

f**AI**nance 5.0

GenAI e Agentic AI na transformação do setor





QUEM SOMOS:

Autoria & Equipe

A equipe da Autoria é composta por profissionais multidisciplinares com expertise em dados, tecnologia, negócios e inovação. Todo time Artefact trabalha de forma colaborativa para entregar soluções de alto impacto, sempre alinhadas às necessidades de nossos 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



SOBRE AUTORIA

Artefact

A Artefact acelera a adoção dos dados e da Inteligência Artificial para impactar positivamente as pessoas e as organizações. Oferecemos uma ampla gama de serviços, desde a estratégia até as operações, com a implementação de soluções de IA por setores industriais, que ajudam as empresas a capturar a vantagem competitiva da transformação data e AI.





Sumário

[🏠 Introdução](#)

1

[🏠 O Cenário Atual](#)

4

O Cenário da IA no Setor Financeiro

A IA em números: impactos e adoção no setor financeiro

5

O estado da arte da tecnologia no setor financeiro

8

[🏠 Agentic AI & Oportunidades](#)

13

O que é Agentic AI e por que ela representa uma oportunidade tão grande?

[🏠 GenAI & Agentic AI Aplicações](#)

24

GenAI e Agentic AI: aplicações e a Arte do Possível

[🏠 IA Está Redefinindo o Setor](#)

36

Casos de Sucesso: como a IA Está Redefinindo o Setor Financeiro

Otimizando a experiência do cliente com GenAI: um novo paradigma em atendimento ao consumidor

37

Desafios e Obstáculos

38

Desenvolvendo a Solução

38

Benefícios Alcançados

40

Destaques

40

Eficiência operacional: otimizando processos de Middle e Backoffice com GenAI

41

Desafios e Limitações

41

Solução Desenvolvida

42

Benefícios Alcançados

43

Reflexões

44



🏠 **Desafios no Setor & como superá-los**

45

Desafios e Estratégias para a Implementação de IA no Setor Financeiro

Convicções da Artefact sobre o uso de GenAI e Agentic AI no setor financeiro	46
Principais desafios para Implementação de IA no Setor Financeiro	48
Boas práticas para superar desafios e garantir uma implementação bem-sucedida	51
Machine Learning: Aprimorando a Imparcialidade e a Interpretabilidade	54
GenAI: Mitigando Alucinações e Garantindo a Segurança dos Dados	54
Agentic AI: Fortalecendo a Governança e os Limites Éticos	55

🏠 **Estratégia na implementação de AI**

56

Estratégia na implementação de AI no Setor Financeiro

Diagnóstico e Planejamento Estratégico	57
Engajamento das Partes Interessadas	59
Pilotos e Iterações: uma abordagem ágil para reduzir riscos e maximizar valor	61
A Segurança e Privacidade de Dados como uma Prioridade Estratégica	64
Soluções em Nuvem	66
Soluções On-Premise	67
Fatores Críticos para a Decisão	68
Tendências e Recomendações	68
Cuidados Especiais e Boas Práticas	69

🏠 **Framework Estratégico & Negócio**

72

Framework Estratégico para Implementação de Agentic AI por Área de Negócio

Introdução: Da Visão à Ação Estratégica	73
O Imperativo Estratégico: Por que Focar em Agentic AI Agora?	74
Dimensões Chave da Estratégia de Agentic AI	76
Perguntas Estratégicas Essenciais por Dimensão	77
Framework Visual: O Ciclo Estratégico de Implementação de Agentic AI	81
Adaptando o Framework por Área de Negócio	82
Conclusão: Navegando o Futuro Agentic com Estratégia	83



[🏠 Artefact como Parceiro](#)

84

Por que a Artefact?

Clientes que já impactamos com GenAI e Agentic AI

86

[🏠 Apêndice | Conteúdo Técnico](#)

89

Framework para a Implementação de Agentic e GenAI on-premise

Necessidades de Hardware

91

Tech Stack

93

Segurança e Conformidade

96

Escalabilidade e Otimização de Desempenho

98

Monitoramento e Manutenção

100

Visão Geral do Fluxo de Implementação de Agentic AI / Gen AI On Premise

102

[🏠 Glossário](#)

103

[🏠 Links & Referências](#)

106

INTRODUÇÃO

O setor financeiro está entre os mais dinâmicos e transformadores da economia global

Diante de um cenário em constante evolução, instituições bancárias, corretoras e fintechs enfrentam o desafio de se adaptarem rapidamente às novas demandas do mercado, avanços tecnológicos e expectativas dos consumidores, enquanto permanecem competitivas em um ambiente cada vez mais complexo.

Neste eBook, nosso objetivo é demonstrar como a **Inteligência Artificial (AI)** e suas diversas aplicações podem se tornar aliadas estratégicas para as instituições financeiras. Exploraremos desde as aplicações mais básicas, como automação de processos e análise de dados, até as tendências e inovações mais recentes impulsionadas pela **Agentic AI**, como o monitoramento em tempo real de transações financeiras e a detecção proativa de fraudes, redefinindo a forma como as empresas interagem com dados e tomam decisões.

Enquanto a **GenAI** revolucionou o mundo da tecnologia, possibilitando que sistemas pudessem criar conteúdos e interagir com o ser humano inclusive no mundo da fala, escrita e na arte, a **Agentic AI** adiciona uma camada de autonomia e tomada de decisão, permitindo que esses sistemas ajam de forma proativa, aprendam com a experiência e executem tarefas complexas com mínima intervenção, elevando ainda mais o impacto da tecnologia e AI na vida e relações cotidianas. No setor de serviços financeiros, isso significa assistentes que não apenas geram insights, mas também os aplicam, otimizando operações, detectando fraudes e aprimorando a gestão de riscos em tempo real, entre muitas outras oportunidades.

Com anos de experiência atuando ao lado dos maiores players do setor, a Artefact oferece soluções estratégicas e implantação de estudos de caso relevantes que integram de forma abrangente e inovadora essas tecnologias. Além disso, essa experiência prévia, juntamente com o espírito inovador e empreendedor que permeia a empresa, nos possibilita enxergar um vasto rol de oportunidades de casos de uso que pareciam impossíveis, mas que, com a evolução tecnológica, não só são possíveis, como serão a realidade do setor de serviços financeiros em pouco tempo. Essas iniciativas são projetadas para impulsionar o crescimento sustentável, promover a inovação contínua e maximizar resultados consistentes e duradouros



Esperamos que este material te auxilie a compreender o **valor estratégico da Inteligência Artificial** e como sua empresa pode colocar a tecnologia a serviço dos negócios e impulsionar os resultados através das aplicações dessa tecnologia para navegar com sucesso no **mercado financeiro**.

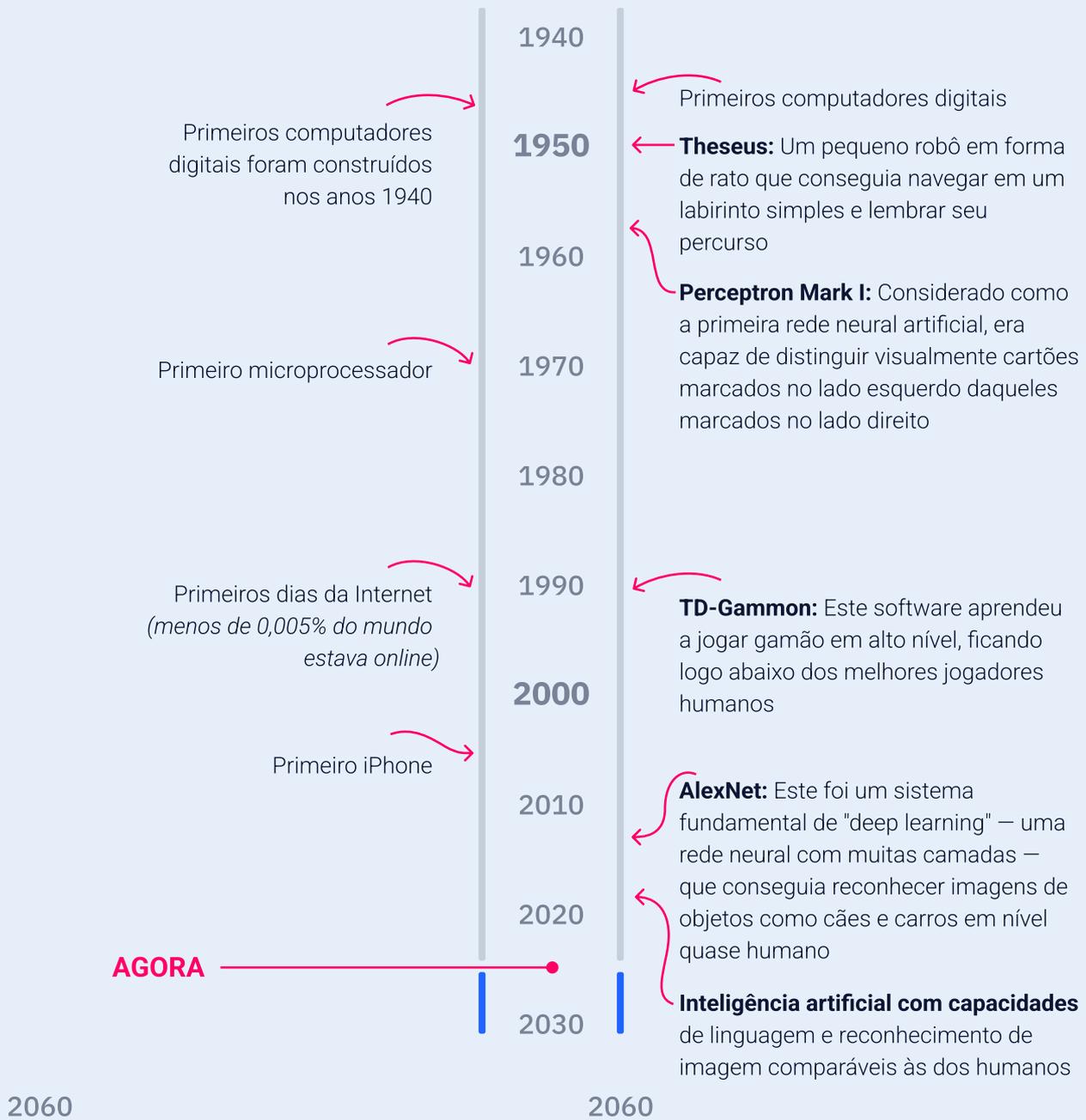


Boa leitura!



LINHA DO TEMPO

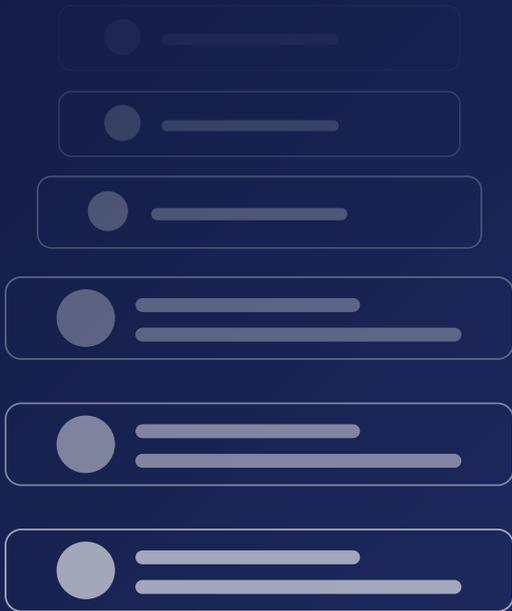
Sistemas Notáveis de Inteligência Artificial ^[1]



[1] Baseado em: Our World in Data – The brief history of artificial intelligence.

CAPÍTULO 1

O Cenário da AI no Setor Financeiro



A AI em números: impactos e adoção no setor financeiro

As instituições financeiras sempre lideraram o uso de dados, impulsionando inovações tecnológicas no setor. Desde os computadores mainframe nos anos 1980 até as atuais soluções avançadas como Inteligência Artificial Generativa (GenAI) e Inteligência Artificial Agêntica (Agentic AI), o setor financeiro tem sido pioneiro na aplicação dessas tecnologias.

EVOLUÇÃO DA TECNOLOGIA - LINHA DO TEMPO



1980's

Linguagem de Máquina



1998

Linguagem Assembly



2007

Linguagem de Instruções



2011

Linguagens Orientadas a Objetos



2022

GenAI



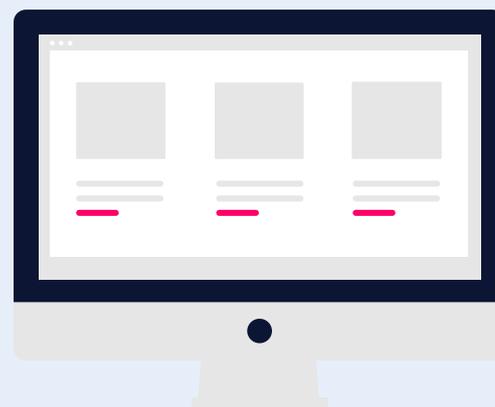
2023

Agentic AI

Os resultados obtidos por instituições financeiras com o uso de GenAI em 2024 são notáveis e comprovam o potencial dessa tecnologia em gerar valor no setor. Casos concretos mostram retornos expressivos: otimização de trading e portfólio com 25% de ROI, melhoria na experiência e no engajamento do cliente alcançando 21%, e ganho de eficiência em áreas operacionais críticas, como processamento de documentos e geração de relatórios, atingindo 11%. Esses números deixam claro que a GenAI já é uma realidade transformadora. Empresas que souberem aproveitar essa oportunidade agora estarão à frente da concorrência. [2]



Quais são os principais casos de uso de GenAI com maior retorno sobre investimento (ROI)?



25% Negociação e otimização de portfólio

21% Experiência e engajamento do cliente

11% Processamento de documentos

11% Geração de relatórios

[2] Fonte: NVIDIA – State of AI in Financial Services: 2025 Trends.

Além dos ganhos observados com GenAI, a evolução para soluções baseadas em agentes inteligentes está abrindo novas e poderosas oportunidades. Esses agentes remodelam o trabalho em três frentes principais: redução de custos, ao simplificar operações e eliminar etapas desnecessárias; diminuição de lead time, eliminando gargalos humanos e automatizando processos; e melhoria na qualidade do serviço, por meio do uso otimizado de ferramentas e ambientes de trabalho. Exemplos concretos reforçam esse impacto: redução de 25% no tempo de ciclos de R&D, economia de até 10 vezes em operações de call center, aumento de até 50 vezes na velocidade de produção de conteúdo de marketing e impulsos de até 40% na produtividade de equipes de TI. Esses resultados comprovam que a adoção estratégica de agentes é um passo decisivo para gerar valor sustentável e mensurável no setor financeiro.

QUAIS (MOTORES DE VALOR) PODEM SER ESPERADOS?

Agentes estão transformando a forma como concebemos o trabalho e desbloqueando 3 oportunidades de valor:



Redução de Custos

Racionalização e Redução



Redução de Tempo de Execução

Otimização de processos e eliminação de gargalos humanos



Qualidade de Serviço

Uso otimizado de diferentes ferramentas e ambientes

ALGUNS EXEMPLOS

[3]

25%

de redução no tempo dos ciclos de P&D

X10

redução de custos em call centers com Agentes

X50

aumento na velocidade de criação de posts de blog de Marketing

40%

de aumento de produtividade no departamento de TI

[3] Fonte: Artefact – Conhecimento interno aplicado a projetos no setor financeiro.

Historicamente, o setor financeiro foi construído sobre um pilar essencial: a confiança. Em um ambiente altamente regulado e globalmente interconectado, essa confiança sempre esteve vinculada à gestão estratégica dos dados dos clientes. Contudo, a dinâmica mudou. A lealdade, antes garantida por um mercado menos centrado no consumidor, agora precisa ser conquistada diariamente.

Nesse novo cenário, a GenAI, aliada à Agentic AI, surge como uma ferramenta fundamental para fortalecer a confiança, otimizar operações e reinventar a experiência do cliente. Seja na personalização de serviços, no aprimoramento da segurança ou na eficiência do back office, a adoção estratégica dessas tecnologias está redefinindo o futuro do setor financeiro.

O estado da arte da tecnologia no setor financeiro

As instituições financeiras estão apenas começando a explorar o verdadeiro potencial da AI Generativa (GenAI), uma tecnologia que vai além do processamento de informações, permitindo a criação de conteúdo genuinamente novo a partir de comandos simples. Diferente das revoluções tecnológicas anteriores, onde o ser humano sempre foi o agente de transformação – seja operando máquinas, dirigindo colheitadeiras ou programando sistemas –, agora a tecnologia assume o papel de criador. Impulsionada por Modelos de Linguagem de Grande Escala (LLMs), a GenAI pode gerar insights, elaborar estratégias, produzir

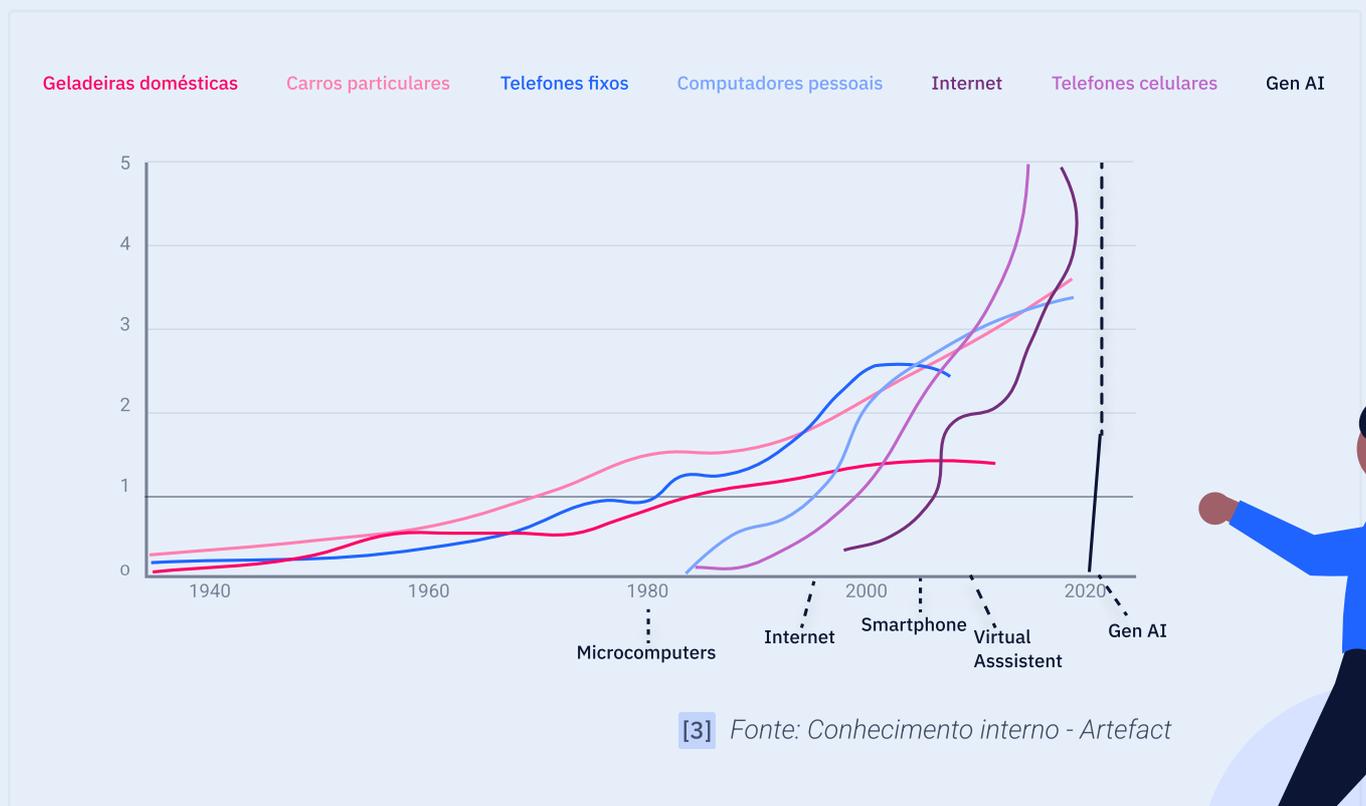
relatórios e até desenvolver código sem a necessidade de intervenção humana ativa. Esse é o verdadeiro divisor de águas: a automação não se limita mais a tarefas repetitivas, mas expande as fronteiras da criatividade e da tomada de decisão, inaugurando um novo paradigma para o setor financeiro. [4]



[4] Adaptado de análise interna – Artefact. Generative AI Survey (documento institucional).

Os LLMs são modelos avançados de inteligência artificial, treinados com grandes volumes de dados textuais, capazes de compreender e gerar linguagem natural com alta precisão. Após adaptações específicas, esses modelos podem ser usados para interação, criação de conteúdo e automação de processos textuais. No ecossistema da GenAI, existem também modelos voltados para outros formatos de dados, como imagens, vídeos e áudios, além dos Modelos Multimodais, que integram diferentes tipos de input simultaneamente, ampliando ainda mais as possibilidades de aplicação.

Assim como o microcomputador, a Internet e o smartphone transformaram nossa relação com a tecnologia, a GenAI está redefinindo a interface entre humanos e máquinas, tornando a interação mais acessível, intuitiva e produtiva. Instituições que adotam essa tecnologia de forma estratégica já estão colhendo benefícios, como ganhos de produtividade, redução de custos e maior agilidade na inovação. Os players que avançam rapidamente na implementação da GenAI estão conquistando uma vantagem competitiva significativa, enquanto outros, que ainda não compreenderam a urgência dessa transformação, ficam para trás. [3]



Mas a sofisticação dessa tecnologia vai além da simples geração de conteúdo. O próximo passo da AI no setor financeiro envolve sistemas mais autônomos e inteligentes, capazes de agir de maneira estruturada e estratégica.

Esse avanço se reflete no conceito de Agentic AI, que amplia as capacidades da GenAI ao permitir que modelos não apenas gerem respostas, mas tomem decisões informadas, realizem tarefas de forma contínua e se adaptem dinamicamente ao contexto. Seja interagindo com o mundo real, acessando informações recentes ou utilizando APIs e ferramentas além dos dados de treinamento, a Agentic AI inaugura uma nova era de autonomia e eficiência.

O esquema a seguir ilustra a transformação significativa que o uso de Agentic AI pode trazer para o processo de seleção de um vencedor de uma RFP (Request for Proposal), comparando o modelo tradicional com o modelo otimizado por agentes inteligentes. [3]

A análise e validação de RFPs (Requests for Proposal) ainda é, em muitos bancos, um processo operacionalmente intenso, descentralizado e fortemente dependente de tarefas manuais. Tradicionalmente, esse processo pode levar até quatro semanas, envolvendo múltiplas interfaces, vários profissionais e um fluxo de trabalho fragmentado. Desde o download e leitura de documentos em diversos formatos até a checagem de conformidade contratual, agrupamento por taxonomia e, finalmente, a decisão sobre o fornecedor vencedor, cada etapa exige esforço humano considerável — especialmente quando realizada em grande escala. Esse modelo tradicional é caracterizado por baixa eficiência, alta carga operacional e pouca integração entre as etapas.



