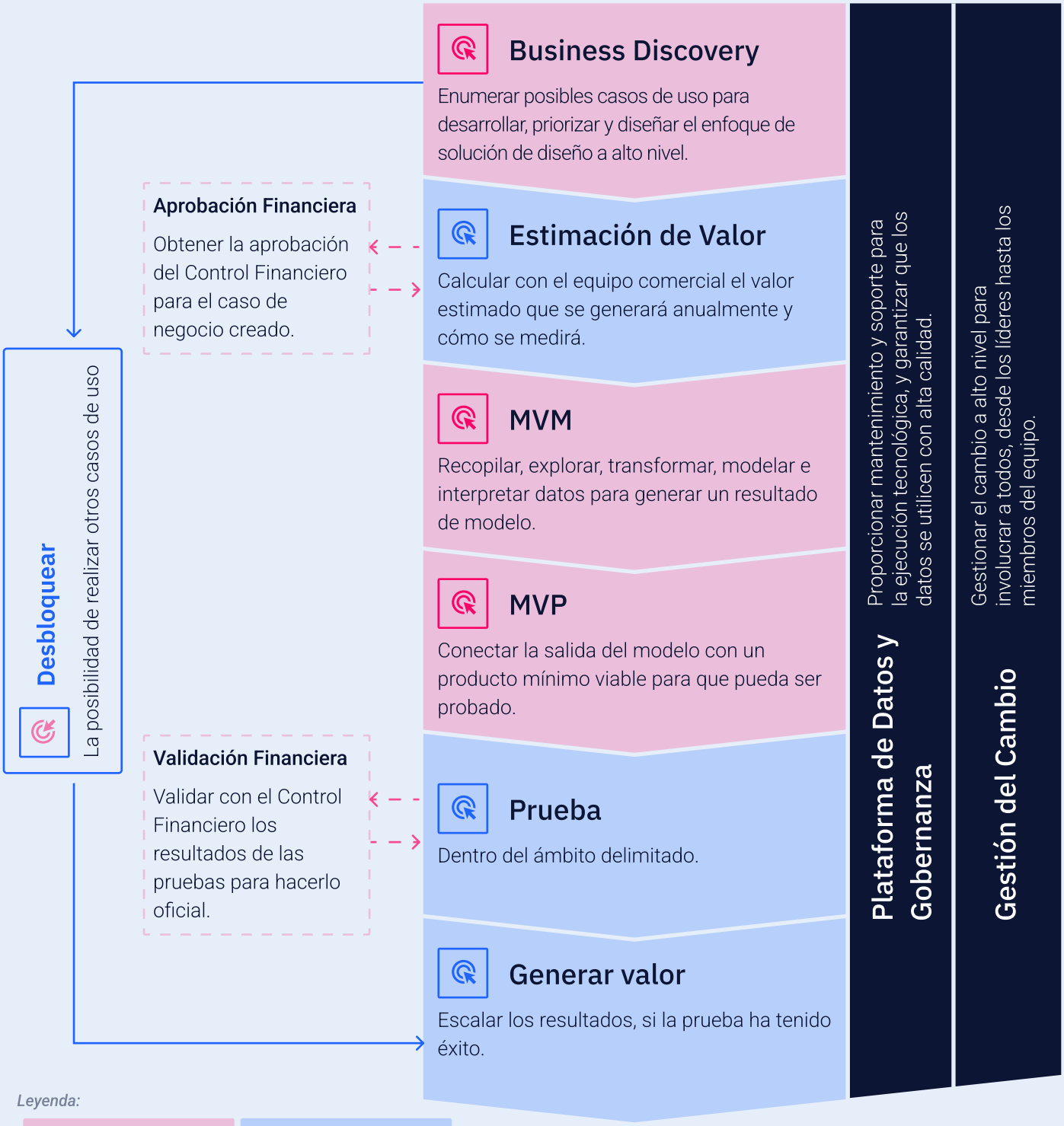
ESTIMACIÓN DEL VALOR

Después del mapeo, es necesario estimar el valor generado por cada Caso de Uso. Esto permite una priorización que garantiza que los esfuerzos estén enfocados en las iniciativas que traerán el mayor retorno en el menor tiempo posible, siempre buscando agregar valor de manera ágil y eficiente. La priorización correcta permite que los primeros proyectos probados sirvan como prueba de concepto para futuras iniciativas, creando un ciclo sostenible de innovación.

Con los casos priorizados, pasamos a la fase de construcción y prueba, siguiendo la lógica de MVM (Minimum Viable Model), MVP (Minimum Viable Product) y experimentación iterativa. Es en este momento que el principio de "empezar pequeño y errar rápido" entra en juego. Esta metodología reduce los costos y minimiza los riesgos, ya que permite ajustes continuos antes de una inversión más robusta en la solución final. A través de múltiples iteraciones, los modelos son refinados para garantizar un alcance bien definido, con menor propensión a errores y mayor confiabilidad.

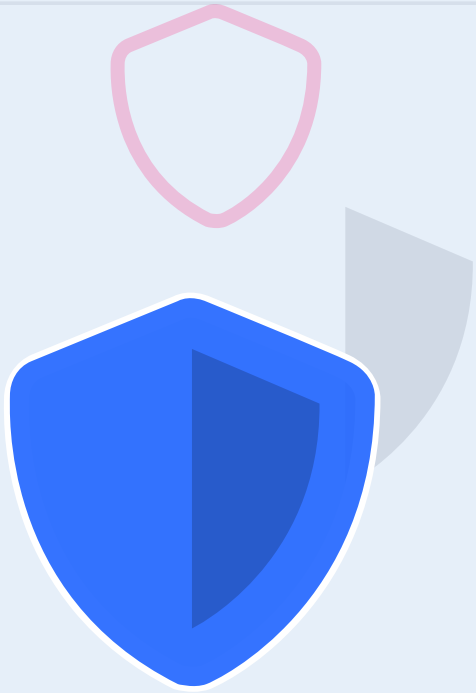
Una vez que la prueba genera resultados positivos y concluyentes, llega el momento de escalar la solución. En esta etapa, la tecnología pasa a ser aplicada de forma más amplia dentro de la organización, maximizando su impacto y generando valor tangible para el negocio. Este ciclo de aprendizaje y validación crea un efecto multiplicador: la ganancia en eficiencia y retorno financiero obtenida con los primeros proyectos permite que se financien nuevas iniciativas, acelerando aún más la transformación digital de la empresa.

Al combinar experimentación controlada, iteración continua y escalabilidad estratégica, las instituciones financieras logran explorar el potencial de la GenAI de forma segura, eficiente y sostenible, garantizando que la innovación traiga resultados concretos y a largo plazo.



Al seguir estas etapas, las instituciones financieras pueden establecer una base sólida para la adopción y creación de casos de uso, creando un ambiente donde la innovación y la seguridad caminan lado a lado. Así, es posible maximizar el valor generado por la tecnología, promoviendo la eficiencia, el compliance y los resultados tangibles para el negocio.

La Seguridad y Privacidad de los Datos como una Prioridad Estratégica



En el sector financiero, la seguridad y la privacidad de los datos son factores innegociables. Los bancos y otras instituciones lidian con información altamente sensible, como historiales financieros, datos personales y transacciones, lo que exige medidas rigurosas de protección. Con el advenimiento de la Inteligencia Artificial (IA), incluyendo sus variantes como el Machine Learning y la IA Generativa, esta necesidad se ha vuelto aún más crítica.

La elección de la infraestructura tecnológica para soportar soluciones de IA es un factor estratégico crucial que impacta directamente en la capacidad de implementar y mantener robustas medidas de seguridad. En este contexto, las instituciones financieras se enfrentan a una elección fundamental: adoptar soluciones basadas en la nube o mantener una infraestructura on-premise.

> **En la implementación de soluciones de IA en el sector financiero, las instituciones enfrentan una decisión crucial entre adoptar infraestructuras basadas en la nube o mantener sistemas on-premise. Cada enfoque ofrece características distintas que pueden impactar significativamente en el éxito y la eficacia de las iniciativas de IA.**

La siguiente tabla presenta una comparación objetiva entre las soluciones basadas en la nube y las soluciones on-premise, destacando los principales aspectos, ventajas y desventajas de cada enfoque. Este análisis busca auxiliar a las instituciones financieras en la elección de la infraestructura más adecuada para soportar sus iniciativas de inteligencia artificial, considerando factores como el costo, la escalabilidad, la seguridad, la conformidad regulatoria y la dependencia tecnológica.

ASPECTO	SOLUCIONES EN LA NUBE	SOLUCIONES LOCALES
Infraestrutura	Gestionado por proveedores externos.	Totalmente gestionada por la institución.
Escalabilidad	Alta, con rápida expansión según la demanda.	Depende de la infraestructura existente.
Costos Iniciales	Bajos, modelo de pago por uso.	Altos, requiere una inversión significativa en hardware y software.
Mantenimiento	Actualizaciones automáticas por parte del proveedor.	Responsabilidad de la institución.
Seguridad y Privacidad	Riesgos de cumplimiento y almacenamiento externo.	Mayor control sobre los datos confidenciales.
Cumplimiento Normativo	Puede ser compleja en relación con la ubicación de los datos.	Facilita el cumplimiento de normativas específicas.
Tecnología	Acceso a marcos y herramientas de vanguardia ofrecidos por proveedores líderes.	Infraestructura adaptada a las necesidades específicas de la institución.
Dependencia	Alta, vinculada a la fiabilidad de los proveedores externos.	Baja, operación independiente de terceros.