Diferença da análise e validação de uma RFP com ou **sem** agentes



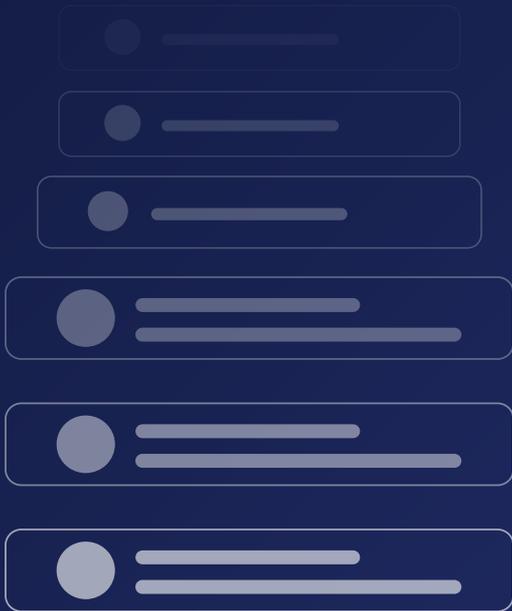
Com a introdução de agentes autônomos de AI, esse cenário se transforma radicalmente. O uso de Agentic AI permite consolidar o processo em uma única estação de trabalho, centralizada no comprador, e reduzir o tempo total de execução para apenas um dia. Agentes especializados automatizam a ingestão e extração de dados a partir de documentos em PDF ou Excel, realizam análises de conformidade com contratos mestres, comparam respostas a benchmarks pré-estabelecidos e propõem recomendações finais com embasamento lógico.

Ao invés de múltiplos profissionais executando tarefas repetitivas e suscetíveis a erro, a atuação humana passa a se concentrar apenas na etapa final de validação – com foco em análise crítica e tomada de decisão. O ganho de eficiência operacional é evidente, mas o impacto estratégico vai além: ao liberar tempo e capacidade cognitiva dos times, bancos passam a ter mais agilidade para responder a oportunidades de mercado, reduzir riscos e tomar decisões mais informadas.

Esta transformação ilustra com clareza o valor da Agentic AI: mais do que automatizar tarefas, ela reposiciona o papel do humano nas operações financeiras, elevando a qualidade do processo decisório e acelerando o ritmo da inovação.

CAPÍTULO 2

Agentic AI & Oportunidades



O que é **Agentic AI** e por que ela representa uma **oportunidade** tão grande?

A transformação que vimos no processo de RFP é apenas uma entre centenas de aplicações possíveis de uma nova geração de inteligência artificial: a Agentic AI. Se a AI tradicional automatizou tarefas específicas e a GenAI trouxe criatividade e geração de conteúdo sob demanda, a Agentic AI leva essa revolução um passo adiante — com agentes capazes de tomar decisões, agir com autonomia e se adaptar em tempo real ao ambiente.

Não se trata mais de responder a perguntas ou gerar textos com base em comandos: trata-se de executar missões complexas de ponta a ponta, conectando sistemas, acessando dados dinâmicos, operando ferramentas e aprendendo com cada interação.

É essa capacidade de operar com propósito, contexto e coordenação que torna a Agentic AI uma oportunidade estratégica gigantesca para o setor financeiro. E embora o conceito possa parecer novo, sua evolução já está em curso — com marcos técnicos que avançam ano após ano e nos aproximam rapidamente de um cenário em que agentes inteligentes poderão atuar como verdadeiros colaboradores digitais. A linha do tempo a seguir mostra como essa transformação está acontecendo em ritmo acelerado.



"Agentes" evoluíram rapidamente, e suas novas capacidades **abriram caminho para oportunidades transformadoras de automação de processos.**

2023: Assistir

2024: Agir

2025: Automatizar



Tipo de agentes	Agentes de recuperação	Agentes de ação	Agentes de orquestração	Agentes de percepção
Capacidades tecnológicas	RAG	Uso de ferramentas (APIs)	Colaboração entre múltiplos agentes / humanos	Uso de interface gráfica (GUI)
Observabilidade	Monitoramento	Avaliação	Supervisão	

Projetar soluções com GenAI	O agente recebe o contexto para responder	O agente pode planejar e executar ações com ferramentas conforme necessário	Agentes especializados trabalham juntos e escalam para humanos quando necessário	Agentes podem interagir com interfaces gráficas sem necessidade de APIs
-----------------------------	---	---	--	---



Essa jornada de evolução pode ser compreendida a partir de três grandes marcos:



Assistir



Agir



Automatizar



Cada fase representa uma ampliação concreta das capacidades dos agentes de AI e das oportunidades que elas desbloqueiam para o setor financeiro:



2023: Assistir

Neste primeiro estágio, surgem os retrieval agents, especializados em recuperar informações relevantes a partir de contextos fornecidos. Com suporte de tecnologias como RAG (retrieval-augmented generation) e ferramentas como LangChain, esses agentes são capazes de buscar dados, consolidar informações e responder com precisão. A observabilidade nesse momento ainda está restrita ao monitoramento básico da performance dos modelos.



2024: Agir

Com os avanços no uso de APIs, emergem os acting agents, capazes de executar ações concretas em ferramentas externas. Essa evolução marca uma virada operacional: os agentes não apenas respondem, mas também agem com base no que interpretam, integrando-se a sistemas como CRMs, ERPs e plataformas analíticas. Nesse mesmo ano, surgem os orchestration agents, que orquestram múltiplos agentes ou colaboram com humanos.



2025: Automatizar

- Em 2025, os agentes evoluem para interações mais complexas com sistemas, dando origem aos perception agents. A principal diferença está na capacidade desses agentes de navegar por interfaces gráficas (GUIs), sem depender exclusivamente de APIs para executar ações. Isso permite automatizar fluxos que envolvem múltiplas aplicações ou sistemas legados – algo especialmente relevante em instituições financeiras com arquiteturas de TI mais heterogêneas. Tecnologias como Runner H e Claude viabilizam esse avanço, enquanto mecanismos de supervisão ativa garantem controle e segurança operacional.

Esse roadmap tecnológico demonstra com clareza que a Agentic AI não é uma tendência distante. Ela já está moldando novas formas de operar, tomar decisões e gerar valor. Para líderes do setor financeiro, entender essas fases é essencial para estruturar estratégias robustas, identificar casos de uso prioritários e garantir uma adoção escalável e segura dessa tecnologia transformadora.

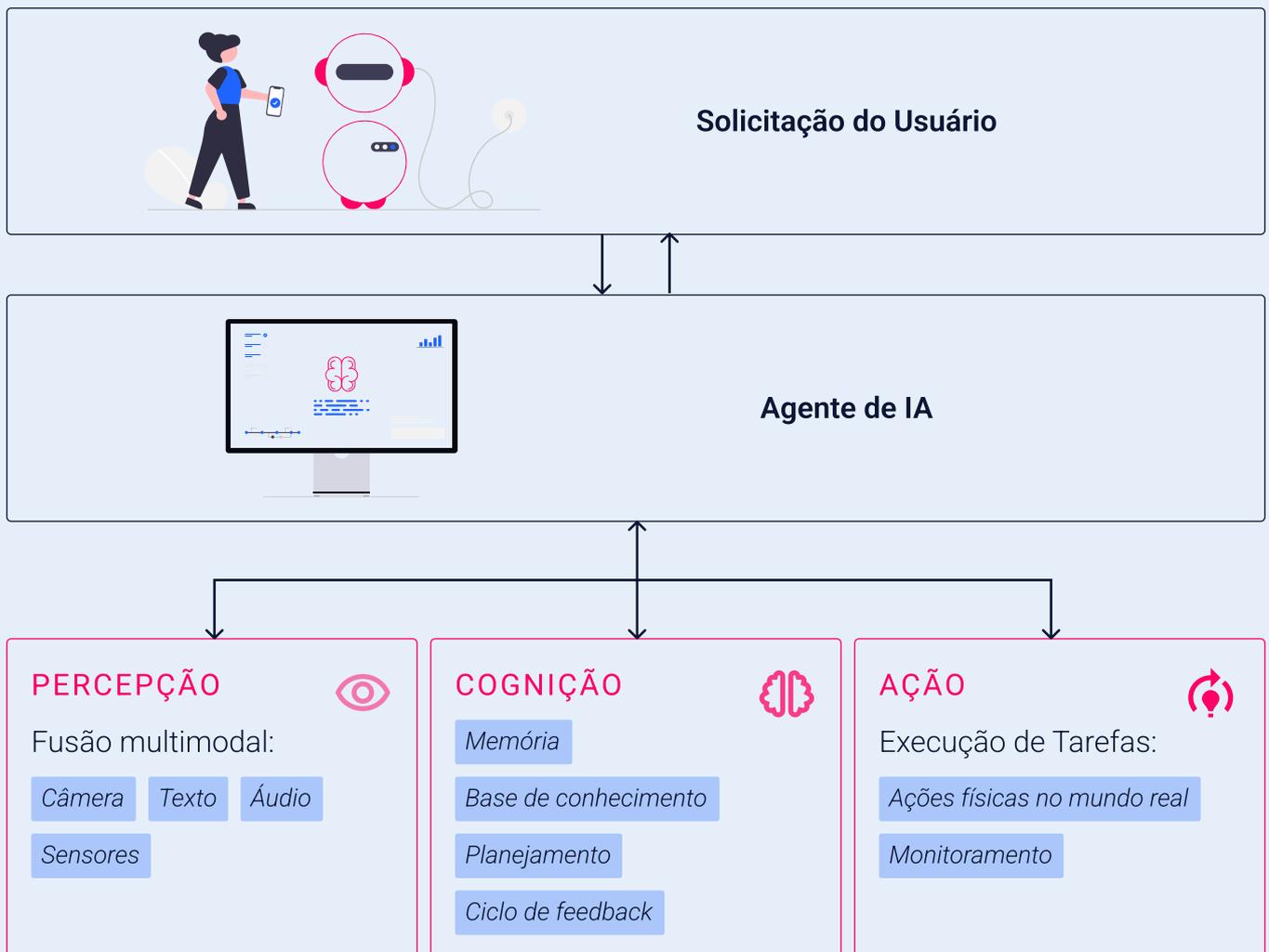
Para compreender o real potencial da Agentic AI, é fundamental entender como esses agentes operam na prática. Diferente de soluções baseadas apenas em grandes modelos de linguagem (LLMs), os agentes são entidades autônomas compostas por três componentes principais:

- ☑ **Percepção;**
- ☑ **Cognição;**
- ☑ **Ação.**

Cada um desses blocos tem um papel essencial na forma como o agente interage com o ambiente, toma decisões e executa tarefas de forma contínua e adaptável.

AGENTE, NOVA DEFINIÇÃO

Uma entidade autônoma que percebe, raciocina, age e se adapta às mudanças.



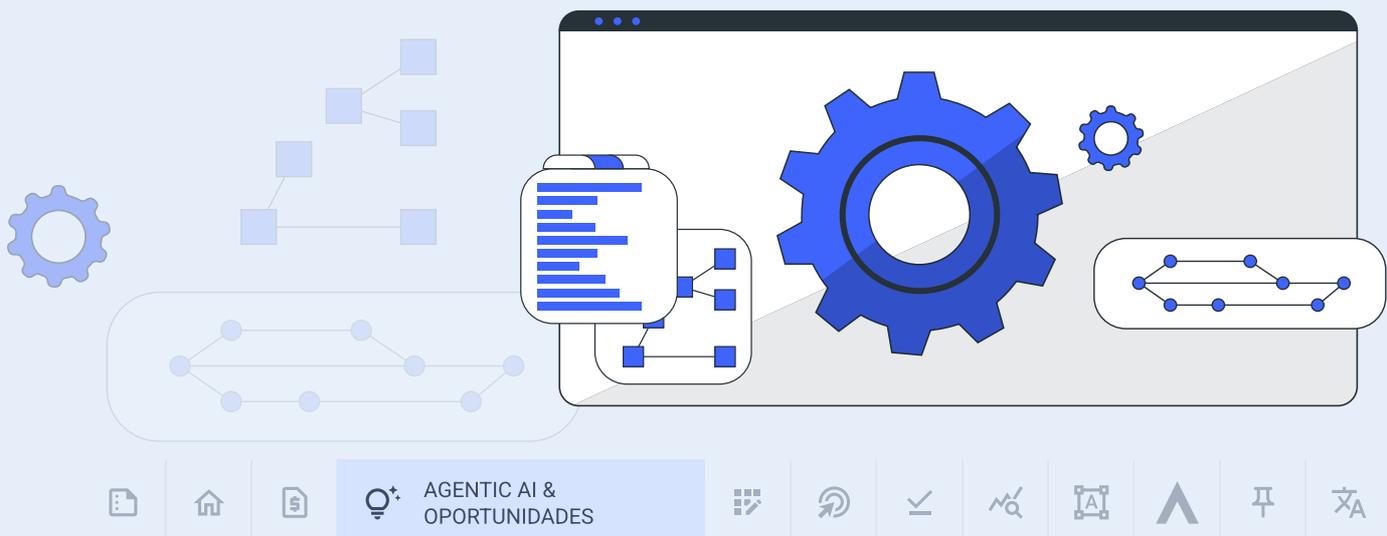
A percepção permite que o agente colete informações de diferentes fontes, por meio da chamada fusão multimodal. Isso inclui dados de texto, áudio, imagens (câmeras) e sensores diversos. Em contextos empresariais, por exemplo, isso significa que um agente pode cruzar informações estruturadas de sistemas (como ERPs e CRMs) com dados não estruturados (como e-mails, gravações de atendimento e documentos), criando uma visão rica e contextualizada da situação.

Com base nos dados percebidos, o componente de cognição entra em ação. Ele combina memória, bases de conhecimento e capacidade de planejamento para interpretar os sinais recebidos, tomar decisões e se adaptar. Um aspecto fundamental aqui é o ciclo de feedback (feedback loop), que permite ao agente aprender com suas próprias ações e refinar seu comportamento ao longo do tempo – essencial para lidar com ambientes complexos, como o mercado financeiro.

Por fim, o agente é capaz de executar tarefas com base no raciocínio desenvolvido. Isso pode incluir desde interações com sistemas via API, navegação em interfaces gráficas, até ações físicas em ambientes conectados. Além disso, o agente mantém a capacidade de monitorar os resultados de suas ações, garantindo uma operação mais segura, eficiente e auditável.

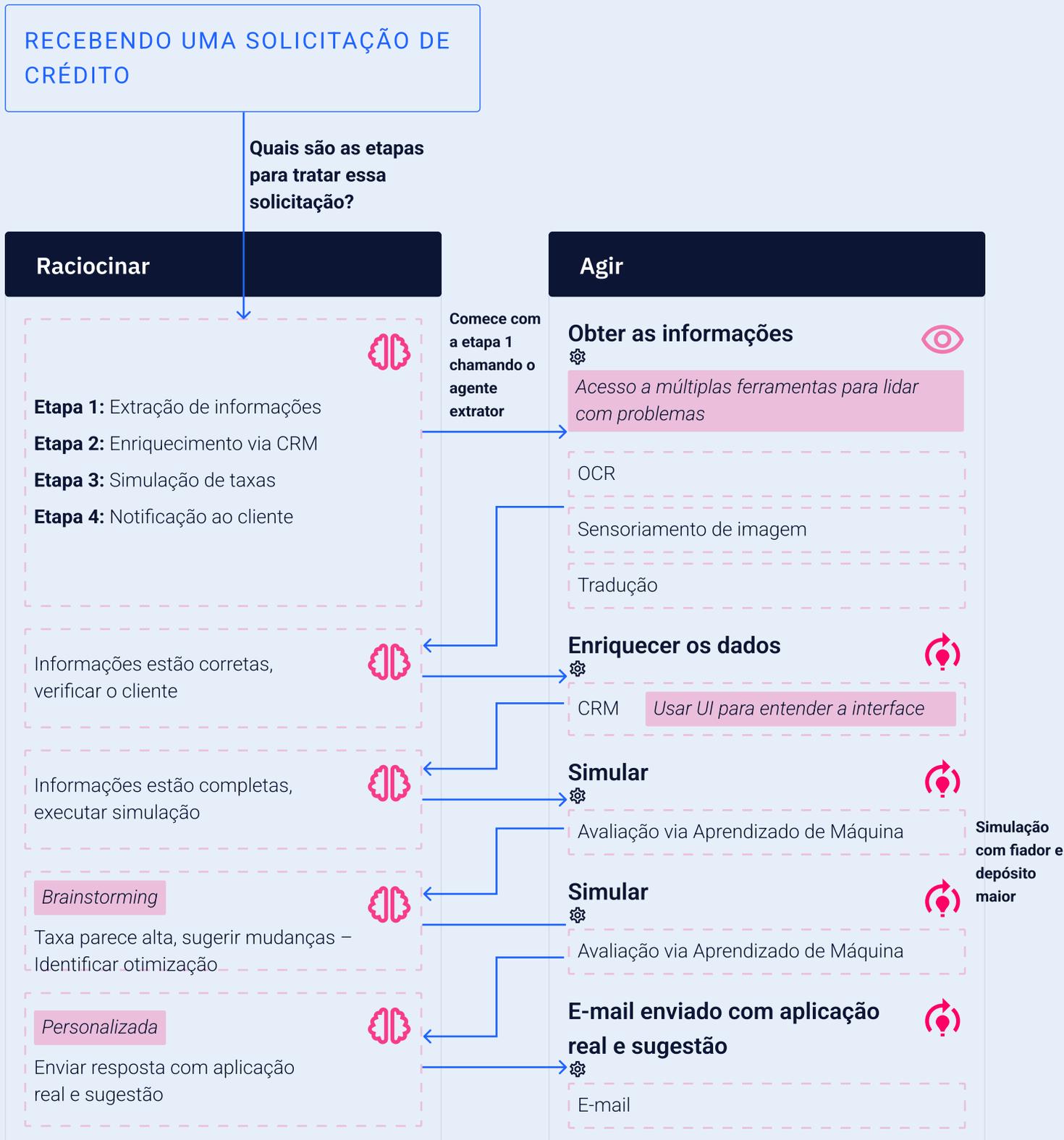
Essa arquitetura permite que agentes ajam de forma proativa, respondam a mudanças em tempo real e evoluam com base na experiência – o que os torna uma das inovações mais promissoras da AI moderna.

A Agentic AI permite automatizar fluxos de trabalho complexos de ponta a ponta por meio da coordenação entre múltiplos agentes. Um exemplo claro e aplicável ao setor financeiro é o processamento automatizado de pedidos de crédito.



Nesse fluxo, os agentes atuam de forma orquestrada, combinando raciocínio (Reason) e execução (Act) em uma sequência lógica de etapas, como mostra o diagrama a seguir:

Ao criar uma cadeia de agentes, podemos automatizar todo um fluxo de trabalho de ponta a ponta com React (Raciocinar & Agir)



Ao receber uma solicitação de crédito, inicia-se uma cadeia de agentes, cada um responsável por uma parte do processo:



1. Extração da informação

Um agente extrai automaticamente dados relevantes do pedido utilizando OCR, tradução e interpretação de interfaces. Ele transforma inputs brutos em informações estruturadas.



2. Enriquecimento de dados (CRM)

Em seguida, outro agente valida e enriquece esses dados com informações do CRM, agregando contexto ao perfil do cliente.



3. Simulação de taxas e condições

Com os dados prontos, um terceiro agente executa uma simulação de crédito baseada em modelos de machine learning para prever a taxa ideal.



4. Otimização e personalização da oferta

Caso a taxa não seja ideal, o sistema identifica alternativas (como a exigência de um fiador ou maior entrada), simula novamente e otimiza a proposta de forma personalizada.



5. Resposta ao cliente

Por fim, um agente elabora e envia automaticamente a resposta ao cliente, contendo a simulação, a análise e a sugestão recomendada.

Esse exemplo mostra como cadeias de agentes podem replicar processos cognitivos, mas com velocidade, precisão e escalabilidade superiores. Cada agente entende seu papel, consulta sistemas, toma decisões com base em regras ou modelos, aprende com os resultados e colabora com os demais para cumprir o objetivo final.

No entanto, para que essa sofisticação seja implementada de maneira segura e eficaz, algumas considerações são cruciais. É indispensável estabelecer frameworks de governança robustos para assegurar que os agentes atuem dentro de limites éticos e regulamentares, prevenindo erros e violações de dados, bem como a construção de “guard rails” que impõem limites ao que é permitido aos agentes e previnem alucinação. Ademais, a integração de agentes aumenta a superfície de ataque cibernético, exigindo soluções como autenticação contextual e monitoramento contínuo. Apesar de sua autonomia, os agentes também requerem supervisão humana em decisões de alto impacto, garantindo que suas ações estejam alinhadas com os objetivos organizacionais.

Ao equilibrar autonomia, segurança e governança, o uso de agentes no setor financeiro promete inaugurar uma nova era de eficiência e inovação, transformando profundamente o modo como os serviços financeiros são oferecidos e consumidos.

De acordo com o Citi GPS, a Agentic AI marca uma mudança em direção à economia do “faça isso por mim”, em que a tecnologia assume a tomada de decisões e a execução de tarefas. [6] A maior autonomia da Agentic AI, em relação à GenAI, permite que ela lide com processos repetitivos e com uso intensivo de dados, otimizando os fluxos de trabalho, aprimorando a conformidade e melhorando a tomada de decisões. Na imagem em seguida fica mais explícito o que significa a transição entre GenAI e Agentic AI e possíveis aplicações no setor de serviços financeiros. [5]

[5] Extraído de: Citi GPS. *Agentic AI: Finance & the 'Do It For Me' Economy*.

[6] Adaptado de artigo: *InsiderFinance Wire – Sahaj Godhani, Medium*.



No setor financeiro, Agentic AI significa, por exemplo, assistentes inteligentes que identificam riscos, ajustam portfólios, automatizam processos complexos e até gerenciam interações com clientes sem a necessidade de intervenção humana.

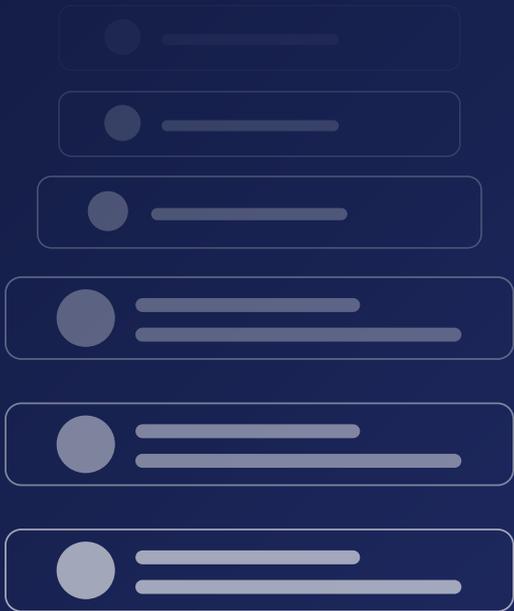
Com essa sofisticação tecnológica, as instituições podem escalar operações, oferecer personalização em massa e aumentar significativamente sua eficiência, mantendo a confiança do cliente em um mercado cada vez mais dinâmico.

O futuro do setor não será definido apenas por quem adotar a AI, mas por quem souber explorar sua máxima sofisticação e inteligência estratégica.



CAPÍTULO 3

GenAI e Agentic AI e suas Aplicações



GenAI e Agentic AI: aplicações e a Arte do Possível

No setor financeiro, a adoção de soluções baseadas em GenAI tem ganhado tração acelerada, com empresas buscando explorar novos casos de uso que impulsionam eficiência, personalização e inovação. Conforme evidenciado na imagem [1], houve um salto expressivo de 2023 para 2024 em diversas frentes. Destacam-se áreas como Experiência do Cliente, que evoluiu de 25% para 60%, e Geração de Relatórios e Pesquisa, que saltou de 27% para 53%. Outros casos, como Processamento de Documentos, Assistência na construção de códigos, e Busca Empresarial, também avançaram rapidamente, refletindo a crescente aposta do mercado financeiro no potencial transformador da GenAI.

Principais Casos de Uso de GenAI

2023 2024



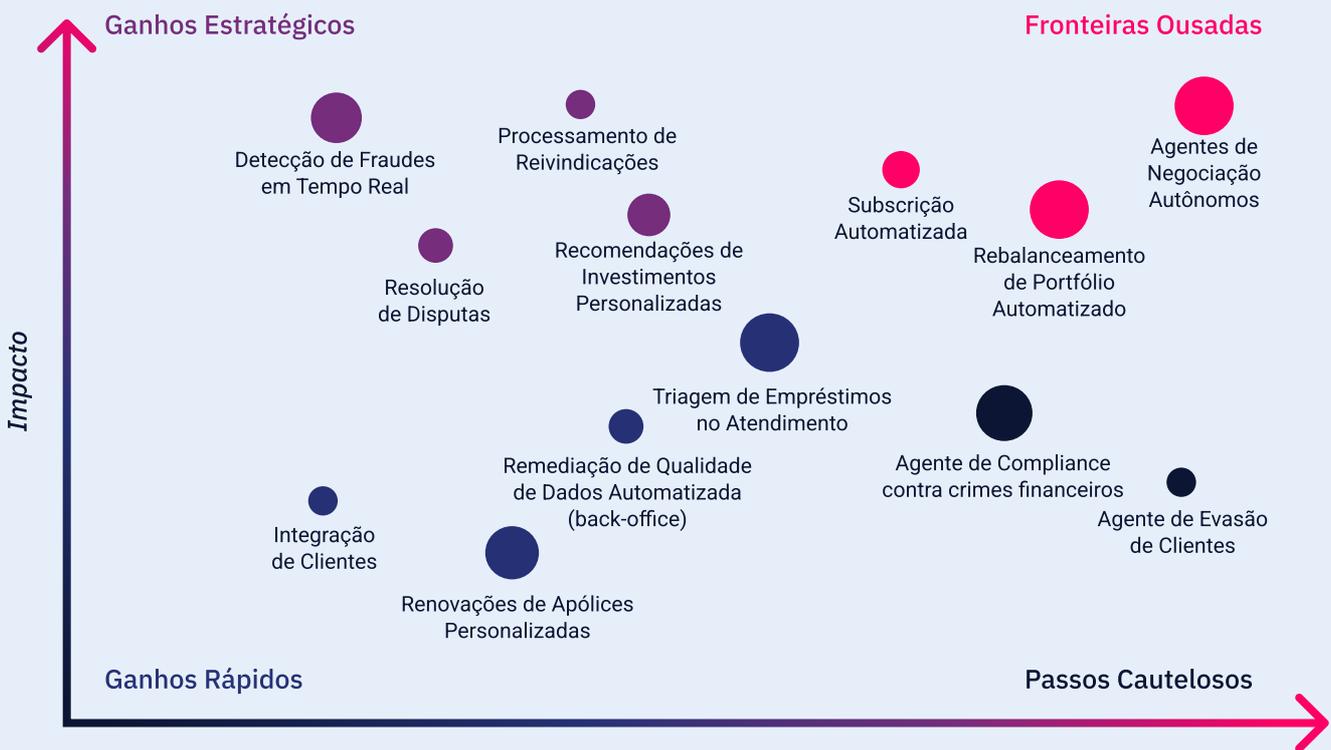
[2] Fonte: State of AI in Financial Services: 2025 Trends - NVIDIA

As aplicações de GenAI no setor financeiro já são uma realidade, como mostra o estudo. Além disso, essa tecnologia cria as bases para que uma série de casos de uso ainda mais transformacionais possam ser desenvolvidos, através de Agentic AI.

Para guiar a adoção estratégica de soluções com Agentic AI nesse setor, é essencial considerar três dimensões centrais: risco, impacto e complexidade. A imagem a seguir ilustra essa lógica, em que o eixo vertical representa o impacto potencial da solução, enquanto o eixo horizontal indica o grau de risco envolvido, por fim o tamanho da bolha reflete o esforço necessário para construir o agente.

Seu roadmap de Agentic AI será refinado por uma combinação de risco, impacto e complexidade

Tamanho da bolha = grau de esforço



[3] Fonte: Conhecimento interno - Artefact

Grau de Risco

Nesse cenário, dois casos se destacam como *Ganhos Rápidos*. A *Integração de Clientes* permite automatizar a jornada inicial do cliente com agentes que coletam documentos, validam informações, integram sistemas (como CRMs e antifraude) e notificam o time interno ou o próprio cliente. Já as *Renovações de Apólices Personalizadas* faz com que agentes revisem históricos de apólices, para sugerir renovações personalizadas com base em perfis e enviar comunicações proativas ao cliente.

Na camada de *Ganhos Estratégicos*, destacam-se aplicações que exigem um esforço moderado, mas entregam alto impacto para o negócio. Casos como *Detecção de Fraudes em Tempo Real*, *Resolução de Disputas* e *Processamento de Reivindicações* otimizam processos críticos com ganhos significativos em velocidade e precisão, sem o nível de risco das fronteiras mais avançadas.

Já nas *Fronteiras Ousadas*, estão os casos de uso que combinam alto impacto com maior risco e complexidade – como *Agentes de Negociação Autônomos*, *Rebalanceamento de Portfólio Automatizado* e *Subscrição Automatizada*. Essas aplicações representam o futuro da inteligência autônoma no setor, exigindo arquiteturas robustas e governança sofisticada.

Para facilitar a priorização e estruturação de iniciativas com Agentic AI, é útil organizar os casos de uso em grandes frentes de valor.

A seguir, apresentamos uma visão estruturada que agrupa os principais casos em quatro categorias-chave:

- 🔘 **Gestão de Riscos**
- 🔘 **Automação de Processos**
- 🔘 **Atendimento ao Cliente e Personalização**
- 🔘 **Tomada de Decisão e Insights**

Para cada caso, é indicado onde ele gera mais valor, com marcações de aderência total ou parcial. Isso permite visualizar rapidamente como cada iniciativa se conecta a objetivos estratégicos e operacionais, ajudando na priorização e aplicação de soluções baseadas em Agentic AI.

Agentes de Gestão de Riscos são voltados à mitigação de riscos financeiros, operacionais e regulatórios, com capacidade de atuação contínua e aprendizado constante. Automação de Processos engloba agentes que otimizam tarefas rotineiras ou operacionais, reduzindo erros e aumentando a eficiência com fluxos mais responsivos e fluidos. Já Atendimento ao Cliente e Personalização inclui agentes que interagem com canais e pessoas para oferecer experiências personalizadas e integradas em tempo real. Por fim, Tomada de Decisão e Insights cobre agentes focados em análise avançada e suporte à decisão, permitindo que a inteligência artificial atue de forma ativa na construção de estratégias e na operação do negócio.



Casos de Uso	Gestão de Riscos	Automação de Processos	Atendimento ao Cliente e Personalização	Tomada de Decisão e Insights
Detecção de Fraudes em Tempo Real	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Agente de Compliance contra crimes financeiros	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Triagem de Empréstimos no Atendimento ao Cliente	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Processamento de Reivindicações (ou Sinistros)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/> <small>Com feedback loops</small>
Subscrição Automatizada	<input checked="" type="checkbox"/> <small>[parcial]</small>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Correção Automatizada de Qualidade de Dados (back-office)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Resolução de Disputas	<input checked="" type="checkbox"/> <small>[parcial]</small>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Integração de Clientes	<input type="checkbox"/>	<input checked="" type="checkbox"/> <small>[parcial]</small>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Renovações de Apólices Personalizadas	<input type="checkbox"/>	<input checked="" type="checkbox"/> <small>[parcial]</small>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Recomendações de Investimento Personalizadas	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Agente de Evasão de Clientes	<input checked="" type="checkbox"/> <small>[parcial]</small>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Rebalanceamento Automático de Portfólio	<input checked="" type="checkbox"/> <small>[parcial]</small>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Agentes de Negociação Autônomos	<input checked="" type="checkbox"/> <small>[alta complexidade]</small>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/> <small>[decisões autônomas]</small>

A categorização acima permite visualizar de forma clara onde cada caso de uso de Agentic AI pode gerar mais valor dentro das operações financeiras. Para aprofundar a análise, a seguir detalhamos cada uma das quatro frentes principais, trazendo exemplos concretos de agentes, suas funções e o impacto esperado. Essa visão ajuda a tangibilizar como os agentes operam na prática, evidenciando os ganhos de eficiência, mitigação de riscos e personalização que eles podem trazer para diferentes áreas do setor.

Gestão de Riscos

Casos-chave



Detecção de Fraudes em Tempo Real

Agente que monitora transações financeiras em tempo real, identificando padrões suspeitos de fraude (como comportamento fora do perfil do cliente) e acionando imediatamente bloqueios ou investigações automatizadas.



Agente de Compliance contra crimes financeiros

Agente especializado em compliance que verifica continuamente a conformidade de transações com normas como AML (anti-lavagem de dinheiro) e KYC (conheça seu cliente), cruzando dados com listas de sanções e perfis de risco.



Triagem de Empréstimos no Atendimento ao Cliente

Agente que realiza a análise preliminar de crédito ou empréstimo no front office, combinando dados estruturados e não estruturados para calcular scores, validar documentação e sugerir decisões iniciais.

Esses casos têm alto impacto e envolvem risco considerável, sendo áreas clássicas de uso de inteligência artificial para mitigação de riscos financeiros e operacionais. Com Agentic AI, é possível:



Criar agentes que monitoram transações em tempo real, detectando padrões anômalos e acionando workflows de investigação automática ou semiautomática.



Habilitar compliance contínuo com agentes que verificam aderência a políticas e regulamentos.



Automatizar o pré-screening de clientes ou empréstimos com base em scores de risco, histórico e comportamento.

Automação de **Processos**

Casos-chave



Processamento de Reivindicações

Agente responsável por receber, classificar, validar e encaminhar sinistros automaticamente, utilizando reconhecimento de documentos, regras de negócio e dados históricos.



Subscrição Automatizada

Agente que analisa automaticamente o perfil do proponente, consulta bases externas, aplica regras atuariais e fornece uma decisão inicial de aceitação ou recusa do risco.



Correção Automatizada de Qualidade de Dados *(back-office)*

Agente focado na identificação e correção de problemas de qualidade de dados (DQ), como dados inconsistentes ou ausentes, mantendo a integridade das bases e alimentando corretamente os sistemas analíticos.



Resolução de Disputas

Agente que atua na mediação e resolução de disputas (por exemplo, em pagamentos ou transações bancárias), organizando evidências, avaliando a situação e sugerindo soluções ou encaminhamentos

Esses casos se beneficiam diretamente da automação inteligente de etapas repetitivas, otimizando tempo e custos. A proposta com Agentic AI é:



Desenvolver agentes especialistas que executam tarefas como validação de documentos, cruzamento de dados e aplicação de regras de negócios.



Incorporar lógica adaptativa, permitindo que os agentes aprendam exceções e melhorem continuamente.



Integrar esses agentes com sistemas BPM (Business Process Management) ou ERPs para orquestração ponta-a-ponta do processo.

Atendimento ao Cliente e Personalização

Casos-chave



Agente de Evasão de Clientes

Agente que monitora comportamentos de clientes (como queda no uso de produtos, reclamações ou mudanças de padrão) para prever o risco de churn e sugerir ações de retenção personalizadas.



Integração de Clientes

Agente que conduz o processo de entrada de novos clientes de forma adaptativa, solicitando documentos conforme o perfil, explicando etapas e integrando dados automaticamente em sistemas internos.



Renovações de Apólices Personalizadas

Agente que avalia o histórico do segurado, mudanças de perfil e contexto atual para oferecer renovações personalizadas e proativas, ajustando coberturas e preços conforme necessidade.



Recomendações de Investimento Personalizadas

Agente que sugere investimentos personalizados com base em perfil de risco, objetivos, preferências e eventos de mercado, podendo inclusive interagir em linguagem natural com o usuário.

Esses casos representam iniciativas de personalização e engajamento proativo do cliente. Usando Agentic AI:



É possível criar agentes que identificam sinais precoces de churn e iniciam ações de mitigação, como ofertas e contato personalizado.



Agentes de onboarding tornam o processo mais fluido, adaptando o fluxo às necessidades do cliente com base em perfis e comportamento.



Agentes de recomendação constroem jornadas personalizadas com base em dados históricos, preferências e contexto atual do usuário.

Tomada de Decisão e Insights

Casos-chave



Rebalanceamento Automático de Portfólio

Agente que avalia carteiras de investimento ou produtos e, com base em critérios predefinidos (como perfil de risco ou eventos externos), propõe e executa realocações automáticas.



Recomendações de Investimento Personalizadas

(Também presente em Personalização) — aqui o foco é na explicabilidade e geração de insights em tempo real para auxiliar o investidor na tomada de decisão autônoma.



Agentes de Negociação Autônomos

Nesse contexto, o agente de sinistros coleta feedback contínuo das decisões e outcomes, aprendendo com erros e acertos para refinar suas futuras decisões.

Esses casos focam mais melhorar a capacidade de tomada de decisões mais embasadas, assertivas e eficientes através da construção de agentes que se organizam em um workflow para levantar todos os pontos necessários. Sendo assim, com Agentic AI:



Agentes financeiros podem reavaliar portfólios automaticamente, com base em eventos de mercado e perfis de risco.



Em sinergia com ferramentas analíticas, agentes oferecem recomendações com explicações (XAI), aumentando a confiança dos usuários.



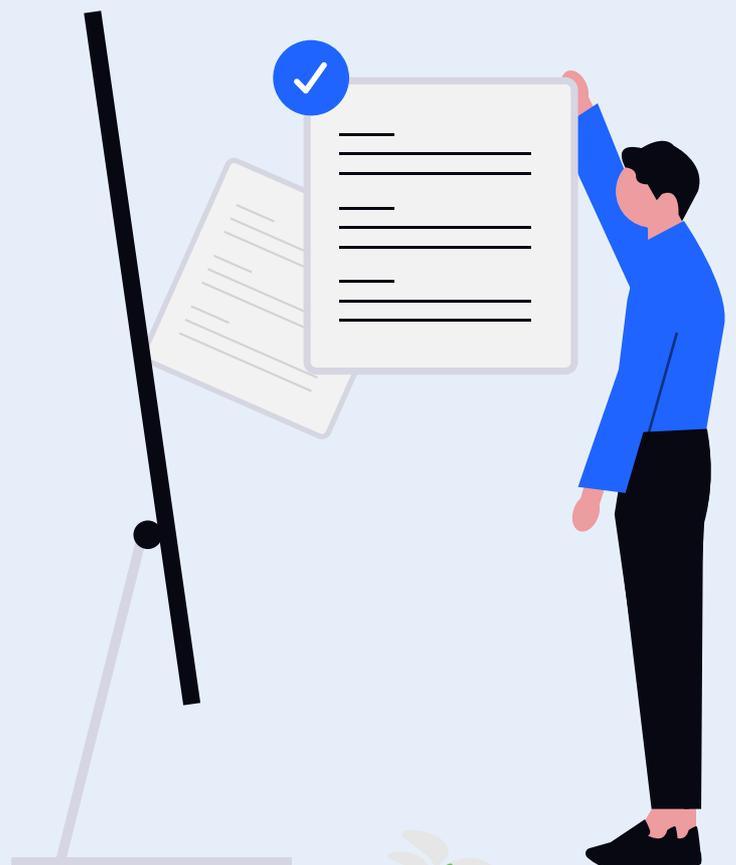
Agentes podem atuar como copilotos de decisão, preparando insights, simulações e até sugerindo ações dentro de um fluxo decisório.

À medida que os casos de uso de Agentic AI se multiplicam, uma pergunta estratégica ganha força: como transformar essas iniciativas em soluções escaláveis e integradas ao dia a dia da organização? A resposta está na construção de produtos digitais inteligentes, em que agentes são projetados para operar dentro de workflows reais, interagindo com sistemas, dados e pessoas. Essa abordagem garante que a AI deixe de ser apenas uma “camada de recomendação” para se tornar parte ativa da operação e da experiência oferecida ao cliente.

INTEGRAÇÃO COM WORKFLOWS DE NEGÓCIO VIA AGENTIC AI

A chave para extrair valor desses casos é tratá-los como produtos digitais compostos por agentes inteligentes, cada um integrado aos processos core da empresa. Isso envolve:

- Definir papéis específicos para os agentes em cada jornada (ex: avaliador de risco, analista de sinistros, recomendador de investimentos).
- Projetar interações entre agentes, humanos e sistemas, com governança e supervisão.
- Monitorar KPIs de performance e aprendizado contínuo dos agentes, como parte de um ciclo ágil de evolução do produto.



Dessa forma, na tabela a seguir temos exemplos de produtos digitais que combinam os casos de uso apresentados anteriormente com enfoque para os desafios mais específicos de quatro segmentos dentro de serviços financeiros: Banco de Varejo, Banco de Investimento, Seguros e Pagamentos. Dessa forma podemos ver com mais clareza como seria a aplicação de Agentic AI na construção de produtos digitais que atendem aos desafios específicos de cada negócio.

Categorias-chave para geração de valor

Categoria	1. Automação de Processos	2. Gestão de Riscos	3. Atendimento, Engajamento e Personalização	4. Tomada de Decisão e Insights Baseados em AI
Varejo	Analista de mesa de crédito virtual capaz de orquestrar e automatizar todo o processo de concessão de crédito de ponta a ponta.	Gestor de risco operacional virtual, que monitora transações, investiga perfis e responde a eventos em tempo real.	Concierge financeiro digital que entende segmentos e perfis, executa e recomenda produtos com base em contexto individual.	Agente autônomo de apoio à decisão, que coleta dados, gera análises, explica tendências, simula cenários e sugere ações.
Banco de Investimento	Deal Book Automation Agent que automatiza a coleta de dados, análise e geração de materiais para apoio a deals.	Risk Intelligence Agent, capaz de monitorar o mercado em tempo real, identificar exposições críticas e recomendar ações.	Relationship Manager Virtual que conhece o cliente, antecipa suas necessidades, responde de forma consultiva e executa tarefas.	Deal Hunter Virtual que analisa tendências, identifica oportunidades e sugere ideias de originação com explicações estruturadas.
Seguros	Claim advisor capaz de receber, analisar, classificar, decidir e comunicar o andamento de um processo de sinistro.	Risk Sentinel, capaz de monitorar sinais de risco em tempo real, identificar padrões anômalos, gerar alertas, simular impactos e sugerir ações mitigadoras.	Concierge digital 24/7, capaz de antecipar necessidades, traçar perfis, entender estilo de vida e personalizar recomendações de seguros.	Agente autônomo que usa dados históricos, tendências de mercado e comportamento de clientes para gerar insights e otimizar portfólio de seguros e produtos.
Pagamentos	Agente autônomo de conciliação financeira, que orchestra, automatiza, analisa e age sobre o processo de reconciliação e liquidação de pagamentos.	Risk Guardian, que atua como um sentinela digital, constantemente monitorando transações, detectando padrões de risco, agindo em tempo real.	PayBot Concierge combinando suporte técnico, orientação financeira e ofertas sob medida a partir dos padrões de transações.	Payments Strategist capaz de analisar grandes volumes de dados, gerar recomendações, simular cenários e interagir com executivos em linguagem natural.



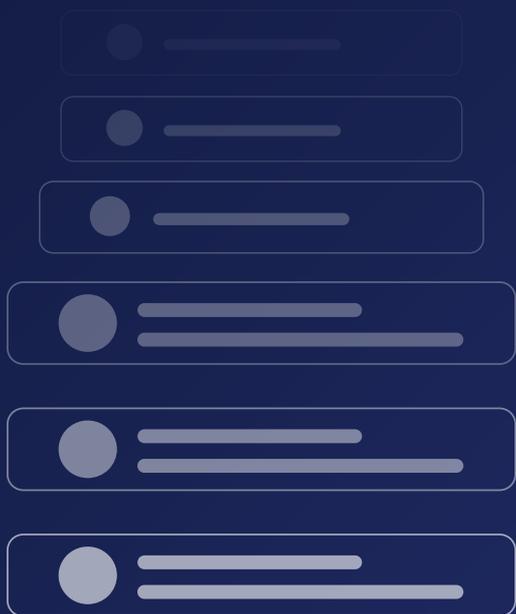
A Agentic AI eleva a construção de produtos digitais no setor financeiro ao permitir que agentes autônomos executem tarefas complexas – como monitorar riscos, personalizar ofertas ou automatizar processos – sempre alinhados aos objetivos de negócio. Diferente da AI tradicional, que entrega análises ou previsões, a Agentic AI transforma esses modelos em motores de ação contínua, ampliando seu impacto e acelerando a geração de valor em áreas como crédito, seguros e investimentos.

Nos próximos anos, espera-se que agentes inteligentes atuem de forma cada vez mais colaborativa e integrada, redefinindo como instituições financeiras orquestram dados, decisões e experiências para clientes e operações.



CAPÍTULO 4

Casos de Sucesso no Setor Financeiro



Casos de Sucesso: como a AI Está Redefinindo o Setor Financeiro

Dentre todas as aplicações possibilitadas pela evolução das ferramentas de GenAI e AgenticAI, a Artefact já desenvolveu implementações práticas para diversos clientes do setor financeiro em nível global. Esses casos de uso não são apenas projeções ou tendências futuras — são aplicações reais, já implementadas por grandes players do setor. Instituições líderes estão utilizando GenAI para transformar suas operações e interações com clientes, gerando resultados concretos em eficiência, personalização e inovação

OTIMIZANDO EXPERIÊNCIA DO CLIENTE: UM NOVO PARADIGMA EM ATENDIMENTO AO CONSUMIDOR

A transformação digital tem imposto às empresas de finanças desafios crescentes na gestão de suas interações com clientes. Em um mercado altamente competitivo, a capacidade de responder às dúvidas dos consumidores com velocidade, precisão e em uma linguagem que atenda a cada contexto dos consumidores tornou-se não apenas um diferencial estratégico, mas um fator determinante para a fidelização. Nesse contexto, a Artefact identificou uma oportunidade significativa: construção de um chatbot que substituísse a necessidade de interação humana constante, garantindo respostas precisas, seguras e em tempo hábil em um grande banco de investimentos europeu.



Desafios e Obstáculos

Apesar de promissor, o projeto enfrenta barreiras bem definidas em sua implementação, independente da estrutura da companhia. A primeira delas é atender à crescente expectativa dos usuários por respostas acessíveis, rápidas e de qualidade, um elemento crítico para a satisfação do cliente. Paralelamente, a solução precisa estar em conformidade com padrões rigorosos de segurança, que incluem não apenas a proteção de dados sensíveis, mas também a aderência às regulamentações específicas do setor financeiro.