En los capítulos siguientes, se explorarán en mayor profundidad los detalles de cada enfoque. Se discutirán las características específicas, los casos de uso más adecuados y las mejores prácticas para la implementación de soluciones en la nube y on-premise, proporcionando una visión integral para apoyar la toma de decisiones estratégicas en el contexto de las instituciones financieras.

Soluciones en la Nube

Se refiere a la utilización de servicios de computación proporcionados por proveedores de terceros a través de Internet. En este modelo, la infraestructura, las plataformas y los softwares son gestionados por el proveedor de servicios en la nube, permitiendo que las instituciones financieras accedan a los recursos de IA bajo demanda, sin la necesidad de mantener hardware físico en sus instalaciones.

VENTAJAS:

Escalabilidad y Flexibilidad: Expansión ágil de la capacidad de procesamiento, ajustándose a las demandas del negocio.

Acceso a Tecnología de Punta: Proveedores como AWS, Google Cloud y Azure ofrecen plataformas optimizadas para la GenAI, con frameworks y herramientas líderes del mercado.

Modelo de Pago por Uso: Reduce los costos iniciales, permitiendo que las instituciones inviertan proporcionalmente al uso.

Actualizaciones Automáticas: Garantía de acceso a las versiones más recientes de seguridad y tecnología, sin necesidad de intervenciones internas.

DESVENTAJAS:

Preocupaciones con la Seguridad y la Privacidad: Los datos sensibles almacenados fuera de la infraestructura interna pueden generar desafíos de compliance y riesgos de seguridad.

Dependencia de Proveedores Externos: La operación queda vinculada a la confiabilidad y la continuidad de los servicios ofrecidos por terceros.

Conformidad Regulatoria: Algunas leyes exigen que los datos permanezcan en el país o sean tratados con controles específicos, lo que puede ser complejo de atender en ciertas plataformas de la nube.

Soluciones On-Premise

Involucra la implementación y el mantenimiento de toda la infraestructura de IA dentro de las instalaciones físicas de la propia institución financiera. En este escenario, la organización es responsable de todos los aspectos de la infraestructura, incluyendo el hardware, el software, la seguridad y el mantenimiento, ofreciendo un control total sobre los sistemas y los datos.

VENTAJAS:

Autonomía y Control Absoluto: Las operaciones permanecen enteramente dentro de la infraestructura de la institución, garantizando una mayor protección de los datos sensibles.

Facilidad de Conformidad: La ubicación de los datos dentro del ambiente interno simplifica el cumplimiento de las regulaciones específicas.

Personalización: Infraestructura adaptada a las necesidades específicas de la institución.

DESVENTAJAS:

Altos Costos Iniciales: Inversiones significativas en hardware, software y equipos especializados.

Escalabilidad Limitada: Las expansiones pueden ser lentas y costosas, dependiendo de la infraestructura existente.

Mantenimiento y Soporte: Todas las actualizaciones y medidas de seguridad dependen de la capacidad interna de la organización.

Factores Críticos para la Decisión

Podemos resumir la toma de decisiones informada en una serie de siete factores críticos:

01 **Conformidad Regulatoria:** Evaluar las restricciones legales sobre la ubicación y el procesamiento de datos.

02 **Escalabilidad:** Considerar el crecimiento futuro y la capacidad de adaptarse rápidamente.

03 **Seguridad:** Analizar los controles necesarios para proteger los datos sensibles.

04 **Costos:** Comparar los costos iniciales y operacionales a largo plazo.

05 **Conocimiento Interno:** Evaluar la capacidad del equipo de TI para gestionar la infraestructura de IA.

06 **Integración:** Considerar la compatibilidad con los sistemas existentes.

07 **Rendimiento:** Evaluar los requisitos de latencia y rendimiento para aplicaciones críticas.

Tendencias y Recomendaciones

Considerando las tendencias actuales del sector financiero y los beneficios ofrecidos por la computación en la nube, muchas instituciones se están inclinando por soluciones basadas en la nube o adoptando un enfoque híbrido. La nube ofrece ventajas significativas en términos de agilidad, escalabilidad y acceso a tecnologías de punta, cruciales para mantener la competitividad en el escenario de rápida evolución de la IA.

Sin embargo, para instituciones con requisitos excepcionales de seguridad, regulaciones muy específicas o necesidad de un control total sobre la infraestructura, una solución on-premise o híbrida puede ser más apropiada. El enfoque híbrido, en particular, está ganando popularidad como un paso intermedio, permitiendo que las instituciones aprovechen los beneficios de la nube mientras mantienen ciertos sistemas críticos on-premise.

Cuidados Especiales y Buenas Prácticas

La tabla a continuación presenta cuidados especiales y buenas prácticas que deben ser observados en la implantación de soluciones de inteligencia artificial tanto en ambientes de nube como en infraestructuras on-premise. Estos aspectos son fundamentales para garantizar la seguridad, la conformidad y la eficiencia operacional, considerando las particularidades de cada enfoque tecnológico.

DIMENSIÓN	SOLUCIONES EN LA NUBE (CLOUD)	SOLUCIONES ON-PREMISE
Protección de Datos	Cifrado de principio a fin (datos en tránsito y en reposo)	Seguridad física de los servidores e infraestructura
Gestión de Acceso	Autenticación multifactor y principio del menor privilegio	Políticas de control de acceso físico y lógico
Auditoría y Monitorización	Auditorías regulares y herramientas de monitorización en tiempo real	Sistemas de detección de intrusión y monitorización interna
Resiliencia y Copia de Seguridad	Estrategias robustas de redundancia y recuperación de desastres	Planes de continuidad de negocios probados regularmente
Actualizaciones y Vulnerabilidades	Dependencia de SLAs y actualizaciones del proveedor de la nube	Actualizaciones frecuentes y escaneos de vulnerabilidades realizados internamente
Segmentación y Aislamiento	Segmentación lógica a través de políticas de red e identidad	Segmentación física y lógica de la red para la protección de sistemas críticos
Compliance Regulatorio	Garantía contractual con los proveedores para atender las regulaciones financieras	Conformidad directamente gestionada y auditada por el equipo interno
Capacitación del Equipo	Entrenamiento continuo en prácticas seguras para ambientes cloud	Capacitación para la administración segura de la infraestructura y la respuesta a incidentes locales

Para Soluciones en la Nube:

Cifrado de Datos: Utilice cifrado de principio a fin para los datos en tránsito y en reposo.

Gestión de Acceso: Implemente controles de acceso rigurosos, incluyendo autenticación multifactor y el principio del menor privilegio.

Auditoría Regular: Realice auditorías frecuentes de seguridad y conformidad.

Monitorización Continua: Utilice herramientas de monitorización en tiempo real para detectar actividades sospechosas.

Redundancia y Copia de Seguridad: Implemente estrategias de copia de seguridad y recuperación de desastres robustas.

Conformidad Contractual: Asegure que los contratos con los proveedores de la nube atiendan a los requisitos regulatorios del sector financiero.

Entrenamiento del Equipo: Capacite al equipo en prácticas de seguridad específicas para ambientes en la nube.

Para Soluciones On-Premise:

Seguridad Física: Implemente medidas rigurosas de seguridad física para proteger los servidores y la infraestructura.

Actualizaciones Regulares: Mantenga los sistemas y el software actualizados con los últimos parches de seguridad.

Segmentación de Red: Utilice la segmentación de red para aislar los sistemas críticos.

Monitorización Interna: Implemente sistemas de detección de intrusión y monitorización de actividades internas.

Plan de Continuidad: Desarrolle y pruebe regularmente los planes de continuidad de negocios.

Gestión de Vulnerabilidades: Realice escaneos regulares de vulnerabilidades y corrija rápidamente cualquier problema identificado.

Control de Acceso Físico y Lógico: Implemente políticas estrictas de control de acceso tanto para el acceso físico como para el lógico a los sistemas.

CONSULTE EL APÉNDICE

Framework de Implementación Técnica (p. 89)

De la estrategia a la arquitectura: ¡tu viaje hacia la IA comienza aquí!

Amplíe su visión y prepare su organización para la adopción práctica y segura de Agentic AI y GenAI en entornos locales. Hemos creado un marco técnico completo, con recomendaciones detalladas para transformar la intención en acción con responsabilidad, escalabilidad y gobernanza.