Desenvolvendo a solução

A solução proposta se baseia em um modelo de GenAI projetado para operar com alta eficiência e segurança. O modelo se ancora em uma base de conhecimento de perguntas e respostas (Q&A), estruturada a partir de dados relevantes da organização. A solução também utiliza algoritmos de filtragem avançados para interpretar a pergunta do usuário, extrair as informações mais pertinentes da base de conhecimento e gerar respostas contextualizadas e precisas. Complementar às estruturas necessárias para o modelo, destaca-se a capacidade de adaptar a solução proposta tanto em ambientes de nuvem (Azure, AWS e GCP) quanto on-premise, preservando a integridade dos dados e alinhando-se às exigências de segurança.



INTERFACE DO USUÁRIO



Adaptado a partir do Artefact One Pager – Agente de Business Intelligence Aumentado aplicado ao setor farmacêutico

Benefícios Alcançados

Como resultados decorrente da implementação do bot de conversas em diferentes cenários, nossos clientes alcançaram uma redução em 90% do tempo médio de resposta aos usuários, promovendo um atendimento mais ágil e satisfatório. Em termos financeiros, o projeto conseguiu gerar uma economia superior a 5 milhões de euros no custo operacional do time de atendimento ao cliente, reafirmando o valor estratégico da GenAI para otimizar processos e alocar recursos de forma mais eficiente.

Destaques

A implementação dessa solução evidencia como as tecnologias de GenAI podem redefinir paradigmas no setor financeiro, combinando eficiência operacional e excelência no atendimento ao cliente. Superando desafios técnicos e operacionais, nossos clientes não apenas modernizaram sua interação com os consumidores, mas também estabeleceram um padrão de inovação que pode servir como referência para toda a indústria.

Além disso, ao integrar o chatbot a APIs de sistemas internos — como consultas de saldo, status de processos ou dados de cadastro — a solução já incorpora um agente operacional dentro de um ecossistema de Agentic AI, capaz de interagir de forma autônoma com diferentes sistemas para executar tarefas e entregar respostas ainda mais personalizadas e resolutivas.

Esse caso ilustra como a GenAI atua como um habilitador técnico nas soluções; viabilizando uma interface que vai além da simples geração de respostas e se posiciona como um motor estratégico de soluções inteligentes. Em um cenário onde dados, sistemas e interações precisam ser orquestrados com inteligência, agilidade e segurança, iniciativas como essa demonstram o potencial da inteligência artificial em transformar operações, melhorar a experiência do cliente e gerar ganhos tangíveis de eficiência e performance das aplicações para o setor financeiro.

Compilado e adaptado a partir de demonstrações funcionais e estudos de MVPs aplicados em bancos digitais e tradicionais, focados em chatbots, APIs conversacionais e assistentes cognitivos.

Eficiência operacional: otimizando processos de Middle e Backoffice com GenAI

O setor financeiro, caracterizado por sua complexidade operacional e alta regulação, enfrenta desafios significativos na gestão de dados não estruturados. Arquivos de transações bancárias, documentos de compliance e relatórios de due diligence representam uma quantidade massiva de informações que demandam uma enorme quantidade de tempo para serem analisados manualmente, e quando não geridas adequadamente, podem limitar a eficiência organizacional em processos automatizados. Para superar essa barreira, uma nova solução baseada em inteligência artificial generativa (GenAI) foi desenvolvida, focada em transformar a maneira como os dados são processados e utilizados no middle e backoffice.

DESAFIOS E LIMITAÇÕES

O projeto enfrentou desafios específicos ao lidar com dados não estruturados, que podem ser divididos em dois fatores principais:

1

A propensão dos modelos de IA generativa a alucinações – respostas que fogem da realidade dos dados – exigindo o desenvolvimento de mecanismos rigorosos para garantir a qualidade das informações extraídas.

2

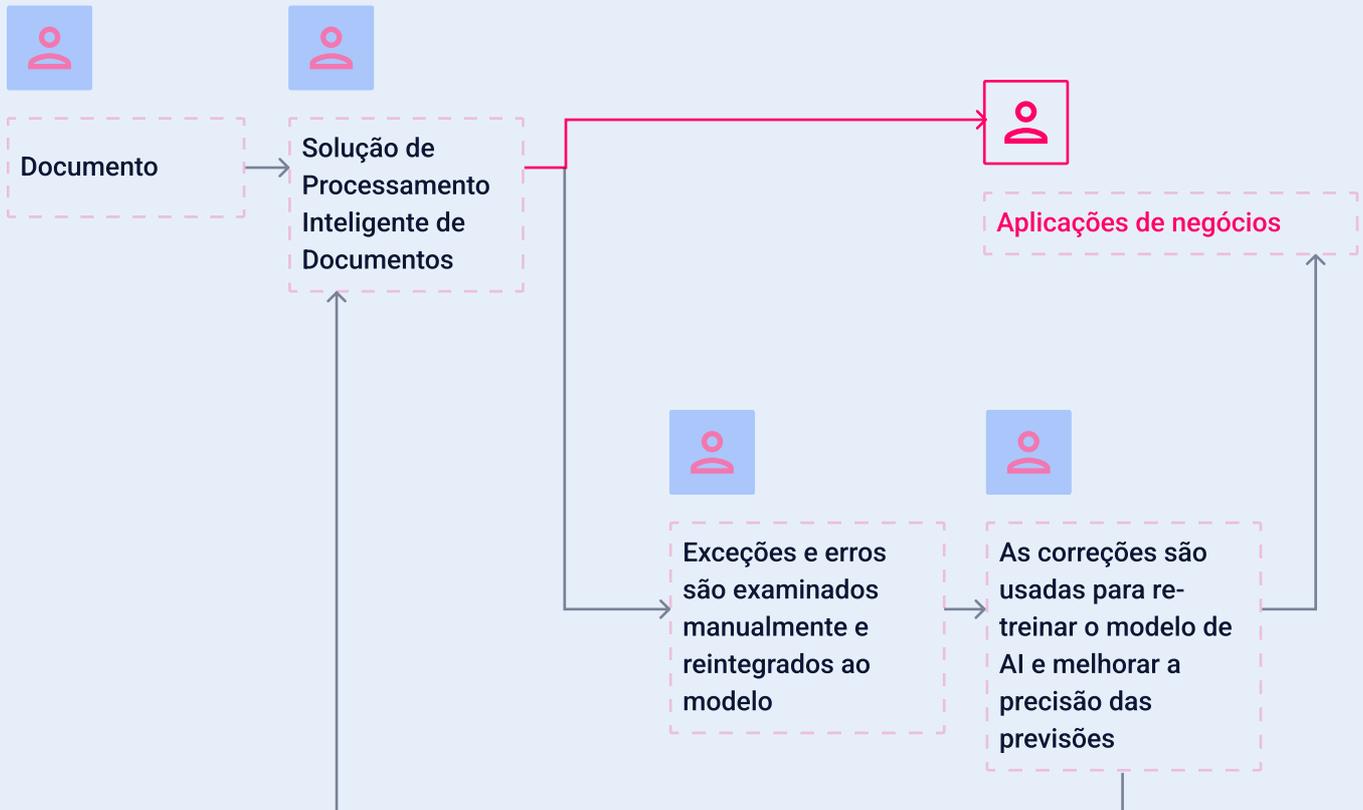
Alinhamento da solução às regulamentações e políticas de compliance específicas do setor financeiro, principalmente neste campo onde dados extremamente sensíveis são tratados, garantindo a segurança e conformidade em todas as etapas do processamento.



Solução Desenvolvida

A abordagem combinou a capacidade da GenAI de interpretar dados não estruturados com técnicas avançadas de processamento e extração de informações. A solução centralizou-se em um modelo treinado para transformar dados não estruturados em formatos estruturados e direcioná-los para tratamentos levando em consideração os objetivos específicos de cada área.

O impacto prático dessa aplicação foi significativo. As empresas puderam obter matrizes de risco com base em documentos de compliance, resumos consolidados de transações bancárias e análises detalhadas de due diligences. Além disso, a solução categorizou reclamações de clientes com precisão, fornecendo recomendações automatizadas para resolução de problemas em chats e chamadas, reduzindo assim a necessidade de interação humana.



Conteúdo adaptado a partir de materiais técnicos e resumos executivos sobre uso de IA aplicada ao processamento inteligente de documentos no setor financeiro.

Benefícios Alcançados

Os resultados do projeto incluem:

- Aumento substancial na **produtividade** devido à facilidade de análise e síntese de grandes volumes de documentos;
- **Redução de 1/3 nas chamadas ao call center**, com tratamento automatizado de reclamações e categorização de problemas;
- Estruturação de matrizes de risco altamente detalhadas, melhorando a **gestão de compliance**;
- **90% de precisão na análise de transações bancárias**, elevando o nível de confiança e eficiência operacional.



Essas soluções não apenas resolveram os desafios imediatos, mas também trouxeram ganhos de eficiência para setores diversos, como open banking, private equity (PE) e seguradoras. A adaptabilidade do modelo permitiu sua aplicação em diferentes contextos empresariais, estabelecendo um novo padrão para o uso de GenAI no gerenciamento de dados críticos.

Reflexões

Este projeto demonstra como tecnologias de GenAI podem transformar processos internos no setor financeiro, oferecendo análises mais rápidas, precisas e alinhadas às exigências regulatórias. Com a automação inteligente de tarefas antes manuais e complexas, as empresas conseguiram otimizar suas operações e redirecionar esforços para áreas estratégicas, reafirmando o papel da GenAI como um motor de inovação e eficiência.

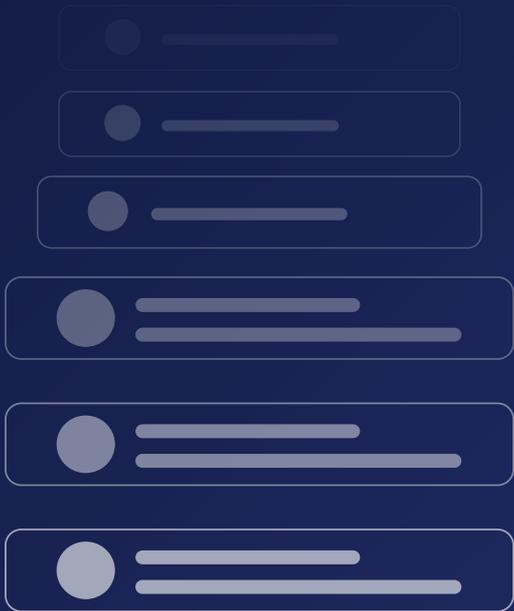
Considerando a arquitetura já implementada, se a solução passasse a não apenas interpretar os documentos recebidos, mas também a buscar de forma autônoma novos dados em sistemas internos ou bases externas via APIs – como consultas automáticas para atualizar relatórios de compliance ou validar informações de transações bancárias – ela já incorporaria as características e o potencial relacionado a um agente. Assim, o modelo evoluiria de uma solução de análise passiva para um agente dinâmico, que não apenas processa dados, mas também age continuamente para enriquecer e validar suas análises em tempo real.

Adaptado a partir de iniciativas práticas e estudos de caso sobre aplicação de GenAI e NLP em automação de documentos, análise de risco e atendimento ao cliente no setor financeiro.



CAPÍTULO 5

Principais Desafios no Setor Financeiro e como superá-los



Principais Desafios no Setor Financeiro e como superá-los

Antes de mostrar os principais desafios na implementação de GenAI e Agentic AI no setor financeiro, é importante ressaltar algumas convicções obtidas a partir da implementação de casos de uso reais e que são determinantes para superar os principais desafios e garantir o sucesso da aplicação de GenAI e Agentic AI na prática:

Convicções da Artefact sobre o uso de GenAI e Agentic AI no Setor Financeiro ^[3]

Uma adoção bem-sucedida de GenAI e Agentic AI requer um planejamento cuidadoso e a superação de desafios significativos, como a integração de sistemas legados e a gestão eficaz dos dados em um ambiente regulado. Para guiar no processo de adoção de GenAI, a Artefact possui sete grandes convicções para indústrias do setor de serviços financeiros:

Provas de conceito rápidas com visão estratégica de longo prazo

A qualidade dos dados como base para o sucesso

Propriedade intelectual como diferencial competitivo

Compliance e segurança como pilar fundamental

Framework de orquestração para integração fluida

Gestão de mudanças para garantir adesão organizacional

Feedback contínuo como motor de evolução

Provas de conceito rápidas com visão estratégica de longo prazo

A criação de provas de conceito (proof-of-concepts) rápidas é essencial para ganhar tração inicial, mas deve estar integrada a uma visão estratégica de longo prazo. Esse alinhamento garante que a execução imediata não se limite a iniciativas isoladas e permita que as empresas capturem valor sustentável da GenAI ao longo do tempo.

A qualidade dos dados como base para o sucesso

O verdadeiro valor da GenAI é desbloqueado quando a tecnologia é alimentada com dados de alta qualidade, preferencialmente de primeira mão. Esses dados fornecem uma base sólida para a geração de insights confiáveis e resultados precisos, fundamentais para atender às demandas de um mercado exigente.

Propriedade intelectual como diferencial competitivo

A criação de propriedade intelectual proprietária é crucial para garantir escalabilidade futura. Isso inclui o desenvolvimento de engenharia de prompt avançada e o registro sistemático de saídas para construir uma base de conhecimento que possa ser usada no refinamento e aprimoramento contínuos da tecnologia.

Compliance e segurança como pilar fundamental

A indústria de serviços financeiros opera em um ambiente altamente regulado, e a adoção de GenAI deve garantir a conformidade com as regulamentações locais e internacionais. Além de gerenciar riscos relacionados à segurança de dados e propriedade intelectual, é essencial que as soluções de GenAI sejam projetadas para atender às exigências regulatórias, preservando a confiança dos stakeholders.

Framework de orquestração para integração fluida

Além da prontidão técnica, um framework de orquestração é indispensável para integrar a GenAI de maneira coesa nos fluxos de trabalho e sistemas existentes. Essa estrutura garante que as soluções de GenAI sejam aplicadas de forma eficaz, sem interromper processos estabelecidos.

Gestão de mudanças para garantir adesão organizacional

A implementação de GenAI exige mais do que tecnologia; demanda uma transformação cultural e organizacional. A gestão de mudanças é crucial para envolver stakeholders, treinar equipes e assegurar que a adoção da tecnologia seja integrada aos processos de forma eficaz. Sem um plano estruturado de change management, o risco de resistência e baixa adesão cresce significativamente.

Feedback contínuo como motor de evolução

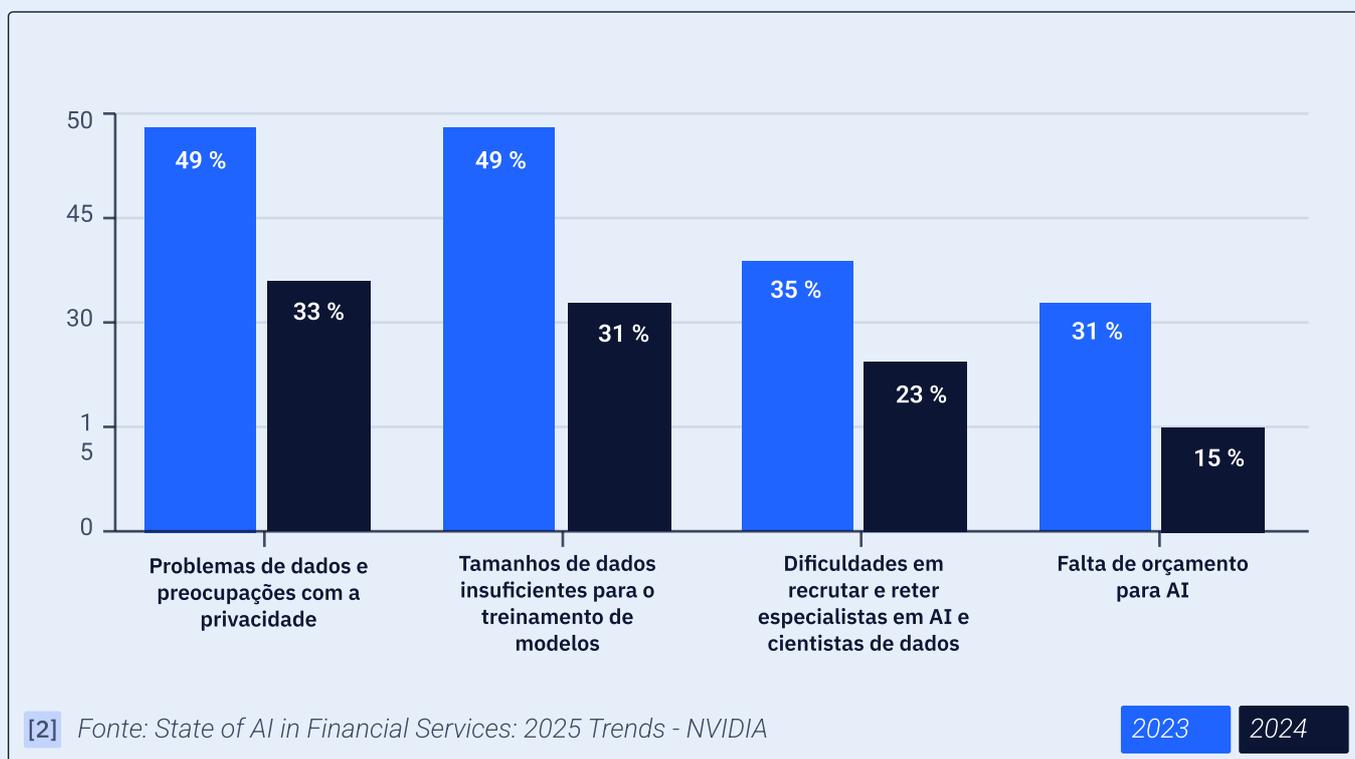
Mecanismos de feedback contínuo, com participação humana no processo, são essenciais para garantir melhorias constantes, alinhamento com os objetivos de negócio e a criação de propriedade intelectual. Esse ciclo de aprimoramento é um fator determinante para o sucesso de longo prazo.

Com essas convicções, a Artefact reforça sua capacidade de guiar empresas do setor financeiro em uma adoção estratégica e eficiente da GenAI, maximizando o impacto positivo da tecnologia enquanto supera os desafios associados à sua implementação.

Principais desafios para Implementação de AI no Setor Financeiro

Os desafios tradicionais para a implementação de AI no setor financeiro, como preocupações com dados, limitações de orçamento e dificuldade em recrutar talentos, vêm apresentando uma tendência de queda.

Segundo o relatório State of AI in Financial Services: 2025 Trends da NVIDIA, questões como privacidade de dados e tamanho insuficiente de bases para treinamento, que antes eram barreiras expressivas, registraram uma redução significativa de 2023 para 2024. Esse movimento indica uma maturidade crescente no mercado: as organizações estão avançando do estágio de experimentação para a consolidação de projetos em larga escala, refletindo um ambiente mais preparado para capturar o valor estratégico da inteligência artificial. [2]



Apesar da redução dos obstáculos iniciais, no setor financeiro a atenção aos riscos permanece fundamental. A natureza altamente regulada e sensível da indústria exige que iniciativas de AI sejam desenhadas com rigor adicional para garantir confiança, compliance e estabilidade de mercado.

A imagem a seguir reforça essa necessidade, mostrando que a implementação de AI e GenAI vai muito além da tecnologia – envolve a gestão ativa de riscos estratégicos. Navegar com sucesso por esses desafios exige estruturas de governança robustas, sistemas de AI transparentes e uma força de trabalho qualificada, capaz de usar e supervisionar efetivamente essas tecnologias.

Protegendo o valor da AI/GenAI ao mesmo tempo em que se aborda o **imperativo de risco**

Nos serviços financeiros, o risco da AI não se trata apenas de tecnologia – trata-se de confiança, conformidade e estabilidade do mercado.

Risco Regulatório e de Conformidade	Risco de Mercado e de Crédito	Risco de Reputação	Risco Operacional
<ul style="list-style-type: none"> Modelo de Governança Relatórios Regulatórios 	<ul style="list-style-type: none"> Risco de Decisão Algorítmica Gestão de Portfólio 	<ul style="list-style-type: none"> Serviços de Consultoria Interação com Clientes 	<ul style="list-style-type: none"> Qualidade de Dados e Viés Integração de Sistemas

AI Use Case	Inherent Risk Type	Risk Description	Existing Control	Residual Risk Score
Fraud Detection	Viés e Justiça	Risco de produzir previsões de fraude enviesadas devido a conjuntos de dados de treinamento desequilibrados	Conjuntos de treinamento diversos e representativos com auditorias regulares	BAIXO
Loan Approval	Transparência e Explicabilidade	Falta de clareza nas decisões de aprovação de crédito pode levar a insatisfação do cliente ou escrutínio regulatório	Ferramentas de explicabilidade integradas (ex.: valores SHAP); supervisão humana	MÉDIO
GenAI Financial Report Summary Generator	Alucinação	Risco de gerar relatórios financeiros incorretos ou enganosos, resumos ou insights	Verificação de fatos integrada no fluxo de saída; modelo ajustado com dados financeiros verificados; revisão humana para relatórios críticos	ALTO

O framework apresentado organiza os principais riscos em quatro grandes categorias: Regulatório e Compliance, Mercado e Crédito, Reputação e Operacional. Dentro dessas categorias, são mapeados riscos específicos como governança de modelos, risco de decisões algorítmicas, qualidade de dados e integração de sistemas. Cada tipo de risco é exemplificado com casos de uso: na detecção de fraudes, o principal desafio é mitigar vieses de dados; na aprovação de crédito, garantir transparência e explicabilidade; e na geração de relatórios financeiros com GenAI, controlar alucinações – risco que, mesmo com medidas de mitigação, ainda apresenta alta criticidade residual.

Esse panorama reforça que, mesmo em um ambiente de maturação tecnológica, o sucesso da AI no setor financeiro depende diretamente da capacidade das organizações em proteger o valor gerado, gerenciando riscos de forma estruturada e contínua.

Além dos riscos inerentes aos modelos, a necessidade de lidar com informações sensíveis e atender às exigências de órgãos reguladores torna a implementação de AI ainda mais complexa. Setores como prevenção à fraude e análise de crédito exigem altos níveis de precisão, transparência e compliance, uma vez que erros podem gerar impactos financeiros e reputacionais relevantes [2]

Apesar dos avanços tecnológicos e da maturidade crescente do mercado, a implementação de AI Generativa (GenAI) no setor financeiro ainda enfrenta desafios consideráveis, especialmente em relação à escalabilidade. Apenas cerca de 10% das instituições financeiras conseguem escalar suas iniciativas de AI com sucesso. Embora ferramentas como ChatGPT ou Copilot facilitem a criação de provas de conceito (POCs), o grande obstáculo está na transformação dessas soluções em produtos que gerem impacto real. Para que a GenAI tenha um efeito significativo, ela precisa ser capaz de responder a questões estratégicas de negócios de maneira robusta e escalável, indo além da simples integração com processos existentes. Sem essa estruturação, medir o retorno sobre o investimento, reduzir custos e aumentar a satisfação do cliente se torna um desafio complexo.