CONTENIDO TÉCNICO

Framework para la Implementación de IA Agente y GenAI on-premise

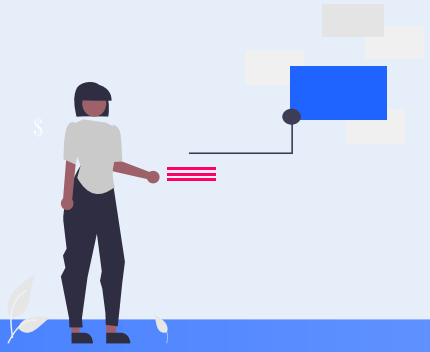
CAPÍTULO 7

Framework Estratégico para la Implementación de Agentic AI por área de negocio

Introducción: De la Visión a la Acción Estratégica

> *Los capítulos anteriores demostraron el potencial transformador de la Inteligencia Artificial (IA), culminando en la emergencia de la IA Agente – sistemas capaces no solo de analizar y generar conocimientos, sino de actuar autónomamente para ejecutar tareas complejas y alcanzar objetivos definidos. Mientras que la IA Generativa (GenAI) revolucionó la creación de contenido y la interacción, la IA Agente representa el siguiente salto: la automatización cognitiva y la ejecución proactiva.*

Sin embargo, liberar el verdadero potencial de la IA Agente en el diversificado escenario financiero exige más que una adopción tecnológica; demanda una estrategia dirigida y específica para cada área de negocio. Ya sea en la Banca Minorista, Banca de Inversión, Seguros o Pagos (conforme a lo detallado en la Tabla de la página 34), los desafíos, las oportunidades y los imperativos regulatorios varían significativamente.



Este capítulo proporciona un framework estratégico para guiar a las instituciones financieras en la construcción de un roadmap de implementación de IA Agente a la medida para sus unidades de negocio específicas. El objetivo es capacitar a los líderes para que hagan las preguntas correctas, consideren las dimensiones críticas y estructuren un enfoque que maximice el valor y mitigue los riesgos inherentes a esta poderosa tecnología, inspirándose en los ejemplos concretos ya posibles.

El Imperativo Estratégico: Desbloqueando el Potencial de la IA Agente en los Cuatro Frentes Clave

La IA Agente representa un cambio fundamental, capacitando a los sistemas no solo para analizar, sino para actuar de forma autónoma y coordinada para alcanzar objetivos complejos. El imperativo estratégico para el sector financiero reside en explorar cómo esta capacidad de acción transforma radicalmente los cuatro frentes clave de valor, conforme a lo categorizado (y ejemplificado inicialmente) en la Tabla de la página 34. El potencial va mucho más allá de los ejemplos actuales:

01 Automatización de Procesos – Hacia la Orquestación Inteligente y Adaptativa:

La IA Agente eleva la automatización de la simple ejecución de tareas a la orquestación dinámica de flujos de trabajo de principio a fin. El horizonte futuro involucra a agentes que no solo siguen procesos (como la concesión de crédito en la Banca Minorista o la gestión de siniestros en Seguros), sino que los optimizan en tiempo real, aprenden de las excepciones y negocian autónomamente dentro de parámetros seguros. Imagine a los agentes gestionando interacciones complejas entre instituciones en Pagos o automatizando la resolución de excepciones en ciclos de trade en la Banca de Inversión, adaptando el proceso dinámicamente a las condiciones e informaciones emergentes.

02 Gestión de Riesgos – De la Respuesta a la Prevención y Adaptación Autónoma:

El verdadero salto estratégico en la gestión de riesgos con la IA Agente es moverse de la detección y la respuesta en tiempo real a la prevención predictiva y la adaptación autónoma de las defensas. En el futuro, los agentes en la Banca Minorista podrán anticipar los riesgos de impago e iniciar planes de mitigación antes incluso de que el problema se manifieste. En Seguros, los agentes podrán modelar los riesgos emergentes (como los climáticos) y activar respuestas preventivas. En la Banca de Inversión, los agentes podrán ajustar autónomamente los límites y los colaterales con base en señales predictivas complejas. En Pagos, la capacidad de aprender y adaptar las defensas continuamente permitirá el desmantelamiento autónomo de redes de fraude cada vez más sofisticadas.

03 Atención, Compromiso y Personalización: el Agente como Navegador Financiero

Proactivo:

La IA Agente posibilita un nivel de compromiso verdaderamente proactivo y holístico, mucho más allá de las recomendaciones personalizadas actuales. El futuro contempla a los agentes en la Banca Minorista gestionando el ciclo de vida financiero completo del cliente, anticipando necesidades y optimizando productos de forma integrada. En Seguros, los agentes podrán usar datos (con consentimiento) para sugerir proactivamente ajustes de cobertura o acciones preventivas. En la Banca de Inversión (Wealth Management), los agentes podrán ejecutar autónomamente optimizaciones de portafolio alineadas con los objetivos a largo plazo del cliente. En Pagos, los agentes podrán actuar como gestores financieros autónomos, optimizando gastos y ahorros proactivamente.

04 Toma de Decisiones e Insights Basados en IA – Del Análisis a la Iniciativa Estratégica

Autónoma:

La frontera más avanzada de la IA Agente reside en la autonomía de toma de decisiones en dominios estratégicos complejos. Esto incluye a los agentes en la Banca de Inversión capaces de desarrollar, probar y ejecutar nuevas estrategias de trading. En la Banca Minorista y en Seguros, los agentes podrán realizar la optimización dinámica de precios y la suscripción, respondiendo autónomamente a las condiciones del mercado. En Pagos, los agentes podrán analizar vastos conjuntos de datos para identificar e incluso iniciar la exploración de nuevos modelos de negocio o asociaciones estratégicas, transformando los conocimientos directamente en una acción estratégica.

Por lo tanto, al definir la estrategia de IA Agente para su área de negocio, es crucial mirar a los ejemplos actuales (como los de la Tabla página 34) como inspiración y validación, pero apuntar al vasto potencial aún inexplorado. La cuestión no es solo "¿Cómo podemos automatizar lo que hacemos hoy?", sino "¿Qué nuevas capacidades de acción, optimización y creación de valor la autonomía de los agentes nos permite alcanzar mañana?". Las instituciones que respondan a esta pregunta de forma visionaria y estratégica liderarán la transformación del sector financiero.

Dimensiones Clave de la Estrategia de IA Agente

Para construir un roadmap robusto, la estrategia de implementación de IA Agente para un área de negocio específica debe abordar las siguientes dimensiones interconectadas:

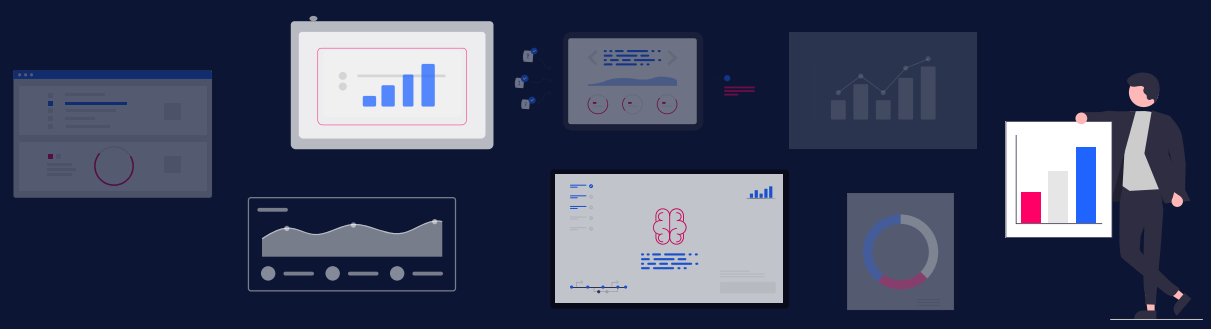
01 **Identificación y Priorización de Casos de Uso:** Mapear los procesos y las oportunidades donde la autonomía y la capacidad de acción de los agentes traerán un mayor impacto, inspirándose en los ejemplos existentes y apuntando al potencial futuro.

02 **Preparación Tecnológica y de Datos:** Evaluar y planificar la infraestructura, los datos, las APIs y las herramientas necesarias para soportar las operaciones de los agentes.

03 **Gobernanza, Riesgo y Conformidad (GRC):** Establecer los límites éticos, regulatorios y de seguridad para la operación autónoma de los agentes.

04 **Capacitación Humana y Gestión de Cambios:** Preparar a los equipos para colaborar, supervisar y confiar en los nuevos sistemas de agentes.

05 **Mensuração de Desempenho e Melhoria Contínua:** Definir métricas para acompanhar o impacto e estabelecer ciclos de feedback para a evolução dos agentes e da estratégia.



Preguntas Estratégicas Esenciales por Dimensión

Para cada dimensión, los líderes del área de negocio deben buscar respuestas claras y específicas:

01 Alineamiento Estratégico y Propuesta de Valor:

¿Cuáles son los 2-3 principales objetivos estratégicos de nuestra área de negocio (ejemplo: reducir los costos operacionales en X%, aumentar la retención de clientes en Y%, lanzar Z nuevos productos) para los próximos 1-3 años?

¿Cómo la autonomía y la capacidad de acción de la IA Agente (yendo más allá del análisis de la GenAI) pueden acelerar directamente el alcance de estos objetivos, considerando los cuatro frentes clave (Automatización, Riesgo, Atención, Decisión)?

¿Cuál es la propuesta de valor única de la IA Agente para nuestros clientes y operaciones internas en esta área específica? ¿Qué problema resuelve mejor que las soluciones actuales, especialmente pensando en el potencial "Más Allá del Horizonte"?

¿Cómo encaja la IA Agente en nuestra estrategia digital y de datos más amplia?

02 2. Identificación y Priorización de Casos de Uso:

¿Qué procesos de principio a fin en nuestra área (que involucran múltiples etapas, sistemas y decisiones) son más adecuados para ser orquestados por agentes autónomos? **(Consulte la Tabla de la página 34 para ejemplos de partida y categorías como Automatización de Procesos, Gestión de Riesgos, Atención/ Personalización y Toma de Decisiones/Insights).**

Dentro de nuestra área de negocio específica (Banca Minorista, Banca de Inversión, Seguros o Pagos), ¿cuáles de los ejemplos "Más Allá del Horizonte" descritos en la sección 8.2 representan las mayores oportunidades transformacionales o resuelven los problemas más profundos?

¿Dónde la capacidad de un agente para interactuar con interfaces (GUI), APIs, diversas bases de datos, y ejecutar acciones basadas en estas interacciones y en el aprendizaje continuo, traería una mayor ganancia de eficiencia, reducción de errores o creación de nuevo valor?

¿Qué tareas cognitivas complejas, que exigen juicio y adaptación, podrían ser aumentadas o eventualmente delegadas a agentes especializados, liberando el conocimiento humano para la supervisión estratégica y la innovación?

¿Cómo priorizar los casos de uso identificados (actuales y futuros) con base en el impacto potencial (valor transformacional), la complejidad de la implementación, los riesgos asociados (incluyendo los riesgos de autonomía) y el alineamiento estratégico? (Usar la matriz Riesgo x Impacto x Complejidad - Cap 3, p. 26).

03 Preparación Tecnológica y de Datos:

¿Nuestros sistemas principales (CRM, ERP, plataformas de negociación, etc.) poseen APIs robustas, accesibles y seguras para que los agentes puedan interactuar y ejecutar acciones? ¿Existen lagunas?

¿Tenemos acceso a datos (estructurados y no estructurados) de alta calidad, actualizados, relevantes y suficientemente diversos para entrenar a los agentes capaces de generalizar y operar eficazmente en escenarios complejos? (Ref. Convicción 2, Cap 5)

¿Cuál es nuestra arquitectura preferencial (Nube, On-Premise, Híbrida - Capítulo 6 - Seguridad y Privacidad de los Datos como una Prioridad Estratégica) para hospedar y operar estos agentes, considerando la seguridad, la latencia, la soberanía de los datos y la conformidad?

¿Qué herramientas de orquestación de agentes (ejemplo: LangChain, CrewAI - Cap 2), monitorización, observabilidad y control (para intervención humana) necesitamos implementar?

04 Gobernanza, Riesgo y Conformidad (GRC):

¿Cómo definiremos los "límites de protección" éticos y operacionales para los agentes autónomos en esta área? ¿Qué acciones están permitidas, cuáles exigen una supervisión de "humano-en-el-bucle", y cuáles están estrictamente prohibidas? (Ref. Cap 5 y Framework de Gobernanza p.53)

¿Qué requisitos regulatorios específicos (ejemplo: KYC, AML, Basilea, Solvencia II, LGPD/GDPR, reglas de idoneidad) se aplican a las acciones autónomas que los agentes ejecutarán? ¿Cómo garantizaremos la conformidad continua y auditable?

¿Cómo garantizamos la auditabilidad, la explicabilidad (XAI donde sea aplicable) y la trazabilidad de las decisiones y acciones tomadas por los agentes autónomos?

¿Cuál es el proceso para lidiar con errores, sesgos, comportamientos inesperados o "deriva" del modelo en los agentes? ¿Quién es responsable de la intervención, el reentrenamiento y el desmantelamiento, si es necesario?

05 Capacitación Humana y Gestión de Cambios:

¿Qué nuevas habilidades serán necesarias para que nuestros equipos diseñen, entrenen, supervisen, validen y colaboren efectivamente con los agentes de IA autónomos? (Ref. Framework de Change Management p.60)

¿Cómo comunicaremos la introducción de la IA Agente y gestionaremos el impacto en los roles y responsabilidades existentes, enfatizando la colaboración humano-máquina?

¿Qué programas de entrenamiento y aculturación (Ref. Cap 7.2) son necesarios para construir confianza, desmitificar la tecnología y garantizar la adopción eficaz por todas las partes interesadas?

¿Cómo fomentaremos una cultura de experimentación controlada, aprendizaje continuo y responsabilidad compartida en la operación de la IA Agente?

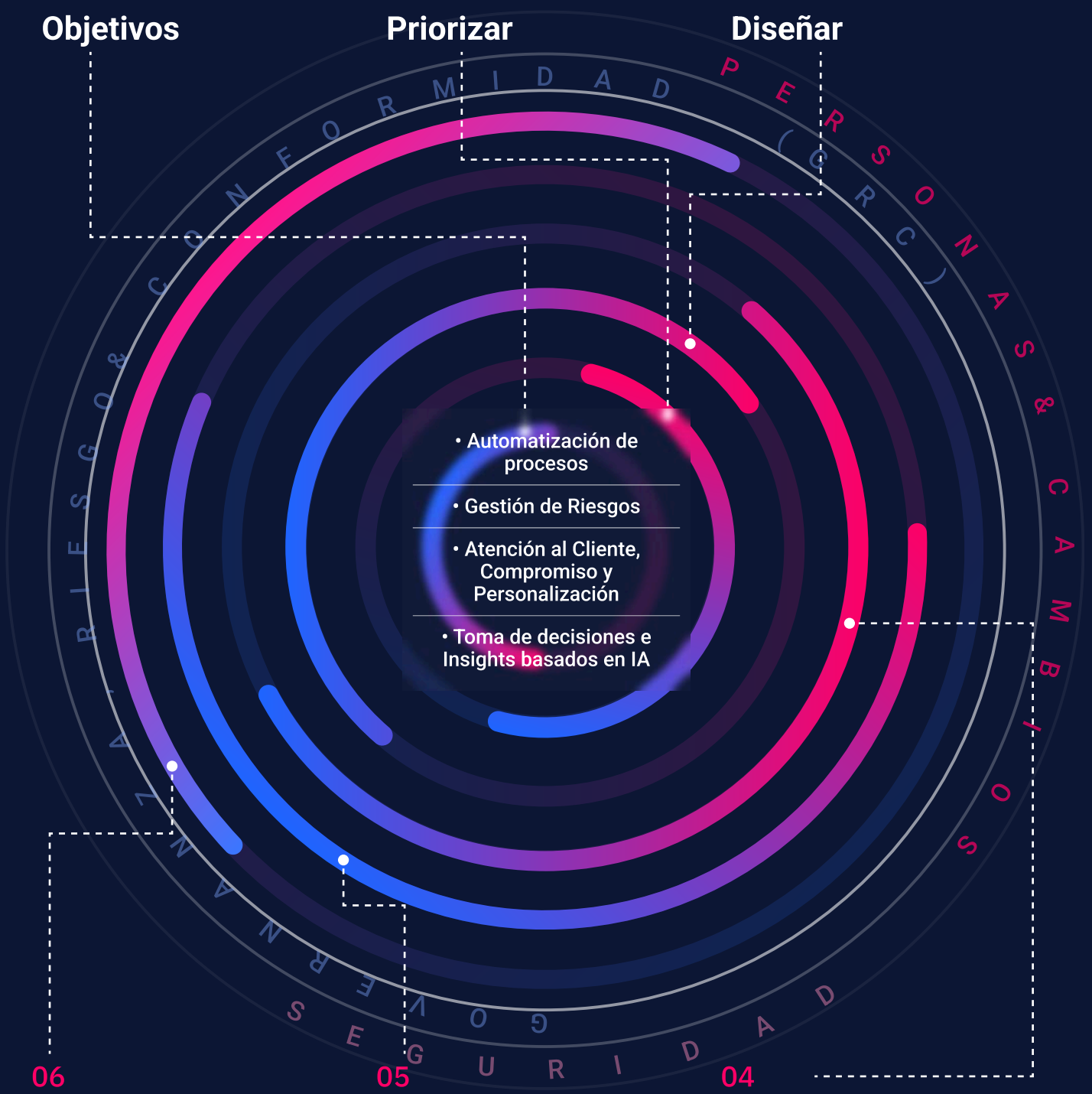
06 Medición del Rendimiento y Mejora Continua:

¿Qué KPIs específicos medirán el éxito de la implementación de la IA Agente en relación a los objetivos de negocio definidos (ejemplo: tiempo de ciclo de principio a fin, costo por proceso automatizado, NPS, tasa de error en decisiones autónomas, ROI, nuevas recetas generadas)? (Ref. Convicción 6, Cap 5)

¿Cómo implementaremos mecanismos de retroalimentación robustos (humanos y automatizados) para refinar continuamente el rendimiento, el comportamiento y la seguridad de los agentes? (Ref. Feedback Loops, Cap 2 y Cap 3)

¿Con qué frecuencia revisaremos el rendimiento de los agentes, la eficacia de la estrategia general y el alineamiento con el ambiente regulatorio y de mercado, ajustando el roadmap según sea necesario? (Ref. Enfoque Ágil, página 61)

Framework Visual: El Ciclo Estratégico de Implementación de IA Agente



Adaptando el Framework por Área de Negócio (con ejemplos de la Tabla p.34)

> Aunque el framework sea general, el énfasis en cada etapa y las respuestas a las preguntas variarán, como se ejemplifica por las ideas base (Tabla de la p. 34) y el potencial futuro (p. 74)

Banca Minorista:

Enfoque intenso en automatizar la jornada del cliente y el crédito, personalizar las interacciones y las recomendaciones proactivas, y gestionar los riesgos operacionales y de impago de forma predictiva. GRC crítico en los datos del consumidor (LGPD/GDPR) y la explicabilidad para el crédito.