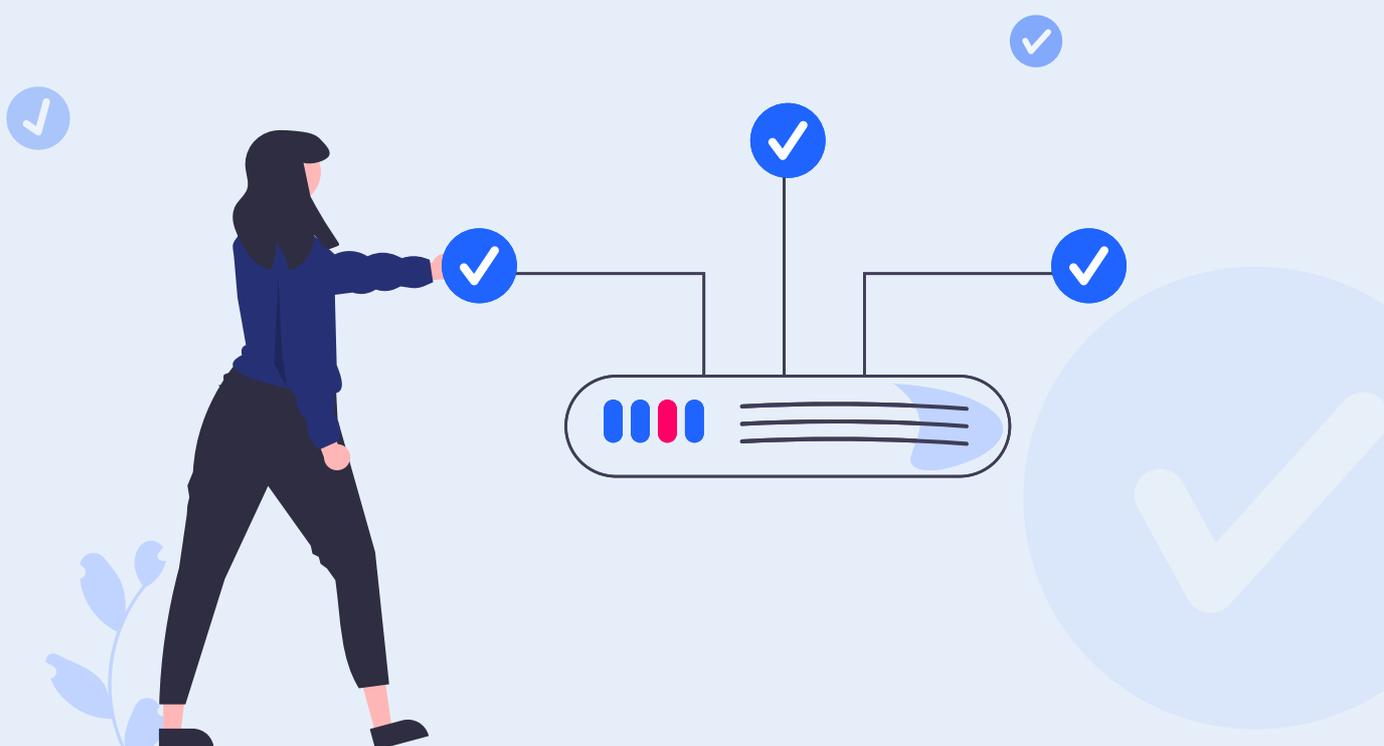
[7] Fonte: AI for Finance Event 2024 – Palestra de Joffrey Martinez (Artefact).

Boas práticas para superar desafios e garantir uma implementação bem-sucedida

Para garantir uma implementação bem-sucedida de AI no setor financeiro, as instituições devem adotar uma abordagem estratégica que enfrente os principais desafios de forma integrada.

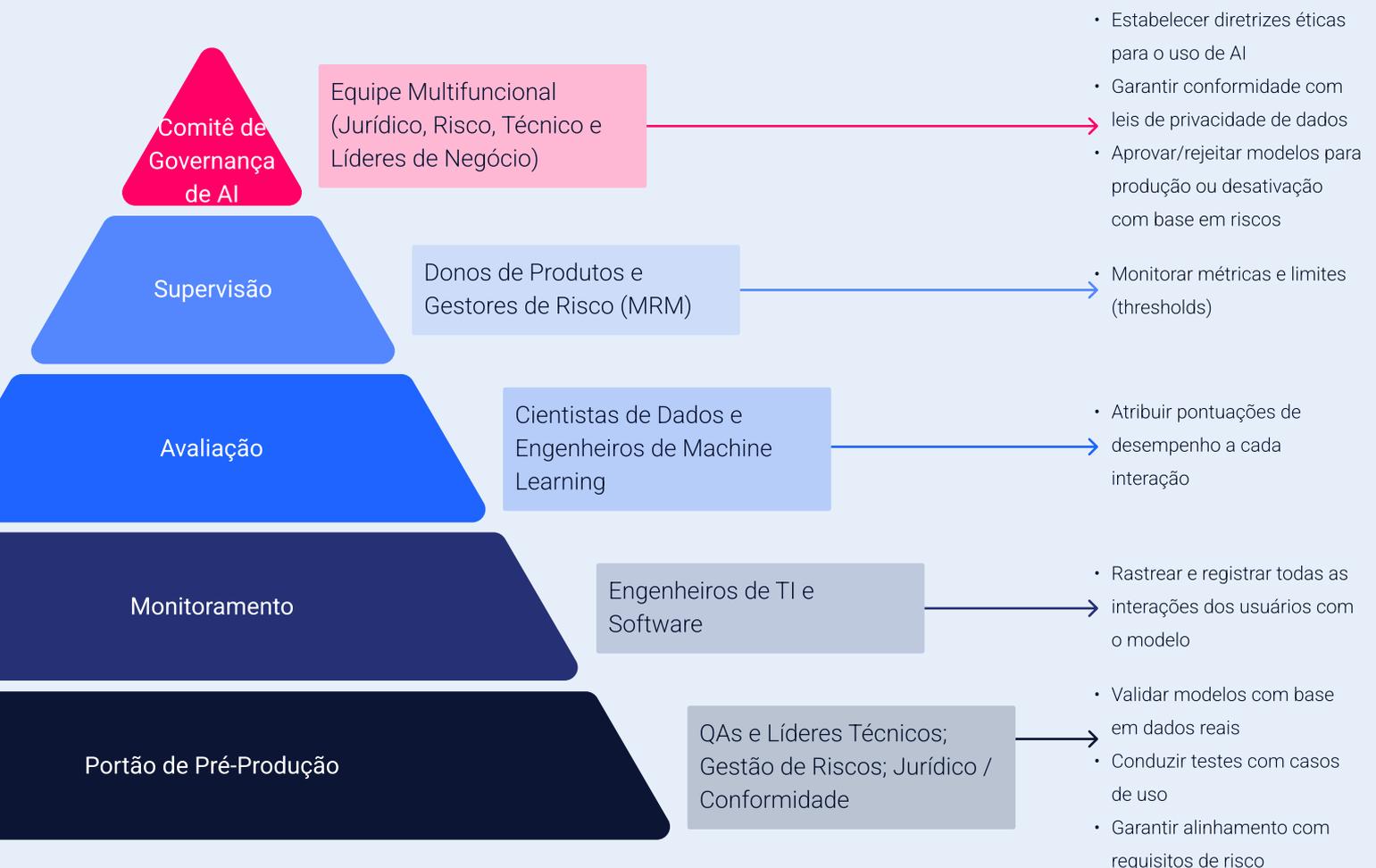
A construção de uma governança sólida, com políticas claras e mecanismos de supervisão adequados, é indispensável para gerenciar riscos e assegurar conformidade regulatória.

Além disso, investir no desenvolvimento de talentos especializados em AI pode mitigar a escassez de profissionais qualificados, facilitando a adoção dessa tecnologia de maneira eficaz. A melhoria na qualidade dos dados também é crucial: políticas robustas de governança de dados, práticas aprimoradas de coleta e proteção da privacidade e segurança são fatores determinantes para o sucesso. Por fim, a transparência nos sistemas de AI aumenta a confiança, garante supervisão eficaz e, junto com práticas éticas bem definidas, reduz vieses e promove decisões justas.



A gestão estruturada de riscos e a definição clara de papéis e responsabilidades continuam sendo pilares essenciais para o sucesso sustentável das iniciativas de AI. O framework de governança de AI descrito no capítulo oferece uma visão prática e hierárquica, mostrando como diferentes níveis de responsabilidade — desde comitês de governança até equipes técnicas — colaboram para garantir a implementação segura e eficaz dessas tecnologias. Essa estrutura destaca a importância de alinhar competências multidisciplinares, como as áreas legal, risco e engenharia, a processos bem definidos, que abrangem desde validações pré-produção até a supervisão contínua em produção. Essa abordagem não apenas mitiga riscos, mas também promove transparência e compliance, assegurando que as iniciativas de AI estejam alinhadas aos objetivos estratégicos e regulatórios do setor financeiro.

Governar para Crescer: definindo papéis e responsabilidades para o sucesso sustentável da AI



Uma governança eficaz de AI deve ser construída sobre pilares multifuncionais que integrem valor de negócio, escalabilidade técnica e conformidade ética. Conforme ilustrado na imagem, três frentes fundamentais sustentam essa abordagem: gestão da demanda e geração de valor, que inclui frameworks claros para desenvolvimento de casos de uso, bibliotecas centralizadas de prompts otimizados e métricas para medir adoção e ROI; escalabilidade e integração, com foco em segurança desde o design, supervisão contínua e governança integrada de dados vetoriais (VDBs) para garantir consistência e qualidade; e ética e regulamentação, que asseguram alinhamento com legislações globais, como GDPR e AI Act, e promovem práticas de mitigação de vieses para decisões mais justas e inclusivas. Essa estrutura integrada é essencial para escalar o uso de GenAI com eficiência, segurança e aderência aos padrões de governança corporativa e regulatórios



AI GOVERNANCE



Gestão de Demanda & Valor

Framework de Desenvolvimento

Guia estruturado com boas práticas, marcos e etapas de validação para o desenvolvimento de casos de uso de GenAI.

Biblioteca e Diretrizes de Prompts

Biblioteca centralizada de prompts otimizados com regras claras de uso para assegurar padronização.

Adoção e Medição de ROI

Promover a adoção por meio de gestão de mudanças e medir o sucesso com KPIs de ROI e satisfação do usuário.



Escalabilidade & Integração

Controle e Supervisão

Modelos confiáveis e eficientes com monitoramento contínuo, avaliação de KPIs e insights de experiência do usuário.

Seguro por Design

Segurança de nível corporativo com autenticação segura e separação de direitos de acesso.

Governança Integrada de Dados para VDBs

Bases de dados vetoriais (VDBs) herdam governança de fontes de dados, garantindo consistência, sem duplicação e com qualidade de dados.



Ética & Regulatório

Conformidade com Regulamentações

Alinhar o uso de GenAI com regulamentações globais (AI Act, GDPR, CCPA) por meio de gestão de riscos e rastreabilidade.

Mitigação de Viés e Promoção de Justiça

Garantir resultados imparciais detectando vieses, projetando para inclusão e avaliando continuamente a justiça.

Machine Learning: Aprimorando a Imparcialidade e a Interpretabilidade

Para superar os desafios do viés algorítmico e da interpretabilidade do modelo em machine learning, as instituições financeiras devem priorizar algoritmos com reconhecimento de imparcialidade e técnicas de AI explicáveis (XAI) [8] [9]. A implementação de procedimentos rigorosos de teste e validação pode ajudar a identificar e mitigar vieses nos dados de treinamento e nas saídas do modelo [8]. A utilização de técnicas como os valores SHAP e LIME para quantificar o peso das features nos resultados dos modelos, pode fornecer insights sobre os processos de tomada de decisão dos modelos de machine learning, aprimorando a interpretabilidade e a transparência [8]. Além disso, o envolvimento com as partes interessadas e os reguladores para estabelecer diretrizes claras para a ética e a transparência da AI é essencial para a implantação responsável da AI [8].

GenAI: Mitigando Alucinações e Garantindo a Segurança dos Dados

Para enfrentar os desafios da GenAI, é necessário focar na segurança, privacidade e precisão dos dados. A implementação da Geração Aumentada de Recuperação (RAG) pode melhorar a precisão da resposta e o contexto, extraindo informações de fontes internas da empresa. Os bancos devem implementar medidas robustas de segurança de dados e obter o consentimento explícito do cliente para o uso da AI, a fim de aderir às rígidas regulamentações de privacidade de dados [8]. Além disso, o envolvimento ativo entre os bancos e os órgãos reguladores é necessário para navegar no cenário regulatório em evolução e mitigar possíveis imprecisões nas previsões da AI, estabelecendo estruturas transparentes e eficazes [8].

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

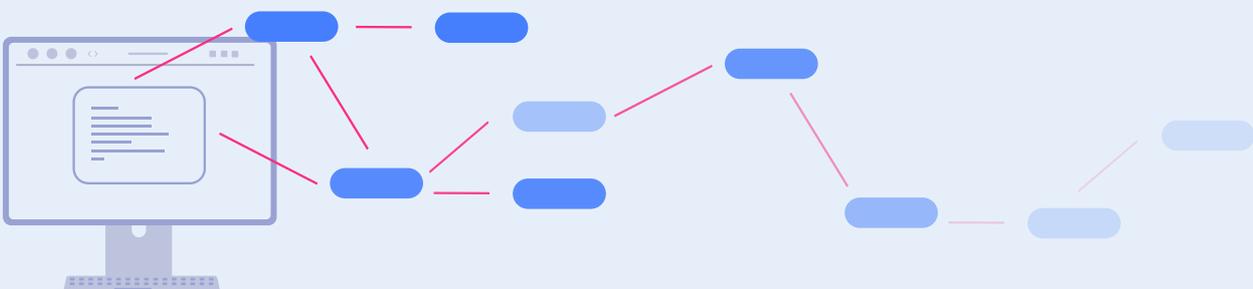
[9] Fonte: WhatNext.Law – Legal challenges of AI in finance.

Agentic AI: Fortalecendo a Governança e os Limites Éticos

Para superar os desafios da Agentic AI, é preciso fortalecer a governança, os limites éticos e a infraestrutura de segurança. As instituições financeiras devem estabelecer diretrizes éticas claras para a tomada de decisões da AI, incorporando a supervisão humana e os mecanismos de responsabilização [8] [9]. A implementação de medidas de segurança robustas e os denominados “guard rails” é crucial para evitar o uso indevido, como a lavagem de dinheiro ou a negociação com informações privilegiadas [8]. O viés algorítmico na Agentic AI pode ser abordado por meio de monitoramento cuidadoso, estratégias de mitigação e processos transparentes de tomada de decisão [8].

Por fim, a escalabilidade da AI como um todo envolve questões éticas, ambientais e sociais, além de depender fortemente de uma gestão eficaz de mudanças (change management), especialmente devido à complexidade organizacional dessas instituições.

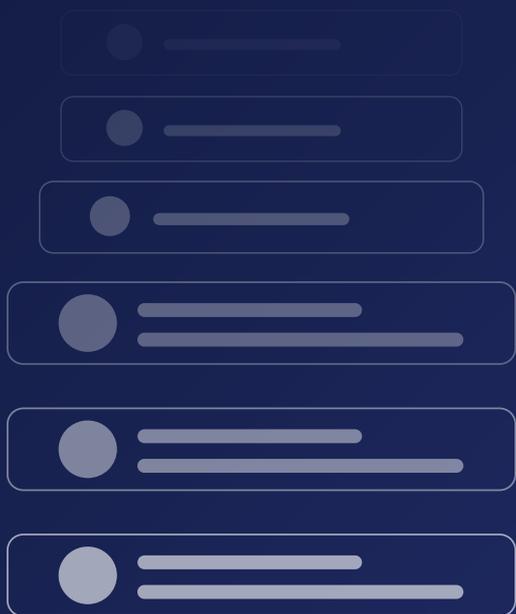
A promoção de uma cultura de responsabilidade e propriedade sobre os dados é essencial, mas muitas vezes esbarra na falta de alinhamento entre diferentes áreas e na resistência à mudança. Assim, o sucesso na implementação da AI exige uma abordagem holística que equilibre inovação tecnológica, responsabilidade social e ambiental e o fomento e valorização de uma cultura organizacional orientada a dados.





CAPÍTULO 6

Estratégia na **implementação** **de AI** no Setor Financeiro

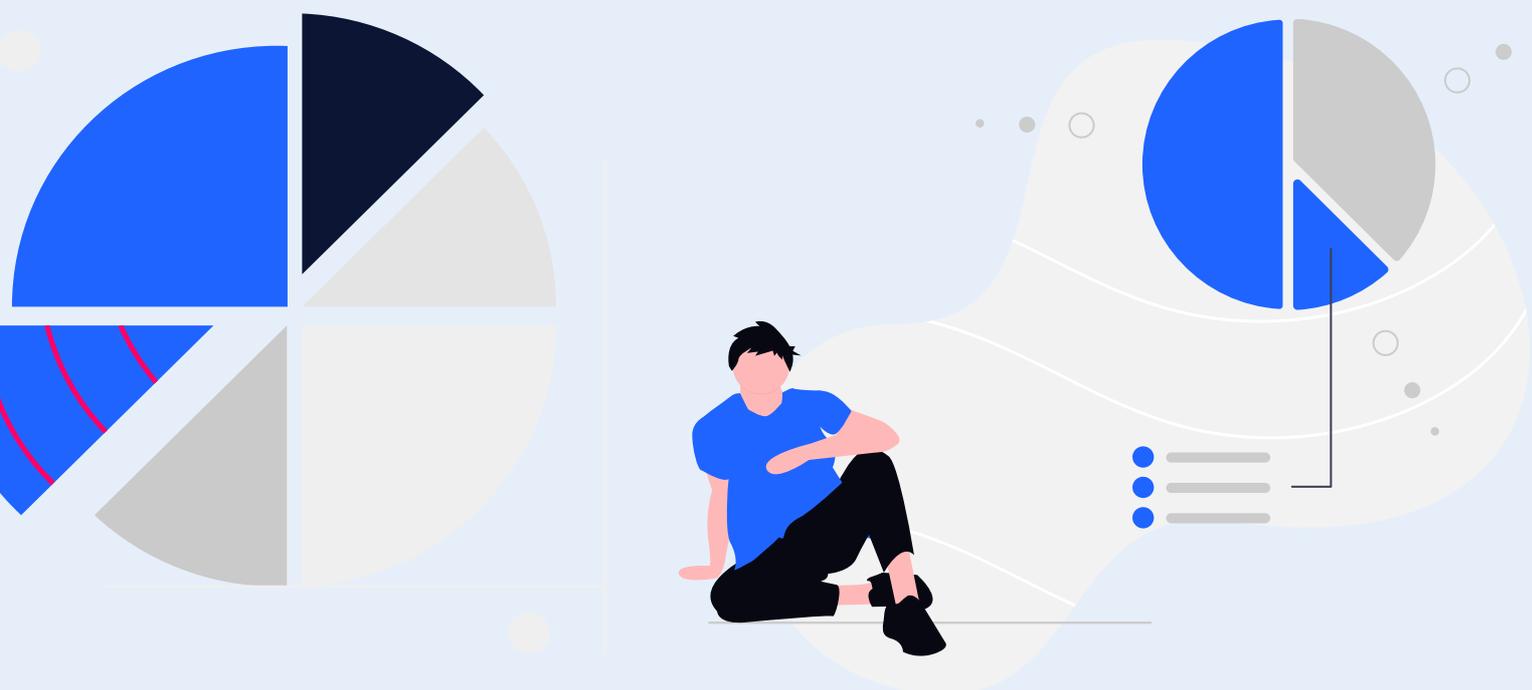


Estratégia na implementação de **AI** no Setor Financeiro

O sucesso na implementação de casos de uso de AI começa com uma avaliação precisa da maturidade organizacional e do modelo operacional existente. Esse diagnóstico é fundamental para identificar lacunas e potenciais riscos que possam comprometer o projeto. Com base nessa análise, as instituições financeiras podem adotar as seguintes estratégias:

DIAGNÓSTICO E PLANEJAMENTO ESTRATÉGICO

Para garantir uma adoção bem-sucedida de qualquer tecnologia, é essencial compreender onde a empresa se encontra hoje em sua jornada de maturidade digital. Somente com um diagnóstico preciso é possível estruturar um planejamento sólido que direcione a organização para o patamar desejado. Esse entendimento permite que a implementação de GenAI seja feita de maneira estratégica, maximizando valor e minimizando riscos.



Na Artefact, avaliamos a maturidade digital das empresas por meio de três dimensões fundamentais:

 <p>Estratégia</p> <p>Os dados são valorizados como um ativo essencial para impulsionar a estratégia de negócios, garantindo que a tomada de decisão seja orientada por insights robustos.</p>	 <p>Entrega</p> <p>A organização deve ser capaz de transformar dados em resultados concretos, aplicando inteligência analítica para otimizar processos e gerar impacto positivo.</p>	 <p>Gerenciamento e Governança de Dados</p> <p>O gerenciamento adequado dos dados assegura qualidade, compliance e segurança, fatores indispensáveis para um ambiente altamente regulado como o setor financeiro.</p>
--	--	---

Essas dimensões são avaliadas dentro de 3 aspectos:

 <p>Pessoas</p> <p>Avalia-se o comprometimento da liderança com uma estratégia de dados bem estruturada e financiada, na capacitação dos colaboradores para o uso estratégico dos dados e no patrocínio executivo para a governança e gestão eficazes, garantindo alinhamento organizacional e geração de valor para o negócio</p>	 <p>Processos</p> <p>Avalia-se o uso sistemático dos ativos de dados para gerar valor, melhorar a proposta ao cliente e otimizar operações. Além disso, analisa a integração de analytics e insights confiáveis nos processos de negócio, bem como a existência de uma governança estruturada, com definição clara, documentação, execução e monitoramento dos processos de gestão de dados, garantindo sua confiabilidade.</p>	 <p>Tecnologia</p> <p>Avalia-se a existência de uma estratégia técnica clara e um roadmap que define as capacidades necessárias para atingir os objetivos de dados da organização. Também considera se a arquitetura tecnológica suporta a entrega de casos de uso prioritários e se está preparada para o futuro. Além disso, avalia a conformidade, segurança, estabilidade e controle de acesso aos dados.</p>
--	---	---

De forma bidimensional podemos exemplificar as dimensões e aspectos da maturidade levados analisado de acordo com a ilustração abaixo:

	<i>Estratégia</i>	<i>Entrega</i>	<i>Gestão e Governança de Dados</i>
	Os dados são valorizados como um ativo estratégico	Os dados impulsionam a entrega	Gestão e governança de dados garantem conformidade e impulsionam a qualidade dos dados
<i>Pessoas</i>	A liderança assume ativamente uma estratégia de dados ambiciosa e prática, com o nível certo de financiamento e apoio organizacional.	Todos na organização, usuários de negócios e especialistas em dados, têm o conhecimento necessário para usar os dados e gerar resultados de negócio.	A gestão e governança de dados contam com patrocínio executivo.
<i>Processos</i>	Ativos de dados são usados sistematicamente para criar novo valor, melhorar a proposta ao cliente e impulsionar melhorias operacionais.	Processos de negócios são impulsionados por análises e insights confiáveis.	Os processos de gestão e governança de dados são de responsabilidade clara, definidos, documentados, entregues e monitorados. Os dados são confiáveis.
<i>Tecnologia</i>	Existe uma estratégia técnica clara e um roadmap que define as capacidades necessárias para atingir as ambições de dados da organização.	O stack tecnológico permite a entrega de casos de uso prioritários de forma adequada e preparada para o futuro.	Os dados são conformes, seguros, estáveis e com controle de acesso.

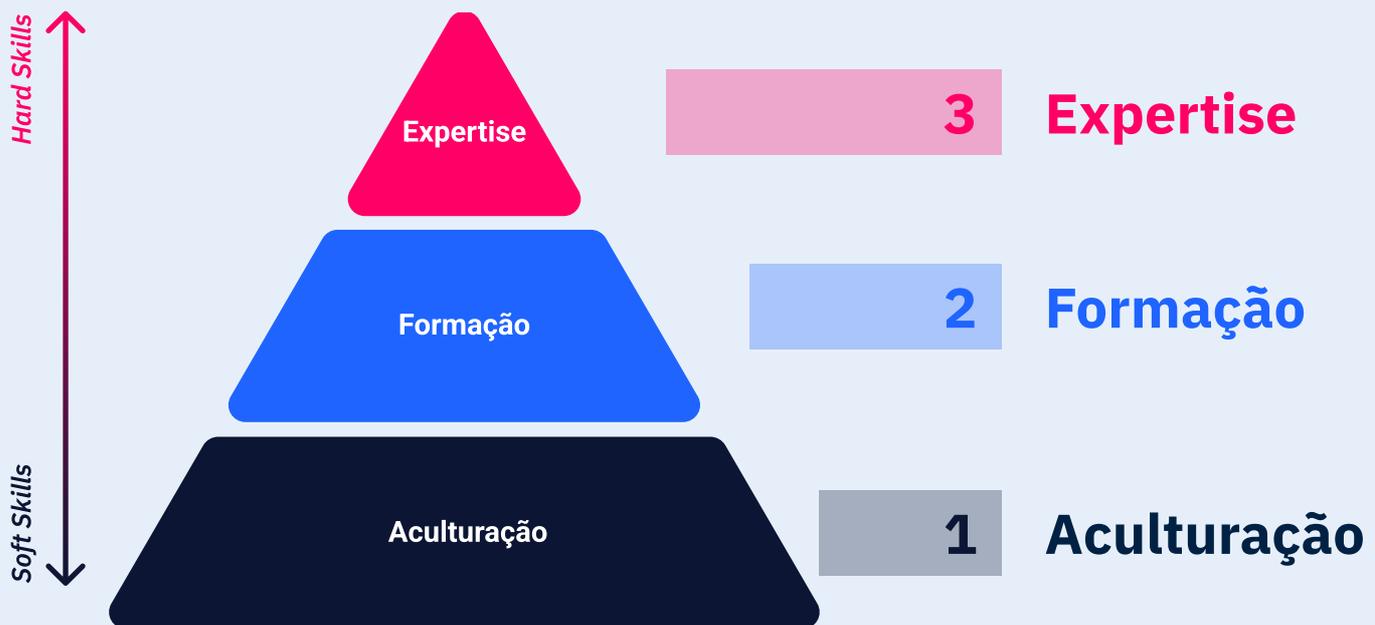
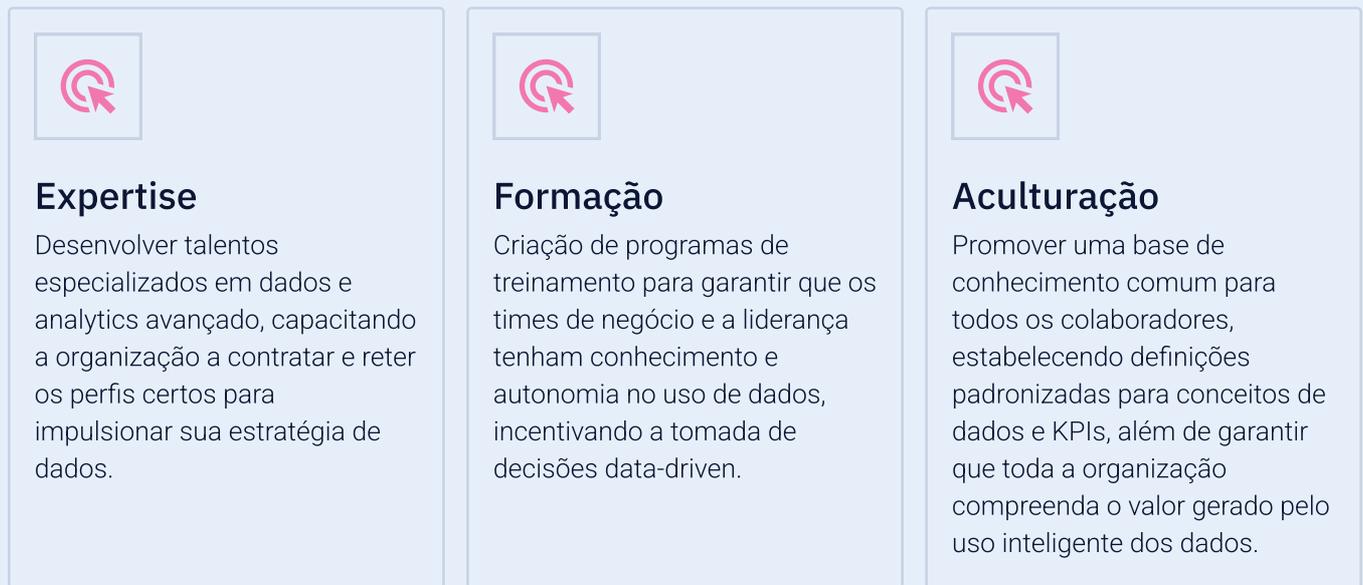
Com essa abordagem estruturada, as instituições financeiras podem não apenas implementar IA de maneira eficiente, mas também garantir que a tecnologia esteja alinhada à estratégia de negócios, fortalecendo a competitividade e a resiliência organizacional.

Engajamento das Partes Interessadas

O engajamento das partes interessadas é um elemento essencial para o sucesso na adoção de IA no setor financeiro. Envolver líderes das áreas de negócios, tecnologia, compliance e segurança desde o início do processo permite alinhar expectativas e garantir suporte contínuo ao longo do projeto. Além disso, uma abordagem colaborativa, que promova o engajamento e a responsabilidade compartilhada entre as equipes, é fundamental para integrar diferentes perspectivas e construir uma visão unificada dos objetivos estratégicos.

A transformação digital e a adoção de IA vão além da tecnologia: exigem uma gestão de mudança estruturada dentro da organização. Isso envolve tanto o desenvolvimento de soft skills, como a adaptação cultural e a mentalidade orientada a dados, quanto hard skills, como a capacitação técnica necessária para operar as novas ferramentas de inteligência artificial.

Na Artefact, aceleramos esse processo de mudança organizacional para nossos clientes por meio do nosso framework que possui três alavancas essenciais:



Pilotos e Iterações: uma abordagem ágil para reduzir riscos e maximizar valor

Quando falamos de soluções de AI, é essencial reconhecer que existem riscos inerentes, como alucinações, viés nos modelos e desafios de interpretabilidade. Para mitigar esses desafios, a gestão de mudanças mencionada anteriormente deve criar um ambiente propício à experimentação, permitindo que as equipes testem e validem novas soluções de maneira controlada.

BUSINESS DISCOVERY

O primeiro passo nessa jornada é a fase de descobrimento do negócio (Business Discovery). Nesse estágio inicial, é fundamental mapear possíveis casos de uso da tecnologia dentro da organização tentando ser o mais exaustivo possível. Essa etapa deve estar diretamente alinhada à visão estratégica da empresa, garantindo que a adoção da AI tenha um propósito claro e gere impacto real.



ESTIMATIVA DE VALOR

Após o mapeamento, é necessário estimar o valor gerado por cada Caso de Uso. Isto permite uma priorização que garanta que os esforços estejam focados nas iniciativas que trarão o maior retorno no menor tempo possível, sempre buscando agregar valor de maneira ágil e eficiente. A priorização correta permite que os primeiros projetos testados sirvam como prova de conceito para futuras iniciativas, criando um ciclo sustentável de inovação.

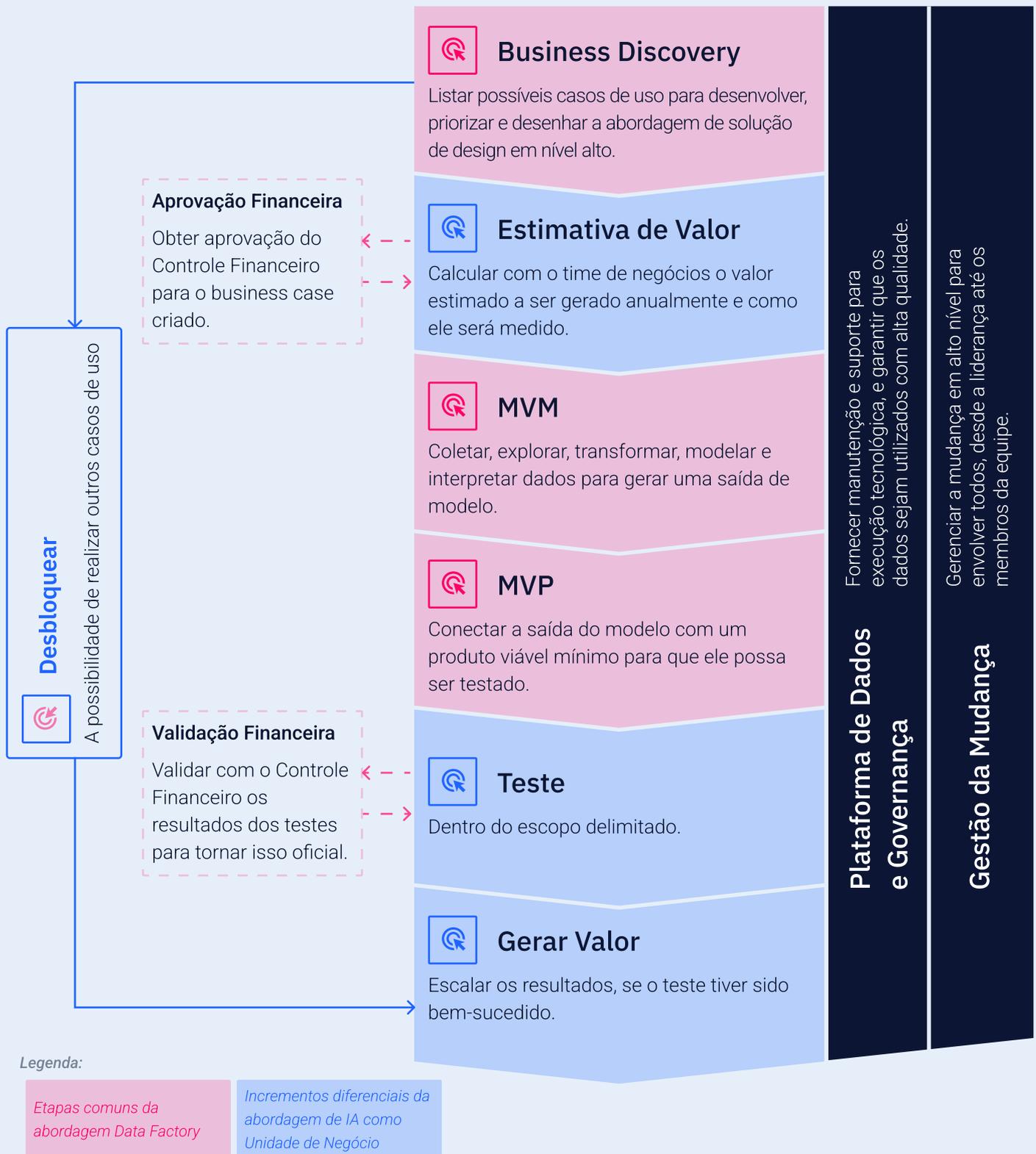


Com os casos priorizados, passamos para a fase de construção e teste, seguindo a lógica de MVM (Minimum Viable Model), MVP (Minimum Viable Product) e experimentação iterativa. É nesse momento que o princípio de "começar pequeno e errar rápido" entra em jogo. Essa metodologia reduz custos e minimiza riscos, pois permite ajustes contínuos antes de um investimento mais robusto na solução final. Através de múltiplas iterações, os modelos são refinados para garantir um escopo bem definido, com menor propensão a erros e maior confiabilidade.



Uma vez que o teste gera resultados positivos e conclusivos, chega o momento de escalar a solução. Nessa etapa, a tecnologia passa a ser aplicada de forma mais ampla dentro da organização, maximizando seu impacto e gerando valor tangível para o negócio. Esse ciclo de aprendizado e validação cria um efeito multiplicador: o ganho em eficiência e retorno financeiro obtido com os primeiros projetos permite que novas iniciativas sejam financiadas, acelerando ainda mais a transformação digital da empresa.

Ao combinar experimentação controlada, iteração contínua e escalabilidade estratégica, as instituições financeiras conseguem explorar o potencial da Gen AI de forma segura, eficiente e sustentável, garantindo que a inovação traga resultados concretos e de longo prazo.

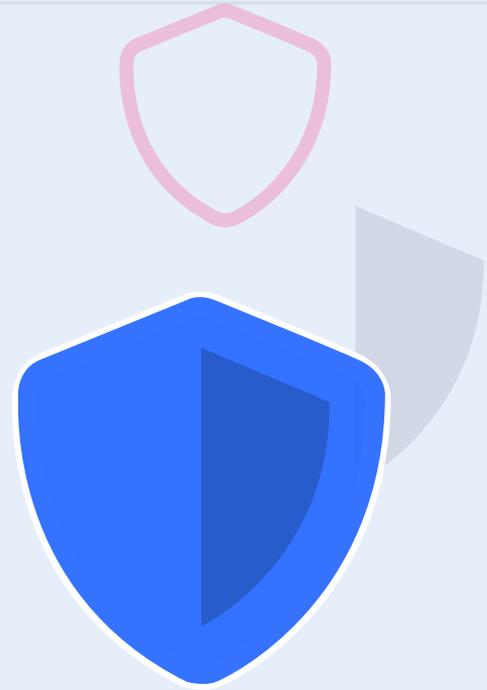


Ao seguir essas etapas, as instituições financeiras podem estabelecer uma base sólida para a adoção e criação de casos de uso, criando um ambiente onde inovação e segurança caminham lado a lado. Assim, é possível maximizar o valor gerado pela tecnologia, promovendo eficiência, compliance e resultados tangíveis para o negócio.

A Segurança e Privacidade de Dados como uma **Prioridade Estratégica**

No setor financeiro, segurança e privacidade de dados são fatores inegociáveis. Bancos e outras instituições lidam com informações altamente sensíveis, como históricos financeiros, dados pessoais e transações, o que exige medidas rigorosas de proteção. Com o advento da Inteligência Artificial (AI), incluindo suas variantes como Machine Learning e GenAI, essa necessidade tornou-se ainda mais crítica.

A escolha da infraestrutura tecnológica para suportar soluções de AI é um fator estratégico crucial que impacta diretamente na capacidade de implementar e manter robustas medidas de segurança. Neste contexto, as instituições financeiras se deparam com uma escolha fundamental: adotar soluções baseadas em nuvem ou manter uma infraestrutura on-premise.



Na implementação de soluções de AI no setor financeiro, as instituições enfrentam uma decisão crucial entre adotar infraestruturas baseadas em nuvem ou manter sistemas on-premise. Cada abordagem oferece características distintas que podem impactar significativamente o sucesso e a eficácia das iniciativas de AI.

A tabela a seguir apresenta uma comparação objetiva entre soluções baseadas em nuvem e soluções on-premise, destacando os principais aspectos, vantagens e desvantagens de cada abordagem. Essa análise busca auxiliar as instituições financeiras na escolha da infraestrutura mais adequada para suportar suas iniciativas de inteligência artificial, considerando fatores como custo, escalabilidade, segurança, conformidade regulatória e dependência tecnológica.



ASPECTO	SOLUÇÕES EM NUVEM	SOLUÇÕES ON-PREMISE
Infraestrutura	Gerenciada por provedores terceirizados.	Totalmente gerenciada pela instituição.
Escalabilidade	Alta, com expansão ágil conforme demanda.	Depende da infraestrutura existente.
Custos Iniciais	Baixos, modelo de pagamento por uso.	Altos, requer investimento significativo em hardware e software.
Manutenção	Atualizações automáticas pelo provedor.	Responsabilidade da instituição.
Segurança e Privacidade	Riscos de compliance e armazenamento externo.	Maior controle sobre dados sensíveis.
Conformidade Regulatória	Pode ser complexa em relação à localização de dados.	Facilita o cumprimento de regulamentações específicas.
Tecnologia	Acesso a frameworks e ferramentas de ponta oferecidos por provedores líderes.	Infraestrutura adaptada às necessidades específicas da instituição.
Dependência	Alta, atrelada à confiabilidade dos provedores externos.	Baixa, operação independente de terceiros.



Nas próximas páginas, serão explorados em maior profundidade os detalhes de cada abordagem. Serão discutidas as características específicas, os casos de uso mais adequados e as melhores práticas para a implementação de soluções em nuvem e on-premise, fornecendo uma visão abrangente para apoiar a tomada de decisão estratégica no contexto das instituições financeiras.

Soluções em Nuvem

Refere-se à utilização de serviços de computação fornecidos por provedores terceirizados através da internet. Neste modelo, a infraestrutura, plataformas e softwares são gerenciados pelo provedor de serviços em nuvem, permitindo que as instituições financeiras acessem recursos de AI sob demanda, sem a necessidade de manter hardware físico em suas instalações.



✔ VANTAGENS:

Escalabilidade e Flexibilidade: Expansão ágil da capacidade de processamento, ajustando-se às demandas do negócio.

Acesso a Tecnologia de Ponta: Provedores como AWS, Google Cloud e Azure oferecem plataformas otimizadas para Gen AI, com frameworks e ferramentas líderes de mercado.

Modelo de Pagamento por Uso: Reduz custos iniciais, permitindo que instituições invistam proporcionalmente ao uso.

Atualizações Automáticas: Garantia de acesso às versões mais recentes de segurança e tecnologia, sem necessidade de intervenções internas.

⊗ DESVANTAGENS:

Preocupações com Segurança e Privacidade: Dados sensíveis armazenados fora da infraestrutura interna podem gerar desafios de compliance e riscos de segurança.

Dependência de Provedores Externos: A operação fica atrelada à confiabilidade e continuidade dos serviços oferecidos por terceiros.

Conformidade Regulatória: Algumas leis exigem que dados permaneçam no país ou sejam tratados com controles específicos, o que pode ser complexo de atender em certas plataformas de nuvem.

Soluções On-Premise

Envolve a implementação e manutenção de toda a infraestrutura de AI dentro das instalações físicas da própria instituição financeira. Neste cenário, a organização é responsável por todos os aspectos da infraestrutura, incluindo hardware, software, segurança e manutenção, oferecendo controle total sobre os sistemas e dados.



✔ VANTAGENS:

Autonomia e Controle Absoluto: As operações permanecem inteiramente dentro da infraestrutura da instituição, garantindo maior proteção de dados sensíveis.

Facilidade de Conformidade: A localização de dados dentro do ambiente interno simplifica o atendimento a regulamentações específicas.

Customização: Infraestrutura adaptada às necessidades específicas da instituição.

⊗ DESVANTAGENS:

Altos Custos Iniciais: Investimentos significativos em hardware, software e equipes especializadas.

Escalabilidade Limitada: Expansões podem ser lentas e custosas, dependendo da infraestrutura existente.

Manutenção e Suporte: Todas as atualizações e medidas de segurança dependem da capacidade interna da organização.

Fatores Críticos para a Decisão

Podemos resumir a tomada de decisão informada em uma série de sete fatores críticos:

01 Conformidade Regulatória: Avaliar restrições legais sobre localização e processamento de dados.

02 Escalabilidade: Considerar o crescimento futuro e a capacidade de adaptar-se rapidamente.

03 Segurança: Analisar controles necessários para proteger dados sensíveis.

04 Custos: Comparar custos iniciais e operacionais de longo prazo.

05 Expertise Interna: Avaliar a capacidade da equipe de TI para gerenciar infraestrutura de AI.

06 Integração: Considerar a compatibilidade com sistemas existentes.

07 Desempenho: Avaliar requisitos de latência e performance para aplicações críticas.

Tendências e Recomendações

Considerando as tendências atuais do setor financeiro e os benefícios oferecidos pela computação em nuvem, muitas instituições estão inclinando-se para soluções baseadas em nuvem ou adotando uma abordagem híbrida. A nuvem oferece vantagens significativas em termos de agilidade, escalabilidade e acesso a tecnologias de ponta, cruciais para manter a competitividade no cenário em rápida evolução da AI.

No entanto, para instituições com requisitos excepcionais de segurança, regulações muito específicas ou necessidade de controle total sobre a infraestrutura, uma solução on-premise ou híbrida pode ser mais apropriada. A abordagem híbrida, em particular, está ganhando popularidade como um passo intermediário, permitindo que as instituições aproveitem os benefícios da nuvem enquanto mantêm certos sistemas críticos on-premise.

Cuidados Especiais e Boas Práticas

A tabela abaixo apresenta cuidados especiais e boas práticas que devem ser observados na implantação de soluções de inteligência artificial tanto em ambientes de nuvem quanto em infraestruturas on-premise. Esses aspectos são fundamentais para garantir a segurança, a conformidade e a eficiência operacional, considerando as particularidades de cada abordagem tecnológica.

DIMENSÃO	SOLUÇÕES EM NUVEM	SOLUÇÕES ON-PREMISE
Proteção de Dados	Criptografia de ponta a ponta (dados em trânsito e em repouso)	Segurança física de servidores e infraestrutura
Gerenciamento de Acesso	Autenticação multifator e princípio do menor privilégio	Políticas de controle de acesso físico e lógico
Auditoria e Monitoramento	Auditorias regulares e ferramentas de monitoramento em tempo real	Sistemas de detecção de intrusão e monitoramento interno
Resiliência e Backup	Estratégias robustas de redundância e recuperação de desastres	Planos de continuidade de negócios testados regularmente
Atualizações e Vulnerabilidades	Dependência de SLAs e atualizações do provedor de nuvem	Atualizações frequentes e varreduras de vulnerabilidades realizadas internamente
Segmentação e Isolamento	Segmentação lógica via políticas de rede e identidade	Segmentação física e lógica da rede para proteção de sistemas críticos
Compliance Regulatório	Garantia contratual com provedores para atender regulações financeiras	Conformidade diretamente gerida e auditada pela equipe interna
Capacitação da Equipe	Treinamento contínuo em práticas seguras para ambientes cloud	Capacitação para administração segura da infraestrutura e resposta a incidentes locais



Para soluções em nuvem:

Criptografia de Dados: Utilize criptografia de ponta a ponta para dados em trânsito e em repouso.

Gerenciamento de Acesso: Implemente controles de acesso rigorosos, incluindo autenticação multifator e princípio do menor privilégio.

Auditoria Regular: Realize auditorias frequentes de segurança e conformidade.

Monitoramento Contínuo: Utilize ferramentas de monitoramento em tempo real para detectar atividades suspeitas.

Redundância e Backup: Implemente estratégias de backup e recuperação de desastres robustas.

Conformidade Contratual: Assegure que os contratos com provedores de nuvem atendam aos requisitos regulatórios do setor financeiro.