Pagos

Enfoque extremo en la automatización de la conciliación/liquidación compleja, la detección/desmantelamiento autónomo de fraudes sofisticados, y el soporte/ gestión financiera proactiva para los clientes. La velocidad, la precisión, la seguridad y la resiliencia de los agentes son primordiales, con un GRC enfocado en la seguridad de las transacciones y el compliance regulatorio (PSD2, etc.).

Seguros

Casos de uso clave en la automatización de principio a fin y la optimización de siniestros (incluyendo la negociación), la suscripción dinámica y predictiva, la personalización proactiva de pólizas y la gestión de riesgo/prevención basada en datos continuos. GRC enfocado en los datos del asegurado, la equidad en la suscripción y la prevención de fraudes.

Banca de Inversión

Prioridad en la automatización de procesos de trade complejos, la inteligencia de riesgo de mercado en tiempo real con acción autónoma, la gestión de la relación consultiva y activa (incluyendo la ejecución discrecional) y la identificación/ejecución proactiva de oportunidades. Énfasis en un GRC riguroso, la seguridad cibernética y la explicabilidad (XAI) para las decisiones de alto valor.

Conclusión: Navegando el Futuro Agente con Estrategia

La implementación exitosa de la IA Agente no sucederá por casualidad. Requiere un enfoque deliberado, estratégico y adaptado a los matices de cada área de negocio dentro de la institución financiera, mirando más allá de las aplicaciones actuales hacia el potencial transformador de la autonomía inteligente. Al utilizar este framework para hacer las preguntas correctas, evaluar la preparación, priorizar las oportunidades visionarias y, crucialmente, establecer una gobernanza robusta y una cultura de aprendizaje y responsabilidad, las organizaciones pueden navegar con confianza en esta nueva frontera.

La IA Agente ofrece la promesa de transformar radicalmente la eficiencia, la experiencia del cliente, la gestión de riesgos y la propia naturaleza de la toma de decisiones financieras. Una estrategia bien definida, anclada en las posibilidades futuras y adaptada a su contexto específico, es la brújula necesaria para transformar esa promesa en una realidad tangible y sostenible, posicionando su área de negocio y su institución en la vanguardia de la innovación financiera.

CAPÍTULO 8

¿Por qué Artefact?

Por que confiar en Artefact:

Experiencia comprobada en la industria:

Poseemos una vasta experiencia en el sector Financiero, entregando soluciones prácticas y resultados medibles para las principales empresas líderes. Actuamos en todas las etapas de la cadena de valor, con énfasis en la innovación, la eficiencia y el crecimiento sostenible.

Metodología Estructurada y Personalizada:

Nuestro enfoque combina análisis detallados con soluciones a medida, equilibrando la visión estratégica y la ejecución práctica. Monitorizamos continuamente los resultados para garantizar el éxito de nuestros clientes.

Visión de la IA como Unidad de Negocios:

Incorporamos la visión de Datos e Inteligencia Artificial como activos generadores de valor, de forma que la tecnología esté al servicio del negocio. Con un enfoque en resolver los desafíos de las áreas de negocio, las aplicaciones deben proveer resultados tangibles y medibles, ya sea con tomas de decisiones más asertivas, optimización y automatización de procesos y generación de conocimientos procesables en tiempo real, posicionando a su empresa por delante de la competencia.

Compromiso con Resultados Concretos y Sostenibles:

Nuestro objetivo es entregar resultados sólidos y de largo alcance, promoviendo la innovación continua y la eficiencia estratégica para posicionar a su empresa como referencia en el mercado.

Negocios en los que ya hemos tenido un impacto con GenAI y Agentic AI

BNP PARIBAS

ACELERACIÓN DE DATOS DE IA PARA BCEF

Activamos el modelo AI Factory con más de 20 casos de uso enfocados en ahorro. Optimizamos el scoring y bloqueamos fraude mediante procesos automatizados. Desplegamos GenAI y MLOps, logrando una estructura escalable y eficiente.

SOCIETE GENERALE

ACELERACIÓN DE DATOS DE IA PARA RBF

Definimos la monetización de datos y el programa BAAS con casos prioritarios. Potenciamos el cross-selling mediante segmentación de clientes y dashboards. Impulsamos la democratización con formación masiva en analítica y ESG.

AXA

ACELERACIÓN DE DATOS DE IA PARA RBF

Formamos a 140,000 empleados en cultura de datos y lanzamos estrategias de marketing digital. Optimizamos el scoring crediticio, la detección de fraude y el CRM inteligente. Implementamos GenAI para ESG y consolidamos MLOps en solo tres meses.

Algunos de nuestros clientes

Santander · BNP PARIBAS · American Century · Societe Generale · bpifrance · adyen · Bank of Africa BMCE Group · LCL Your Data Center · Vanguard · ANBIMA · Harmonie mutuelle · malakoff humanis · Entrepreneurs d'Assurances · MAIF · Banque Populaire Faite pour vous · AG2R La Mondiale · J.P.Morgan · astorg · AXA · Rothschild & Co · defaqto · Amundi Asset Management · Qonto · Prudential · MACIF · matmut · Standard & Poor's · ARDIAN · Schroders

Aceleración de IA/Datos para BCEF

Estrategia de datos

Definición del modelo operativo de la Fábrica de IA/Datos

Identificación de más de 20 casos de uso (enfoque en la reducción de costes)

Fábrica de IA/Datos: Varios casos de uso entregados/en curso:

Optimización de la puntuación crediticia corporativa

Optimización de la detección de fraudes (+ millones de euros de fraude adicional bloqueado)

CRM autónomo (+ millones de euros de PNB)

Agente conversacional GenAI (impuesto sobre sucesiones)

Fundamentos de datos

Desarrollo de buenas prácticas de MLOps

Aportación de experiencia en la estructuración de plataformas de datos/IA

Aceleración de IA/datos para RBF

Estrategia de datos

Definición de la estrategia de monetización de datos e identificación de prioridades (programa BAAS)

Fábrica de IA/Datos: Desarrollo de múltiples casos de uso

Paneles de control (dashboards)

Agrupación de clientes

Venta cruzada / aumento de ventas (upsell / cross sell)

Democratización de datos

Formación a gran escala de profesionales en datos aplicada a dos ejemplos concretos: producción de paneles de control y uso de datos ESG.

Aceleración de IA/Datos

Aculturación: Formación para los 140 000 empleados del grupo.

Implementación de contenido digital y microaprendizaje que abarca los 7 temas principales de la gestión de datos.

Desarrollo de un manual de gestión de datos.

Definición de la estructura del contenido de la formación adaptada al contexto de AXA.

Formación lanzada a nivel de grupo en 3 meses.

Programa de marketing de datos

Aceleración del marketing de datos para mejorar las ventas en la red francesa: herramientas/organización/medición (en curso)

Apéndice

CONTENIDO TÉCNICO

Framework para la implementación de Agentic y GenAI en las instalaciones

En este apéndice exclusivo, reunimos los pilares técnicos esenciales para guiar un viaje seguro y eficiente en la adopción de GenAI y Agentic AI en el sector financiero. Arquitecturas recomendadas, prácticas de seguridad, cumplimiento normativo y estrategias de escalabilidad, todo ello adaptado al contexto de las soluciones locales.

Framework para la implementación de Agentic y GenAI en las instalaciones

La adopción de soluciones GenAI y Agentic AI en entornos locales ofrece importantes ventajas estratégicas, especialmente en lo que respecta a la seguridad de la información, el cumplimiento normativo y el control de los recursos informáticos. Sin embargo, este enfoque también plantea retos técnicos y operativos que requieren una planificación detallada.

Desde la definición de la infraestructura de hardware hasta las capas de seguridad, red, almacenamiento y mantenimiento, cada componente de la arquitectura debe alinearse cuidadosamente con la realidad del negocio, la sensibilidad de los datos procesados y la frecuencia de uso de los modelos. Esto es aún más crítico cuando se trata de automatizaciones autónomas y flujos de decisión continuos, como los que permiten los agentes inteligentes.

Basándose en experiencias prácticas y conocimientos técnicos especializados, este capítulo presenta los principales aspectos que deben tenerse en cuenta para la implementación exitosa de soluciones GenAI y Agentic AI en entornos locales.

Requisitos de Hardware

En lo que respecta a la potencia computacional, los proyectos más pequeños pueden ejecutarse de manera eficaz en servidores con CPU que tengan un alto número de núcleos, logrando un buen equilibrio entre coste y rendimiento. Sin embargo, a medida que aumenta la demanda, ya sea por el volumen de datos o por la complejidad de los modelos utilizados, se recomienda el uso de GPU especializadas (como las NVIDIA A100 o H100) para reducir drásticamente el tiempo de entrenamiento. Así, se perfilan dos escenarios típicos:

- **Baja demanda:** uso de CPU de alto rendimiento, con potencial de escalabilidad gradual a medida que crecen las necesidades.
- **Alta demanda:** infraestructura basada en GPU de última generación, que incluye clústeres de múltiples servidores y orquestación dinámica (por ejemplo, Kubernetes).

En lo que respecta al almacenamiento, los proyectos iniciales suelen satisfacerse con soluciones menos complejas, como SSD, HDD locales o incluso un NAS (almacenamiento conectado a la red) de pequeño tamaño para compartir datos. Sin embargo, cuando el volumen de información y la criticidad aumentan, surgen mayores exigencias en cuanto a velocidad y tolerancia a fallos. En este escenario, entran en escena los sistemas de almacenamiento distribuido (por ejemplo, Ceph o GlusterFS) y las unidades SSD NVMe, que proporcionan un alto rendimiento de E/S y resiliencia. Una vez más, se pueden destacar dos niveles de demanda:

- **Baja demanda:** SSD locales o un NAS sencillo, asociados a políticas básicas de copia de seguridad.
- **Alta demanda:** almacenamiento distribuido con SSD NVMe, diseñado para ofrecer redundancia, disponibilidad y escalabilidad.

En lo que respecta a la red, la disparidad entre las demandas menores y las operaciones más complejas es bastante clara. Para entornos más pequeños, las conexiones de 1GbE o 10GbE suelen ser suficientes, y para el aislamiento básico, las VLAN o VPN simples pueden ser suficientes. Sin embargo, cuando se busca un alto rendimiento y una baja latencia (por ejemplo, en escenarios de HPC, High-Performance Computing), es fundamental utilizar conexiones a partir de 25GbE o incluso Infiniband. En estas situaciones, también es necesario invertir en segmentación de red avanzada y firewalls robustos para preservar la seguridad a gran escala:

- **Baja demanda:** redes de 1 GbE o 10 GbE y segmentación simple mediante VLAN/VPN.
- **Alta demanda:** conexiones ≥25GbE o Infiniband, con segmentación avanzada y supervisión continua del tráfico.

Al estructurar la arquitectura de IA local desde esta perspectiva modular, diferenciando adecuadamente las cargas de trabajo más pequeñas de las más grandes, la asignación de recursos se vuelve más precisa. De esta manera, se evita el sobredimensionamiento inicial y se garantiza la posibilidad de expansión sin necesidad de interrumpir los servicios. Esto establece una base sólida para proyectos de GenAI y Agentic AI que priorizan la seguridad, el rendimiento y la escalabilidad, además de crear un camino para futuras adaptaciones.

Tech Stack

La definición adecuada de un Tech Stack es un paso fundamental para alcanzar el máximo rendimiento y la escalabilidad deseada en proyectos de IA locales. A continuación se enumeran los principales componentes recomendados:

Sistema Operacional

Las distribuciones Linux como Ubuntu o CentOS son adecuadas para cargas de trabajo de IA, ya que combinan **estabilidad, amplio soporte comunitario y facilidad de optimización de hardware avanzado** (GPU y CPU multinúcleo).

Frameworks de IA

Al elegir un marco de IA, adapte su decisión a las necesidades del proyecto, la experiencia del equipo y la escalabilidad. Para el aprendizaje profundo, TensorFlow ofrece escalabilidad, PyTorch flexibilidad y Keras facilidad de uso. En el aprendizaje automático tradicional, Scikit-Learn es intuitivo, mientras que XGBoost y LightGBM son ideales para datos estructurados. Para la visión artificial, Caffe destaca. Por su parte, marcos como LangChain, OpenAI y Hugging Face se recomiendan para LLM y GenAI. Tenga en cuenta también el soporte de la comunidad, la integración con su pila y el mantenimiento a largo plazo.

Contenedorización, orquestación y empaquetado

El uso de contenedores, como Docker, facilita la creación de entornos aislados para cada modelo o servicio, lo que simplifica tanto el desarrollo como la implementación. Paralelamente, Kubernetes es la opción estándar para la orquestación, debido a su capacidad para gestionar la escalabilidad horizontal y proporcionar mecanismos de tolerancia a fallos, factores críticos en escenarios con múltiples servicios de IA ejecutándose simultáneamente.

Para aprovechar al máximo los recursos de hardware y la pila de software seleccionada, hay algunas configuraciones que son especialmente importantes:

- **Clúster Kubernetes con soporte para GPU:** la adopción del complemento de dispositivo Kubernetes de NVIDIA permite que las aplicaciones en contenedores utilicen los recursos de la GPU según la demanda, lo que optimiza el entrenamiento y reduce el tiempo de inferencia.
- **Monitorización de recursos:** Herramientas como Prometheus y Grafana permiten el seguimiento en tiempo real de métricas de CPU, GPU, memoria y red, lo que facilita la identificación de cuellos de botella en el rendimiento y orienta las mejoras continuas.

Además, la contenedorización garantiza la portabilidad y el aislamiento del entorno, lo que simplifica tanto el desarrollo como el mantenimiento de diferentes versiones del modelo. Herramientas como Docker y plataformas de inferencia (por ejemplo, FastAPI o Flask) ayudan a componer el ecosistema de implementación. Además de permitir el empaquetado del modelo, ya que al crear la imagen Docker se incluyen todas las dependencias necesarias, desde bibliotecas de IA (PyTorch, TensorFlow, Transformers, etc.) hasta el modelo en su estado optimizado.

Ejemplo de Dockerfile:
A continuación, se muestra un ejemplo ilustrativo de Dockerfile que instala las dependencias básicas, copia el modelo al contenedor y expone el servicio de inferencia:

```
Unset
# Ejemplo de Dockerfile
FROM nvidia/cuda:11.8-base
RUN pip install torch transformers fastapi uvicorn
COPY model/ /app/model/
COPY app/ /app
WORKDIR /app
CMD ["uvicorn", "main:app", "--host", "0.0.0.0", "--port", "8000"]
```

Este ejemplo utiliza como base la imagen de NVIDIA con CUDA 11.8, instala las bibliotecas necesarias, copia los archivos del modelo y de la aplicación dentro del contenedor e inicia un servidor Uvicorn para exponer la API de inferencia.

Selección del modelo: Para mantener un control total sobre la implementación y garantizar la adaptabilidad, se recomienda utilizar modelos de código abierto, como Llama 3 o Mixtral. Estos pueden ajustarse (fine-tuned) con datos propietarios, en caso de que sea necesario adaptar el comportamiento o el rendimiento a escenarios específicos.

- **Modelos de código abierto:** evitan la dependencia excesiva de los proveedores y aumentan la transparencia en relación con el funcionamiento interno.
- **Ajuste fino (Fine-tuning):** Cuando sea necesario, aplique técnicas de ajuste fino sobre los datos corporativos, preservando la competitividad y la relevancia del modelo para el ámbito empresarial en cuestión.

Optimización del modelo: Una vez seleccionado el modelo, es posible optimizarlo para reducir la latencia de inferencia y el consumo de recursos computacionales. Entre las estrategias habituales destacan la cuantificación y el paralelismo:

- **Cuantificación:** convertir modelos a formatos de menor precisión, como FP16 o INT8, reduce el uso de memoria y mejora la velocidad de inferencia, especialmente en GPU de última generación.
- **Paralelismo:** En modelos muy grandes, el uso del paralelismo de modelos o tensores puede distribuir el procesamiento entre varios nodos o GPU, reduciendo así el tiempo total de ejecución.

Implementación: Una vez creado el contenedor, el siguiente paso es integrar la plantilla en el entorno Kubernetes. Esto se puede hacer mediante archivos YAML o gráficos Helm (Helm Charts). En ambos casos, hay que tener en cuenta algunos aspectos esenciales:

Gestión de la escalabilidad: Horizontal Pod Autoscaler (HPA) ajusta automáticamente el número de réplicas de pods a medida que varía la carga de trabajo. Supervisa métricas como el uso de la CPU o la GPU y crea o elimina pods para mantener el rendimiento.

Supervisión y registros: vincule la implementación a soluciones de supervisión (como Prometheus y Grafana) para realizar un seguimiento en tiempo real de las métricas de uso de CPU, GPU, memoria y latencia. Esto ayuda a identificar cuellos de botella y a tomar decisiones de escalabilidad y optimización.