Treinamento de Equipe: Capacite a equipe em práticas de segurança específicas para ambientes em nuvem.

Para Soluções On-Premise:

Segurança Física: Implemente medidas rigorosas de segurança física para proteger servidores e infraestrutura.

Atualizações Regulares: Mantenha sistemas e software atualizados com as últimas patches de segurança.

Segmentação de Rede: Utilize segmentação de rede para isolar sistemas críticos.

Monitoramento Interno: Implemente sistemas de detecção de intrusão e monitoramento de atividades internas.

Plano de Continuidade: Desenvolva e teste regularmente planos de continuidade de negócios.

Gerenciamento de Vulnerabilidades: Realize varreduras regulares de vulnerabilidades e corrija prontamente quaisquer problemas identificados.

Controle de Acesso Físico e Lógico: Implemente políticas estritas de controle de acesso tanto para acesso físico quanto lógico aos sistemas.

[CONSULTE NO APÊNDICE](#)

Framework de Implementação
Técnica (p. 89)

Da Estratégia à Arquitetura: sua jornada rumo à AI começa aqui!

Aprofunde sua visão e prepare sua organização para a adoção prática e segura de Agentic AI e GenAI em ambientes on-premise. Criamos um framework técnico completo, com recomendações detalhadas para transformar intenção em ação com responsabilidade, escalabilidade e governança.

CONTEÚDO TÉCNICO

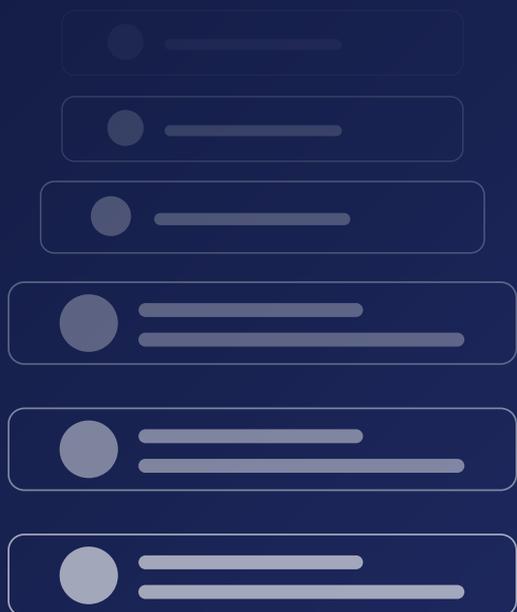
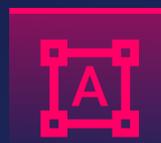
Framework para a Implementação de **Agentic AI e GenAI on-premise**





CAPÍTULO 7

Framework Estratégico para Implementação de **Agentic AI** por Área de Negócio



50K

40K

30K

20K

10K

Jan

Feb

Mar

Apl

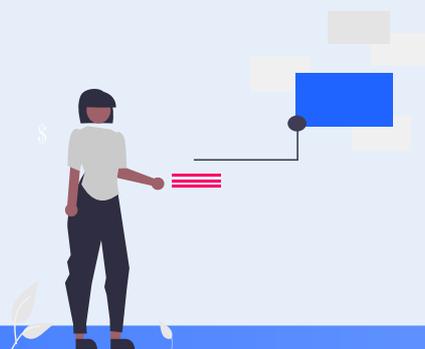
Mai



Introdução: Da Visão à Ação Estratégica

Os capítulos anteriores demonstraram o potencial transformador da Inteligência Artificial (AI), culminando na emergência da Agentic AI – sistemas capazes não apenas de analisar e gerar insights, mas de agir autonomamente para executar tarefas complexas e atingir objetivos definidos. Enquanto GenAI revolucionou a criação de conteúdo e a interação, a Agentic AI representa o próximo salto: a automação cognitiva e a execução proativa.

No entanto, liberar o verdadeiro potencial da Agentic AI no diversificado cenário financeiro exige mais do que adoção tecnológica; demanda uma estratégia direcionada e específica para cada área de negócio. Seja no Banco de Varejo, Banco de Investimento, Seguros ou Pagamentos (conforme detalhado na Tabela da página 23), os desafios, oportunidades e imperativos regulatórios variam significativamente.



Este capítulo fornece um framework estratégico para guiar as instituições financeiras na construção de um roadmap de implementação de Agentic AI sob medida para suas unidades de negócio específicas. O objetivo é capacitar os líderes a fazerem as perguntas certas, considerarem as dimensões críticas e estruturarem uma abordagem que maximize o valor e mitigue os riscos inerentes a essa poderosa tecnologia, inspirando-se nos exemplos concretos já possíveis.



O Imperativo Estratégico: Desbloqueando o Potencial da Agentic AI nas Quatro Frentes Chave

A Agentic AI representa uma mudança fundamental, capacitando sistemas não apenas a analisar, mas a agir de forma autônoma e coordenada para atingir objetivos complexos. O imperativo estratégico para o setor financeiro reside em explorar como essa capacidade de ação transforma radicalmente as quatro frentes chave de valor, conforme categorizado (e exemplificado inicialmente) na Tabela da p. 34. O potencial vai muito além dos exemplos atuais:

01 Automação de Processos – Rumo à Orquestração Inteligente e Adaptativa:

A Agentic AI eleva a automação da simples execução de tarefas para a orquestração dinâmica de fluxos de trabalho ponta-a-ponta. O horizonte futuro envolve agentes que não só seguem processos (como concessão de crédito no Varejo ou gestão de sinistros em Seguros), mas que os otimizam em tempo real, aprendem com exceções e negociam autonomamente dentro de parâmetros seguros. Imagine agentes gerenciando interações complexas entre instituições em Pagamentos ou automatizando a resolução de exceções em ciclos de trade no Banco de Investimento, adaptando o processo dinamicamente às condições e informações emergentes.

02 Gestão de Riscos – Da Resposta à Prevenção e Adaptação Autônoma:

A Agentic AI eleva a automação da simples execução de tarefas para a orquestração dinâmica de fluxos de trabalho ponta-a-ponta. O horizonte futuro envolve agentes que não só seguem processos (como concessão de crédito no Varejo ou gestão de sinistros em Seguros), mas que os otimizam em tempo real, aprendem com exceções e negociam autonomamente dentro de parâmetros seguros. Imagine agentes gerenciando interações complexas entre instituições em Pagamentos ou automatizando a resolução de exceções em ciclos de trade no Banco de Investimento, adaptando o processo dinamicamente às condições e informações emergentes.

03 Atendimento, Engajamento e Personalização – O Agente como Navegador Financeiro

Proativo:

A Agentic AI possibilita um nível de engajamento verdadeiramente proativo e holístico, muito além das recomendações personalizadas atuais. O futuro contempla agentes no Varejo gerenciando o ciclo de vida financeiro completo do cliente, antecipando necessidades e otimizando produtos de forma integrada. Em Seguros, agentes poderão usar dados (com consentimento) para sugerir proativamente ajustes de cobertura ou ações preventivas. No Banco de Investimento (Wealth Management), agentes poderão executar autonomamente otimizações de portfólio alinhadas aos objetivos de longo prazo do cliente. Em Pagamentos, agentes poderão atuar como gestores financeiros autônomos, otimizando gastos e poupanças proativamente.

04 Tomada de Decisão e Insights Baseados em AI – Da Análise à Iniciativa Estratégica

Autônoma:

A fronteira mais avançada da Agentic AI reside na autonomia decisória em domínios estratégicos complexos. Isso inclui agentes no Banco de Investimento capazes de desenvolver, testar e executar novas estratégias de trading. No Varejo e em Seguros, agentes poderão realizar otimização dinâmica de preços e subscrição, respondendo autonomamente às condições de mercado. Em Pagamentos, agentes poderão analisar vastos conjuntos de dados para identificar e até mesmo iniciar a exploração de novos modelos de negócio ou parcerias estratégicas, transformando insights diretamente em ação estratégica.

Portanto, ao definir a estratégia de Agentic AI para sua área de negócio, é crucial olhar para os exemplos atuais (como os da Tabela p. 34) como inspiração e validação, mas mirar no vasto potencial ainda inexplorado. A questão não é apenas "Como podemos automatizar o que fazemos hoje?", mas sim "Que novas capacidades de ação, otimização e criação de valor a autonomia dos agentes nos permite alcançar amanhã?". As instituições que responderem a essa pergunta de forma visionária e estratégica liderarão a transformação do setor financeiro.



Dimensões Chave da Estratégia de Agentic AI

Para construir um roadmap robusto, a estratégia de implementação de Agentic AI para uma área de negócio específica deve abordar as seguintes dimensões interconectadas:

01 Alinhamento Estratégico e Proposição de Valor: Conectar a Agentic AI aos objetivos de negócio fundamentais da área e definir claramente o valor esperado.

02 Identificação e Priorização de Casos de Uso: Mapear processos e oportunidades onde a autonomia e capacidade de ação dos agentes trarão maior impacto, inspirado pelos exemplos existentes e visando o potencial futuro.

03 Prontidão Tecnológica e de Dados: Avaliar e planejar a infraestrutura, dados, APIs e ferramentas necessárias para suportar operações agenticas.

04 Governança, Risco e Conformidade (GRC): Estabelecer os limites éticos, regulatórios e de segurança para a operação autônoma dos agentes.

05 Capacitação Humana e Gestão de Mudança: Preparar as equipes para colaborar, supervisionar e confiar nos novos sistemas agenticos.

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.



Perguntas Estratégicas Essenciais por Dimensão

Para cada dimensão, os líderes da área de negócio devem buscar respostas claras e específicas:

01 Alinhamento Estratégico e Proposição de Valor:

Quais são os 2-3 principais objetivos estratégicos da nossa área de negócio (ex: reduzir custos operacionais em X%, aumentar a retenção de clientes em Y%, lançar Z novos produtos) para os próximos 1-3 anos?

Como a autonomia e a capacidade de ação da Agentic AI (indo além da análise da GenAI) podem acelerar diretamente o alcance desses objetivos, considerando as quatro frentes chave (Automação, Risco, Atendimento, Decisão)?

Qual é a proposição de valor única da Agentic AI para nossos clientes e operações internas nesta área específica? Qual problema ela resolve melhor do que as soluções atuais, especialmente pensando no potencial "Além do Horizonte"?

Como a Agentic AI se encaixa na nossa estratégia digital e de dados mais ampla?

02 Identificação e Priorização de Casos de Uso:

Quais processos ponta-a-ponta em nossa área (envolvendo múltiplas etapas, sistemas e decisões) são mais adequados para serem orquestrados por agentes autônomos? (Consulte a Tabela da p. 34 para exemplos de partida e categorias como Automação de Processos, Gestão de Riscos, Atendimento/Personalização e Tomada de Decisão/Insights).

Dentro da nossa área de negócio específica (Varejo, Investimento, Seguros ou Pagamentos), quais dos exemplos "Além do Horizonte" descritos na seção "O Imperativo Estratégico: Desbloqueando o Potencial da Agentic AI nas Quatro Frentes Chave" p.74 representam as maiores oportunidades transformacionais ou resolvem as dores mais profundas?

Onde a capacidade de um agente interagir com interfaces (GUI), APIs, bases de dados diversas, e executar ações baseadas nessas interações e em aprendizado contínuo, traria maior ganho de eficiência, redução de erros ou criação de novo valor?

Quais tarefas cognitivas complexas, que exigem julgamento e adaptação, poderiam ser aumentadas ou eventualmente delegadas a agentes especializados, liberando expertise humana para supervisão estratégica e inovação?

Como priorizar os casos de uso identificados (atuais e futuros) com base no impacto potencial (valor transformacional), complexidade de implementação, riscos associados (incluindo riscos de autonomia) e alinhamento estratégico? (Usar matriz Risco x Impacto x Complexidade - Cap 3, p.26).

03 **Prontidão Tecnológica e de Dados:**

Nossos sistemas core (CRM, ERP, plataformas de negociação, etc.) possuem APIs robustas, acessíveis e seguras para que os agentes possam interagir e executar ações? Existem lacunas?

Temos acesso a dados (estruturados e não estruturados) de alta qualidade, atualizados, relevantes e suficientemente diversos para treinar agentes capazes de generalizar e operar eficazmente em cenários complexos? (Ref. Convicção, 2 cap 5)

Qual a nossa arquitetura preferencial (Nuvem, On-Premise, Híbrida - Capítulo 6 - Segurança e Privacidade de Dados como uma Prioridade Estratégica) para hospedar e operar esses agentes, considerando segurança, latência, soberania de dados e conformidade?

Quais ferramentas de orquestração de agentes (ex: LangChain, CrewAI - Cap 2), monitoramento, observabilidade e controle (para intervenção humana) precisamos implementar?

04 Governança, Risco e Conformidade (GRC):

Como definiremos os "guard rails" éticos e operacionais para os agentes autônomos nesta área? Quais ações são permitidas, quais exigem supervisão "humano-no-loop", e quais são estritamente proibidas? (Ref. Cap 5 e Framework de Governança p.53)

Quais requisitos regulatórios específicos (ex: KYC, AML, Basileia, Solvência II, LGPD/GDPR, regras de suitability) se aplicam às ações autônomas que os agentes executarão? Como garantiremos a conformidade contínua e auditável?

Como garantiremos a auditabilidade, explicabilidade (XAI onde aplicável) e rastreabilidade das decisões e ações tomadas pelos agentes autônomos?

Qual o processo para lidar com erros, vieses, comportamentos inesperados ou "drift" do modelo nos agentes? Quem é responsável pela intervenção, retreinamento e descomissionamento, se necessário?

05 Capacitação Humana e Gestão de Mudança:

Quais novas habilidades serão necessárias para nossas equipes projetarem, treinarem, supervisionarem, validarem e colaborarem efetivamente com agentes de AI autônomos? (Ref. Framework de Change Management p.60)

Como comunicaremos a introdução da Agentic AI e gerenciaremos o impacto nos papéis e responsabilidades existentes, enfatizando a colaboração humano-máquina?

Que programas de treinamento e aculturação (Ref. p.59 “Engajamento das Partes Interessadas” do Capítulo 6) são necessários para construir confiança, desmistificar a tecnologia e garantir a adoção eficaz por todas as partes interessadas?

Como fomentaremos uma cultura de experimentação controlada, aprendizado contínuo e responsabilidade compartilhada na operação da Agentic AI?

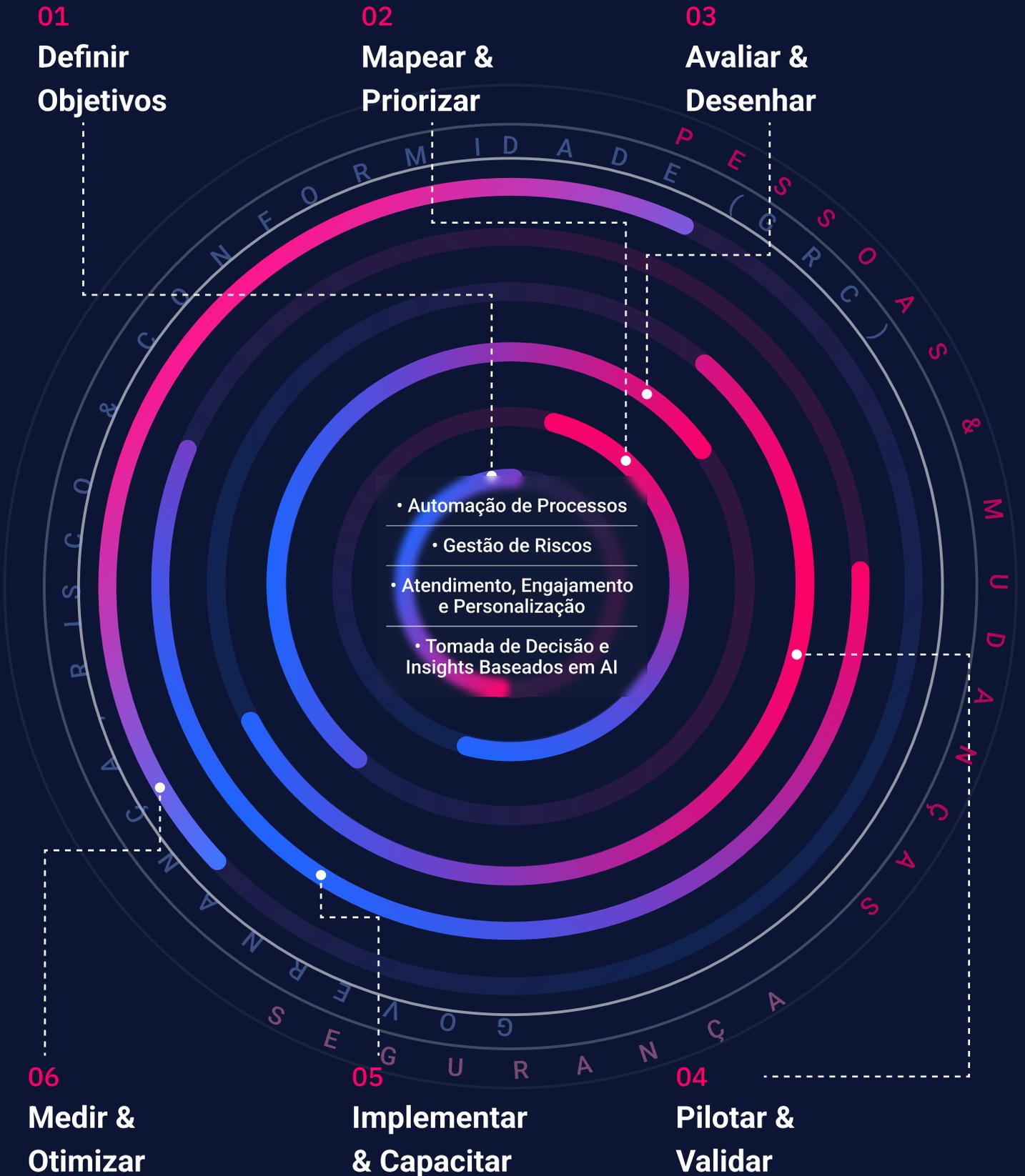
06 Mensuração de Desempenho e Melhoria Contínua:

Quais KPIs específicos medirão o sucesso da implementação da Agentic AI em relação aos objetivos de negócio definidos (ex: tempo de ciclo ponta-a-ponta, custo por processo automatizado, NPS, taxa de erro em decisões autônomas, ROI, novas receitas geradas)? (Ref. Convicção 6 e Cap 5)

Como implementaremos mecanismos de feedback robustos (humano e automatizado) para refinar continuamente o desempenho, o comportamento e a segurança dos agentes? (Ref. Feedback Loops, Cap 2 e Cap 3)

Com que frequência revisaremos o desempenho dos agentes, a eficácia da estratégia geral e o alinhamento com o ambiente regulatório e de mercado, ajustando o roadmap conforme necessário? (Ref. Abordagem Ágil, p.61)

Framework Visual: O Ciclo Estratégico de Implementação de Agentic AI



Adaptando o Framework por **Área de Negócio** (com exemplos da Tabela p. 34)

Embora o framework seja geral, a ênfase em cada etapa e as respostas às perguntas variarão, como exemplificado pelas ideias base (Tabela da p. 34) e potencial futuro (p. 74)



Banco de Varejo

Foco intenso em automatizar a jornada do cliente e crédito, personalizar interações e recomendações proativas, e gerenciar riscos operacionais e de inadimplência de forma preditiva. GRC crítico em dados do consumidor (LGPD/GDPR) e explicabilidade para crédito.



Seguros

- Casos de uso chave em automação ponta-a-ponta e otimização de sinistros (incluindo negociação), subscrição dinâmica e preditiva, personalização proativa de apólices e gestão de risco/prevenção baseada em dados contínuos. GRC focado em dados do segurado, equidade na subscrição e prevenção de fraudes.



Pagamentos

- Foco extremo em automação da reconciliação/liquidação complexa, detecção/desmantelamento autônomo de fraudes sofisticadas, e suporte/gestão financeira proativa para clientes. A velocidade, precisão, segurança e resiliência dos agentes são primordiais, com GRC focado em segurança de transações e compliance regulatório (PSD2, etc.).



Banco de Investimento

Prioridade em automação de processos de trade complexos, inteligência de risco de mercado em tempo real com ação autônoma, gestão de relacionamento consultiva e ativa (incluindo execução discricionária) e identificação/execução proativa de oportunidades. Ênfase em GRC rigoroso, segurança cibernética e explicabilidade (XAI) para decisões de alto valor.

Conclusão: Navegando o Futuro Agentico com Estratégia

A implementação bem-sucedida da Agentic AI não acontecerá por acaso. Requer uma abordagem deliberada, estratégica e adaptada às nuances de cada área de negócio dentro da instituição financeira, olhando além das aplicações atuais para o potencial transformador da autonomia inteligente. Ao utilizar este framework para fazer as perguntas certas, avaliar a prontidão, priorizar oportunidades visionárias e, crucialmente, estabelecer uma governança robusta e uma cultura de aprendizado e responsabilidade, as organizações podem navegar com confiança nesta nova fronteira.

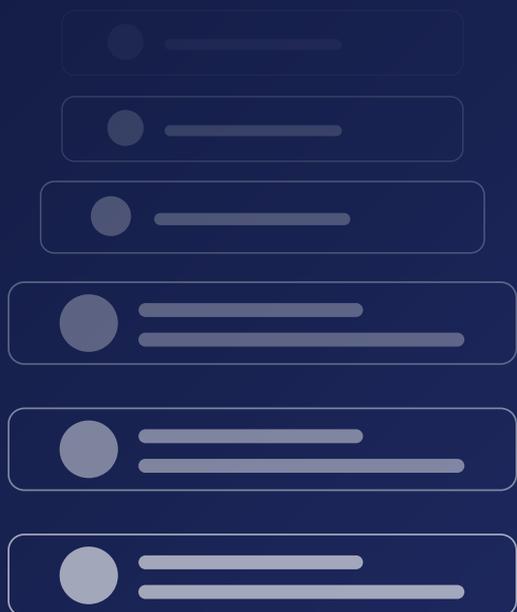
A Agentic AI oferece a promessa de transformar radicalmente a eficiência, a experiência do cliente, a gestão de riscos e a própria natureza da tomada de decisões financeiras. Uma estratégia bem definida, ancorada em possibilidades futuras e adaptada ao seu contexto específico, é a bússola necessária para transformar essa promessa em realidade tangível e sustentável, posicionando sua área de negócio e sua instituição na vanguarda da inovação financeira.





CAPÍTULO 8

Artefact como parceiro



Por que confiar na **Artefact**:



Experiência comprovada na indústria

Possuímos vasta experiência no setor Financeiro, entregando soluções práticas e resultados mensuráveis para as principais empresas líderes. Atuamos em todas as etapas da cadeia de valor, com ênfase em inovação, eficiência e crescimento sustentável.



Metodologia Estruturada e Personalizada

Nossa abordagem combina análises detalhadas com soluções sob medida, equilibrando visão estratégica e execução prática. Monitoramos continuamente os resultados para garantir o sucesso de nossos clientes.



Visão de AI como Unidade de Negócios

Incorporamos a visão de Dados e Inteligência Artificial como ativos geradores de valor, de forma que a tecnologia esteja a serviço do negócio. Com foco em resolver desafios das áreas de negócio, as aplicações devem prover resultados tangíveis e mensuráveis, seja com tomadas de decisão mais assertiva, otimização e automatização de processos e geração de insights acionáveis em tempo real, posicionando sua empresa à frente da concorrência.