Seguridad y Conformidad

La seguridad de los datos y el cumplimiento normativo son pilares fundamentales en los proyectos de IA locales, especialmente cuando se manipula información confidencial o sujeta a regulaciones estrictas. A continuación, se indican las principales prácticas para proteger la integridad de los datos, adaptar la implementación a la legislación pertinente y reforzar la seguridad de la red.

Protección de Datos

Para mantener los datos corporativos en un entorno seguro y controlado, se recomienda un enfoque de múltiples capas:

- **Almacenamiento Cifrado:** Es fundamental implementar el cifrado en reposo (at rest) y en tránsito (in transit), manteniendo los datos ilegibles incluso en situaciones de acceso indebido.
- **Redes Privadas:** Restringir la comunicación del clúster a redes privadas internas, evitando la exposición de puertos y servicios críticos al entorno externo.
- **Control de acceso basado en funciones (RBAC):** configurar Kubernetes con RBAC garantiza que cada función dentro de la organización reciba permisos específicos, lo que evita que los usuarios sin autorización tengan acceso a recursos confidenciales.

Conformidad

Dependiendo del sector, los requisitos legales y normativos pueden variar considerablemente. Cada proyecto de IA debe ajustarse a las normas correspondientes, tales como:

- **RGPD:** Orientado a la protección de los datos personales de los ciudadanos de la Unión Europea, exige transparencia en el tratamiento de los datos.
- **HIPAA:** Aplicable a la salud en EE. UU., establece normas para proteger y mantener la privacidad de la información médica.
- **ISO 27001:** Certificación internacional orientada a la gestión de la seguridad de la información, que contempla políticas de control, supervisión y mejora continua.
- **LGPD:** Ley brasileña que regula el tratamiento de datos personales, garantizando los derechos de los titulares e imponiendo responsabilidades a las organizaciones en cuanto a la recopilación, el uso y el almacenamiento de datos.

Estos puntos de referencia guían la creación de políticas internas sólidas, que garantizan la confidencialidad, integridad y disponibilidad de los datos.

Seguridad de red

La infraestructura de red en entornos locales debe configurarse para contener las amenazas externas y minimizar los riesgos internos:

Nube privada virtual (VPC): Utilizar direcciones IP privadas y segmentar subredes críticas, de modo que solo los servicios fiables accedan al clúster de IA.

Cortafuegos y sistemas de detección de intrusiones (IDS): definir reglas de cortafuegos para filtrar puertos y protocolos sensibles y aplicar IDS para bloquear intentos de intrusión o detectar tráfico anormal.

Monitorización continua: integrar herramientas como Prometheus y Grafana para observar eventos de seguridad en tiempo real y generar alertas en caso de actividades sospechosas.

Escalabilidad y Optimización del Rendimiento

Los proyectos de Agentic AI o Gen AI pueden expandirse rápidamente en términos de volumen de datos, número de usuarios y complejidad de los modelos. Para acompanhar este crecimiento sin comprometer la calidad, es imprescindible aplicar prácticas de escalabilidad y optimización del rendimiento. A continuación se presentan tres recomendaciones clave:

Equilibrio de Carga

El equilibrio de carga distribuye las solicitudes de manera uniforme entre los recursos disponibles, evitando la sobrecarga. Herramientas como NGINX o Traefik pueden configurarse para enrutar el tráfico de manera eficiente en un clúster de Kubernetes, lo que aumenta la fiabilidad y la velocidad de respuesta.

- **NGINX o Traefik:** Ajuste las reglas de enrutamiento para que los pods con mayor capacidad de procesamiento reciban más solicitudes de forma proporcional.
- **Tolerancia a Fallos:** en caso de caída de un nodo, el tráfico se redirige sin afectar al servicio.

Caché

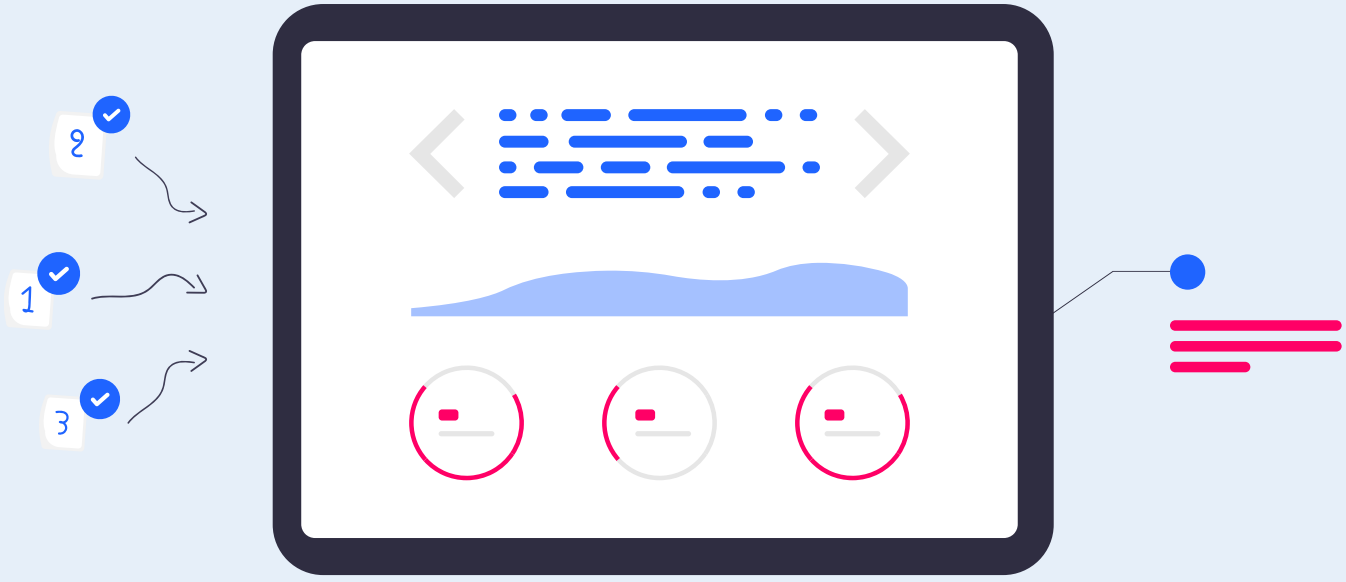
Los mecanismos de caché son extremadamente útiles cuando se accede con frecuencia a determinadas inferencias o resultados de modelos. Herramientas como Redis permiten almacenar en memoria las respuestas más solicitadas, lo que reduce considerablemente el tiempo de respuesta.

- **Reducción de la Latencia:** al servir las respuestas directamente desde la memoria, no es necesario volver a procesar el modelo con cada solicitud.
- **Reducción de la Carga:** la caché reduce la presión sobre los servidores de IA, liberando recursos para otras inferencias o tareas más complejas.

Paralelismo

Para gestionar grandes volúmenes de datos o modelos de alta complejidad, el paralelismo permite dividir la carga de trabajo, lo que acelera la ejecución. Marcos como Ray o Dask simplifican la distribución de tareas de entrenamiento o inferencia en varios nodos, lo que aumenta la eficiencia y reduce el tiempo de procesamiento.

- **Marcos de Computación Distribuida:** Ray y Dask ofrecen API accesibles para paralelizar operaciones sin necesidad de gestionar directamente la comunicación entre nodos.
- **Escalado Horizontal:** con el aumento del número de nodos, el rendimiento tiende a crecer linealmente, lo que permite la ejecución de cargas de trabajo cada vez mayores.



Monitoreo y Mantenimiento

Para mantener la eficiencia y la disponibilidad de las soluciones GenAI o Agentic AI en un entorno local, es fundamental establecer procedimientos sólidos de supervisión y mantenimiento. Estas prácticas permiten detectar cuellos de botella en el rendimiento, realizar un seguimiento de los registros y actualizar los modelos, lo que garantiza que las soluciones sigan satisfaciendo las necesidades del negocio.

Herramientas de Monitoreo

Una arquitectura de monitorización integral facilita la identificación y resolución ágil de problemas. Prometheus y Grafana se utilizan a menudo conjuntamente para recopilar métricas (CPU, memoria, uso de GPU, latencia, etc.) y mostrar paneles que reflejan el estado general del sistema.

Paralelamente, una pila de registros, como ELK (Elasticsearch, Logstash, Kibana), unifica la gestión de registros en un único punto. Esto simplifica el diagnóstico de incidentes, permitiendo identificar rápidamente comportamientos anormales o fallos, además de mantener un historial detallado para auditorías o futuras investigaciones.

- Prometheus + Grafana: Supervisión y visualización en tiempo real de las métricas más críticas del clúster.
- ELK Stack: Centralización y análisis de registros, lo que permite corregir problemas y verificar el historial de manera más eficiente.

Reciclaje y Actualizaciones del modelo

A medida que cambian los datos de entrada y surgen nuevos conocimientos, mantener el modelo actualizado es fundamental para preservar la precisión y la relevancia de la solución de IA. La integración de los procesos de integración continua y entrega continua (CI/CD) en el proceso de reentrenamiento permite automatizar etapas como la carga de datos, la verificación del rendimiento del nuevo modelo y, eventualmente, la implementación en producción.

Herramientas como ArgoCD o Jenkins ayudan a controlar las versiones del modelo, liberando actualizaciones incrementales y ejecutando reversiones cuando se producen inconsistencias o pérdidas significativas de rendimiento. Este ciclo de mejora continua acelera el desarrollo, minimiza los riesgos y garantiza que el modelo en producción siga siendo adecuado para las demandas del negocio.

- Pipelines de CI/CD: Permiten un reentrenamiento frecuente con datos actualizados y pruebas de validación automatizadas.
- Automatización de la implementación: soluciones como ArgoCD o Jenkins facilitan la entrega de modelos revisados, garantizando la trazabilidad y el control de versiones.

Descripción General del Flujo de Implementación de Agentic AI / Gen AI On Premise

ETAPA	TEMA	SUBTEMA	POSIBLES HERRAMIENTAS/ MARCOS UTILIZADOS
Hardware	Potencia Computacional	Baja Demanda	CPUs de alto rendimiento
		Alta Demanda	GPU de última generación (Ejemplo: Kubernetes)
	Almacenamiento	Baja Demanda	SSD locales o un NAS sencillo
		Alta Demanda	NVMe SSD
	Red	Baja Demanda	Redes de 1GbE o 10GbE (segmentación VLAN/VPN)
		Alta Demanda	Conexiones ≥25GbE o Infiniband
Tech Stack	Sistema Operacional	---	Linux (Ubuntu ou CentOS)
	Frameworks	AI	PyTorch, TensorFlow o Keras
		Gen AI	LangChain, OpenAI y Hugging Face
		ML	Scikit-Learn, XGBoost o LightGBM
	Contenedorización, orquestación y empaquetado	---	Docker, Kubernetes
	Selección del modelo	---	Llama 3 o Mixtral
	Optimización del modelo	---	Cuantización y/o Paralelismo
	Implementación	---	Archivos YAML o gráficos Helm (Helm Charts)
Seguridad y Conformidad	Protección de Datos	---	RBAC
	Conformidad	---	---
	Seguridad de red	---	VPC
Escalabilidad y Optimización del Rendimiento	Equilibrio de carga	---	NGINX o Traefik
	Caché	---	Redis
	Paralelismo	---	Ray o Dask
Monitoreo y Mantenimiento	Herramientas de monitoreo	---	Prometheus + Grafana o ELK Stack
	Reciclaje y Actualizaciones del modelo	---	ArgoCD o Jenkins

Glosario

A - G

IA Agente

Un tipo de inteligencia artificial que añade autonomía y capacidad de toma de decisiones a los sistemas, permitiendo que actúen de forma proactiva, aprendan de la experiencia y ejecuten tareas complejas con mínima intervención humana.

Análisis Predictivo

Técnica de análisis de datos que utiliza estadísticas y algoritmos para prever tendencias y patrones futuros con base en datos históricos.

API (Application Programming Interface)

Conjunto de reglas y herramientas que permite la integración de aplicaciones con sistemas de IA, facilitando la comunicación y el intercambio de datos.

Automatización de Procesos

Uso de tecnología para realizar tareas repetitivas y operacionales de forma eficiente, reduciendo la necesidad de intervención humana.

Chatbot

Programa basado en IA que interactúa con los usuarios en lenguaje natural para responder preguntas, proporcionar información o realizar acciones automatizadas.

Compliance

Conjunto de normas y regulaciones que las empresas deben seguir para garantizar la conformidad legal y regulatoria, especialmente en el sector financiero

Cifrado

Método de protección de datos a través de la codificación, garantizando la seguridad y la privacidad en las transacciones digitales.

Datos No Estructurados

Información que no sigue un formato predefinido, como textos, imágenes, audios y videos, haciendo que su procesamiento sea más complejo.

Detección de Fraudes

Uso de inteligencia artificial y análisis de datos para identificar actividades sospechosas y prevenir crímenes financieros, como el lavado de dinero y los fraudes bancarios.

GenAI (Inteligencia Artificial Generativa)

Subcampo de la IA enfocado en la creación de nuevos contenidos, como texto, imágenes, código y audio, a partir de modelos entrenados en grandes volúmenes de datos.

G - R

Governanza de Datos

Conjunto de procesos y políticas que garantizan la calidad, la seguridad y la conformidad del uso de datos dentro de una organización.

GUI (Graphical User Interface)

Interfaz visual que facilita la interacción con la IA de agentes, usando elementos gráficos como botones, menús y paneles para una navegación intuitiva.

Inteligencia Artificial (IA)

Campo de la ciencia de la computación que desarrolla sistemas capaces de realizar tareas que normalmente exigen inteligencia humana, como el aprendizaje, la toma de decisiones y el reconocimiento de patrones.

LLMs (Large Language Models)

Modelos de Lenguaje de Gran Escala que utilizan redes neuronales avanzadas para comprender y generar lenguaje natural, permitiendo aplicaciones como chatbots y asistentes virtuales.

Machine Learning (Aprendizaje Automático)

Técnica de IA que permite que los sistemas aprendan a partir de los datos y mejoren su rendimiento sin necesidad de programación explícita.

Modelos de Lenguaje

Algoritmos de IA entrenados para entender, procesar y generar texto con base en un vasto conjunto de datos de lenguaje natural.

Onboarding de Clientes

Proceso de integración de nuevos clientes en una institución financiera, garantizando que atiendan a los requisitos regulatorios y tengan acceso a productos y servicios personalizados.

Personalización Avanzada

Uso de la IA para adaptar productos, servicios y experiencias con base en las necesidades y preferencias individuales de los clientes.

RAG (Retrieval-Augmented Generation)

Técnica utilizada en sistemas de inteligencia artificial generativa, donde el modelo combina la generación de texto con la recuperación de información relevante de una base de datos o documentos externos para aumentar la precisión y la relevancia de las respuestas.

RPA (Robotic Process Automation)

Tecnología que usa softwares para automatizar procesos empresariales repetitivos, aumentando la eficiencia operacional.

R - X

Seguridad de Datos

Prácticas y tecnologías utilizadas para proteger la información sensible contra los accesos no autorizados, las fugas o los ataques cibernéticos.

Transformación Digital

Proceso de integración de tecnologías digitales en todas las áreas de una empresa para mejorar las operaciones, la eficiencia y la experiencia del cliente.

XAI (Explainable AI)

La Inteligencia Artificial Explicable es una corriente de la Inteligencia Artificial que busca aportar comprensión a los resultados de los modelos y sistemas inteligentes, de manera de justificar y viabilizar la auditabilidad de los procesos con IA incrustada.

Enlaces y Referencias

[1] Artículo

The brief history of artificial intelligence: the world has changed fast — what might be next? Our World in Data. Disponível em: https://ourworldindata.org/brief-history-of-ai

[2] Investigación

NVIDIA. *State of AI in Financial Services: 2025 Trends*.

[3] Material Institucional

Artefact. Conocimiento interno.

[4] Investigación

Artefact. Generative AI Survey – The Technology, the Rewards & the Risks. Documento interno (Artefact-GenAI-Survey.pdf).

[5] Artículo

Citi GPS. Agentic AI: Finance & the 'Do It For Me' Economy. Jan. 2025.

[6] Artículo

GODHANI, Sahaj. Agentic AI will transform financial services with autonomy, efficiency, and inclusion. Medium – InsiderFinance Wire. Disponível em: https://medium.com/insiderfinance-wire/agentic-ai-will-transform-financial-services

[7] Video | Evento

MARTINEZ, Joffrey. AI in Financial Services: Key Market Trends and Insights for 2024. AI For Finance Event 2024. Powered by Artefact.

[8] Reporte

The Alan Turing Institute. The AI Revolution – Opportunities and Challenges for the Finance Sector. Nov. 2024. Disponível em: https://www.turing.ac.uk/sites/default/files/2024-11/the_ai_revolution_-_opportunities_and_challenges_for_the_finance_sector_-_report_1.pdf

[9] Artículo

AI implementation in the financial sector: legal challenges. WhatNext.Law. 8 nov. 2024. Disponível em: https://whatnext.law/2024/11/08/ai-implementation-in-the-financial-sector-legal-challenges/

Materiales técnicos y estudios internos

Recopilado a partir de MVP y demostraciones internas de IA generativa aplicada al sector financiero, con especial atención a la automatización de documentos, API conversacionales, chatbots y copilotos.

Artefact – One Pagers y prototipos funcionales aplicados:

DraftAI – MVP / Atención al cliente con ChatBot cognitivo / Bot conversacional con integración bancaria / Inteligencia empresarial Agente aumentado / Procesamiento inteligente de documentos.

Estudios sobre GenAI y NLP aplicados en:

Mapeo de riesgos emergentes / Comprensión de llamadas con NLP / Análisis en profundidad de la documentación de capital privado / Aceleración del proceso de diligencia debida / Automatización del crédito mediante el procesamiento inteligente de documentos.