Compromisso com Resultados Concretos e Sustentáveis

Nosso objetivo é entregar resultados sólidos e de longo alcance, promovendo inovação contínua e eficiência estratégica para posicionar sua empresa como referência no mercado.



Negócios que já **impactamos** com GenAI e Agentic AI



Forecast de Demanda e Abastecimento

Com a solução de previsão da Artefact, aumentamos a precisão e reduzimos o tempo necessário das equipes. Em apenas três meses conseguimos obter um ganho financeiro significativo.



Previsão de Rupturas de Estoque

Com a previsão de rupturas da Artefact, prevenimos riscos, aumentando a eficiência e assegurando receita ao longo de toda a cadeia de abastecimento em apenas cinco dias.

Alguns de nossos clientes



Aceleração de AI/Dados para BCEF

Estratégia de dados

Definição do modelo operacional da Fábrica de IA/Dados
Identificação de mais de 20 casos de uso (foco na redução de custos)

Fábrica de AI/Dados: Vários casos de uso entregues / em andamento:

Otimização da pontuação de crédito corporativo
Otimização da detecção de fraudes (+ milhões € de fraude adicional bloqueada)
CRM autônomo (+ milhões € PNB)
Agente conversacional GenAI (imposto sobre herança)

Fundamentos de dados

Construção de boas práticas de MLOps
Fornecimento de expertise na estruturação de plataformas de dados/AI

Aceleração de AI/Dados para RBF

Estratégia de dados

Definição da estratégia de monetização de dados e identificação de prioridades (programa BAAS)

Fábrica de AI/Dados: Desenvolvimento de múltiplos casos de uso

Painéis de controle (dashboards)
Agrupamento de clientes
Venda cruzada / aumento de vendas (upsell / cross sell)

Democratização de dados

Treinamento em larga escala de profissionais em dados aplicado a dois exemplos concretos: produção de dashboards e uso de dados ESG

Aceleração de AI/Dados

Aculturação: Treinamento para os 140.000 funcionários do grupo

- Implementação de conteúdo digital e microlearning cobrindo 7 principais temas de gestão de dados
- Desenvolvimento de um manual de gestão de dados
- Definição da estrutura do conteúdo de treinamento adaptada ao contexto da AXA
- Treinamento lançado a nível de grupo em 3 meses

Programa de marketing de dados

- Aceleração do marketing de dados para melhorar as vendas na rede francesa – ferramentas/organização/mensuração (em andamento)

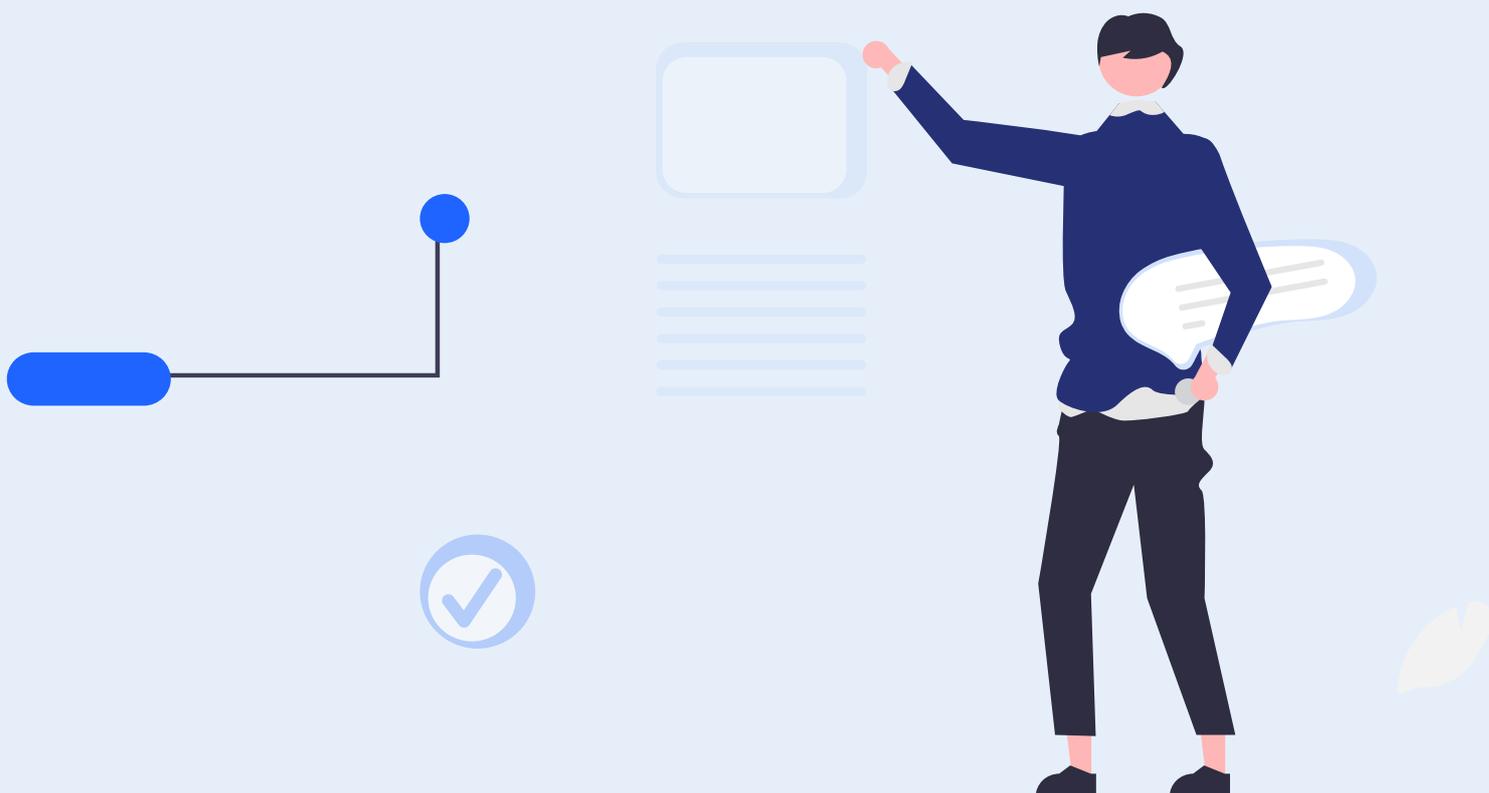


Apêndice

CONTEÚDO TÉCNICO

Framework para a Implementação de **Agentic** e **GenAI on-premise**

Neste apêndice exclusivo, reunimos os pilares técnicos essenciais para orientar uma jornada segura e eficiente na adoção de GenAI e Agentic AI no setor financeiro. Arquiteturas recomendadas, práticas de segurança, compliance e estratégias de escalabilidade – tudo adaptado ao contexto de soluções on-premise.

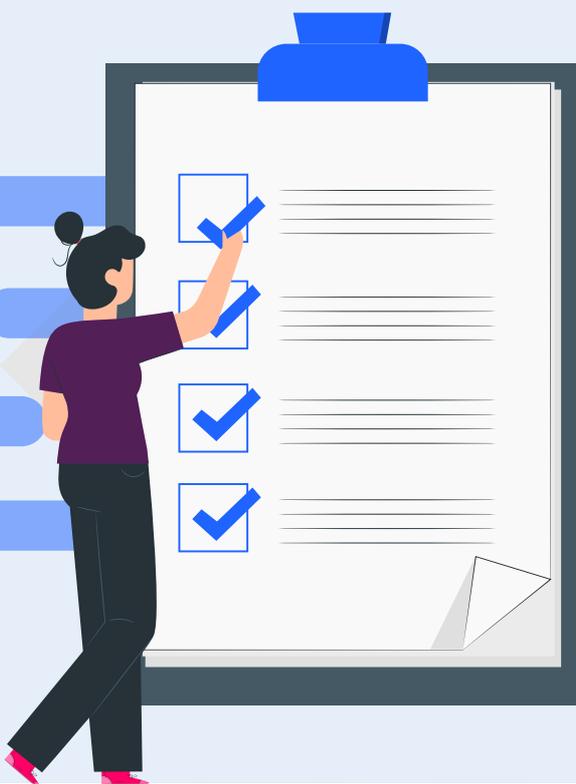


Framework para a Implementação de **Agentic e GenAI on-premise**

A adoção de soluções de GenAI e Agentic AI em ambientes on-premise oferece vantagens estratégicas importantes, especialmente no que diz respeito à segurança da informação, conformidade regulatória e controle sobre os recursos computacionais. No entanto, essa abordagem também impõe desafios técnicos e operacionais que exigem um planejamento detalhado.

Desde a definição da infraestrutura de hardware até as camadas de segurança, rede, armazenamento e manutenção, cada componente da arquitetura deve ser cuidadosamente alinhado à realidade do negócio, à sensibilidade dos dados processados e à frequência de uso dos modelos. Isso é ainda mais crítico quando se trata de automações autônomas e fluxos contínuos de decisão, como os viabilizados por agentes inteligentes.

Com base em experiências práticas e conhecimento técnico especializado, este capítulo apresenta os principais aspectos a serem considerados para a implementação bem-sucedida de soluções GenAI e Agentic AI em ambientes on-premise.





Necessidades de Hardware

No que tange ao poder computacional, projetos menores podem ser executados de forma eficaz em servidores com CPUs que possuam alta contagem de núcleos, alcançando um bom equilíbrio entre custo e desempenho. Porém, à medida que a demanda aumenta — seja por conta do volume de dados ou pela complexidade dos modelos utilizados —, torna-se recomendável o uso de GPUs especializadas (como as NVIDIA A100 ou H100) para reduzir drasticamente o tempo de treinamento. Assim, delineiam-se dois cenários típicos:



Demanda baixa: Uso de CPUs de alto throughput, com potencial de escalonamento gradual conforme a necessidade cresce.



Demanda alta: Infraestrutura baseada em GPUs de última geração, envolvendo clusters de múltiplos servidores e orquestração dinâmica (por exemplo, Kubernetes).

Em relação ao armazenamento, projetos iniciais costumam ser bem atendidos por soluções de menor complexidade, como SSDs, HDDs locais ou até um NAS (Network Attached Storage) de pequeno porte para compartilhamento de dados. Contudo, quando o volume de informações e a criticidade aumentam, surgem demandas mais elevadas por velocidade e tolerância a falhas. Nesse cenário, entram em cena sistemas de armazenamento distribuído (por exemplo, Ceph ou GlusterFS) e unidades NVMe SSD, que fornecem alto desempenho de I/O e resiliência. Novamente, pode-se destacar dois patamares de demanda:



Demanda baixa: SSDs locais ou um NAS simples, associados a políticas básicas de backup.



Demanda alta: Armazenamento distribuído com NVMe SSD, projetado para oferecer redundância, disponibilidade e escalabilidade.



No que se refere à rede, a disparidade entre demandas menores e operações mais complexas é bastante nítida. Para ambientes de menor porte, conexões de 1GbE ou 10GbE geralmente dão conta do recado, e, para isolamento básico, VLANs ou VPNs simples podem ser suficientes. Entretanto, quando se almeja alto desempenho e baixa latência (por exemplo, em cenários de HPC – High-Performance Computing), torna-se fundamental utilizar conexões a partir de 25GbE ou até mesmo Infiniband. Nessas situações, também se faz necessário investir em segmentação de rede avançada e firewalls robustos para preservar a segurança em grande escala:



Demanda baixa: Redes de 1GbE ou 10GbE e segmentação simples via VLAN/VPN.



Demanda alta: Conexões $\geq 25\text{GbE}$ ou Infiniband, com segmentação avançada e monitoramento contínuo do tráfego.

Ao se estruturar a arquitetura de IA on-premise sob essa perspectiva modular — diferenciando adequadamente cargas de trabalho menores das de grande porte —, a alocação de recursos torna-se mais precisa. Dessa forma, evita-se o superdimensionamento inicial e ainda se garante possibilidade de expansão sem necessidade de interromper serviços. Isso estabelece uma base sólida para projetos de GenAI e Agentic AI que priorizam segurança, desempenho e escalabilidade, além de criar um caminho para adaptações futuras.





Tech Stack

A definição de uma Tech Stack adequada constitui um passo fundamental para atingir o máximo desempenho e a escalabilidade desejada em projetos de IA on-premise. Abaixo, listam-se os principais componentes recomendados:

Sistema Operacional

Distribuições Linux como Ubuntu ou CentOS são indicadas para cargas de trabalho em AI, pois combinam **estabilidade, amplo suporte comunitário e facilidade de otimização de hardware avançado** (GPUs e CPUs de múltiplos núcleos).

Frameworks de AI

Ao escolher um framework de AI, alinhe sua decisão às necessidades do projeto, expertise da equipe e escalabilidade. Para deep learning, TensorFlow oferece escalabilidade, PyTorch flexibilidade e Keras facilidade de uso. Em machine learning tradicional, Scikit-Learn é intuitivo, enquanto XGBoost e LightGBM são ideais para dados estruturados. Para visão computacional, Caffe se destaca. Já frameworks como LangChain, OpenAI e Hugging Face são recomendados para LLMs e GenAI. Considere também suporte da comunidade, integração com sua stack e manutenção a longo prazo

Containerização, Orquestração e Empacotamento

O uso de containers, como Docker facilita a criação de ambientes isolados para cada modelo ou serviço, simplificando tanto o desenvolvimento quanto a implantação. Em paralelo, o Kubernetes é a escolha padrão para orquestração, devido à sua capacidade de gerenciar a escalabilidade horizontal e fornecer mecanismos de tolerância a falhas — fatores críticos em cenários com múltiplos serviços de AI executando-se simultaneamente.

Para maximizar o aproveitamento dos recursos de hardware e da pilha de software selecionada, algumas configurações são especialmente importantes:



Cluster Kubernetes com Suporte a GPU: A adoção do plugin de dispositivo Kubernetes da NVIDIA possibilita que as aplicações em contêiner façam uso dos recursos de GPU conforme a demanda, otimizando o treinamento e diminuindo o tempo de inferência.



Monitoramento de Recursos: Ferramentas como Prometheus e Grafana permitem o acompanhamento em tempo real de métricas de CPU, GPU, memória e rede, viabilizando a identificação de gargalos de performance e orientando melhorias contínuas.

Além disso, a containerização assegura portabilidade e isolamento do ambiente, simplificando tanto o desenvolvimento quanto a manutenção de diferentes versões de modelo. Ferramentas como Docker e plataformas de inferência (por exemplo, FastAPI ou Flask) ajudam a compor o ecossistema de implementação. Além de viabilizar o empacotamento do Modelo, já que ao criar a imagem Docker, todas as dependências necessárias são incluídas, desde bibliotecas de AI (PyTorch, TensorFlow, Transformers etc.) até o modelo em seu estado otimizado.

Exemplo de Dockerfile:

Abaixo, um exemplo ilustrativo de Dockerfile que instala as dependências básicas, copia o modelo para o contêiner e expõe o serviço de inferência:

Unset

- `# Exemplo 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"]`

Esse exemplo utiliza como base a imagem da NVIDIA com CUDA 11.8, instala as bibliotecas necessárias, copia os arquivos do modelo e do aplicativo para dentro do contêiner e inicia um servidor Uvicorn para expor a API de inferência.

Seleção do Modelo: Para manter total controle sobre a implantação e garantir adaptabilidade, recomenda-se utilizar modelos de código aberto, como Llama 3 ou Mixtral. Estes podem ser ajustados (fine-tuned) com dados proprietários, caso seja necessário adaptar o comportamento ou o desempenho para cenários específicos.



Modelos Open Source: Evitam a dependência excessiva de fornecedores e aumentam a transparência em relação ao funcionamento interno.



Ajuste Fino (Fine-tuning): Quando for preciso, aplique técnicas de fine-tuning sobre dados corporativos, preservando competitividade e relevância do modelo para o domínio de negócio em questão.

Otimização do Modelo: Após a seleção do modelo, é possível otimizá-lo para reduzir a latência de inferência e o consumo de recursos computacionais. Entre as estratégias usuais destacam-se a quantização e o paralelismo:



Quantização: Converter modelos para formatos de menor precisão, como FP16 ou INT8, diminui o uso de memória e melhora a velocidade de inferência, principalmente em GPUs de última geração.



Paralelismo: Em modelos muito grandes, o uso de paralelismo de modelo ou de tensor pode distribuir o processamento em vários nós ou GPUs, reduzindo o tempo total de execução.

Implementação: Depois de criar o contêiner, o passo seguinte é integrar o modelo ao ambiente Kubernetes. Isso pode ser feito via arquivos YAML ou gráficos Helm (Helm Charts). Em ambos os casos, aspectos essenciais devem ser observados:

Gerenciamento de Escalabilidade: O Horizontal Pod Autoscaler (HPA) ajusta automaticamente o número de réplicas de pods conforme a carga de trabalho varia. Ele monitora métricas como uso de CPU ou GPU e cria ou remove pods para manter o desempenho.

Monitoramento e Logs: Vincule o deployment a soluções de monitoramento (como Prometheus e Grafana) para rastrear métricas de uso de CPU, GPU, memória e latência em tempo real. Isso auxilia na identificação de gargalos e na tomada de decisões de escalabilidade e otimização.

Segurança e Conformidade

A segurança dos dados e a conformidade regulatória são pilares em projetos de AI on-premise, principalmente quando se manipulam informações sensíveis ou sob regulamentações rígidas. A seguir, apontam-se as principais práticas para proteger a integridade dos dados, adequar a implantação à legislação pertinente e reforçar a segurança de rede.

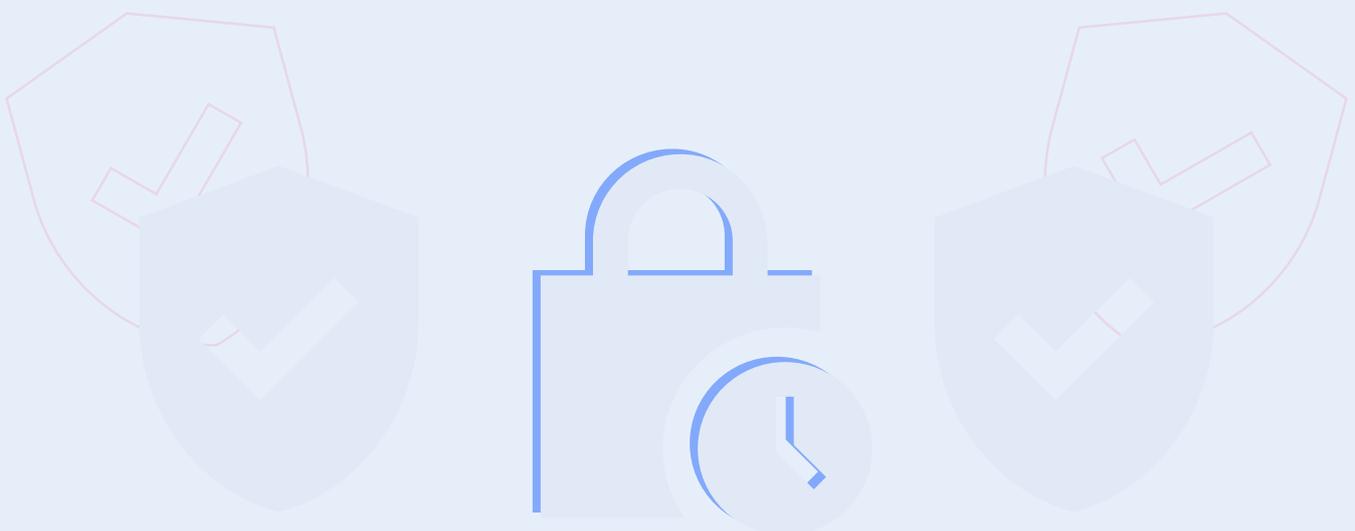
Proteção de Dados

Para manter dados corporativos em um ambiente seguro e controlado, recomenda-se uma abordagem em múltiplas camadas:

> **Armazenamento Criptografado:** É crucial implementar criptografia em repouso (at rest) e em trânsito (in transit), mantendo os dados ilegíveis mesmo em situações de acesso indevido.

> **Redes Privadas:** Restringir a comunicação do cluster a redes privadas internas, prevenindo a exposição de portas e serviços críticos ao ambiente externo.

> **Controle de Acesso Baseado em Função (RBAC):** Configurar Kubernetes com RBAC garante que cada função dentro da organização receba permissões específicas, evitando que usuários sem autorização tenham acesso a recursos confidenciais.



Conformidade

Dependendo do setor, as exigências legais e regulatórias podem variar consideravelmente. Cada projeto de AI deve alinhar-se às normas correspondentes, como:



GDPR: Voltada à proteção de dados pessoais de cidadãos da União Europeia, exige transparência no tratamento de dados.



HIPAA: Aplicável à saúde nos EUA, define padrões para proteger e manter a privacidade de informações médicas.



ISO 27001: Certificação internacional orientada à gestão de segurança da informação, contemplando políticas de controle, monitoramento e melhoria contínua.



LGPD: Lei brasileira que regula o tratamento de dados pessoais, garantindo direitos aos titulares e impondo responsabilidades às organizações quanto à coleta, uso e armazenamento de dados.

Esses referenciais guiam a criação de políticas internas robustas, garantindo confidencialidade, integridade e disponibilidade dos dados.

Segurança de Rede

A infraestrutura de rede em ambientes on-premise precisa ser configurada para conter ameaças externas e minimizar riscos internos:

Nuvem Privada Virtual

(VPC): Utilizar endereços IP privados e segmentar sub-redes críticas, de modo que apenas serviços confiáveis acessem o cluster de IA.

Firewalls e Sistemas de Detecção de Intrusão

(IDS): Definir regras de firewall para filtrar portas e protocolos sensíveis e aplicar IDS a fim de bloquear tentativas de invasão ou detectar tráfego anormal.

Monitoramento Contínuo:

Integrar ferramentas como Prometheus e Grafana para observar eventos de segurança em tempo real e gerar alertas em caso de atividades suspeitas.

Escalabilidade e Otimização de Desempenho

Projetos de Agentic AI ou Gen AI podem se expandir rapidamente em termos de volume de dados, número de usuários e complexidade de modelos. Para acompanhar esse crescimento sem comprometer a qualidade, é imprescindível aplicar práticas de escalabilidade e otimização de desempenho. Seguem três recomendações-chave:

Balanceamento de Carga

O balanceamento de carga distribui requisições uniformemente entre os recursos disponíveis, evitando sobrecarga. Ferramentas como NGINX ou Traefik podem ser configuradas para rotear tráfego de maneira eficiente em um cluster Kubernetes, aumentando a confiabilidade e a velocidade de resposta.



NGINX ou Traefik: Ajuste as regras de roteamento para que os pods com maior capacidade de processamento recebam mais requisições proporcionalmente.



Tolerância a Falhas: Em caso de queda de um nó, o tráfego é redirecionado sem prejudicar o serviço.

Cache

Mecanismos de cache são extremamente úteis quando determinadas inferências ou resultados de modelos são acessados com frequência. Ferramentas como Redis permitem armazenar em memória as respostas mais solicitadas, reduzindo consideravelmente o tempo de resposta.



Redução de Latência: Ao servir respostas diretamente da memória, não é preciso reprocessar o modelo a cada solicitação.



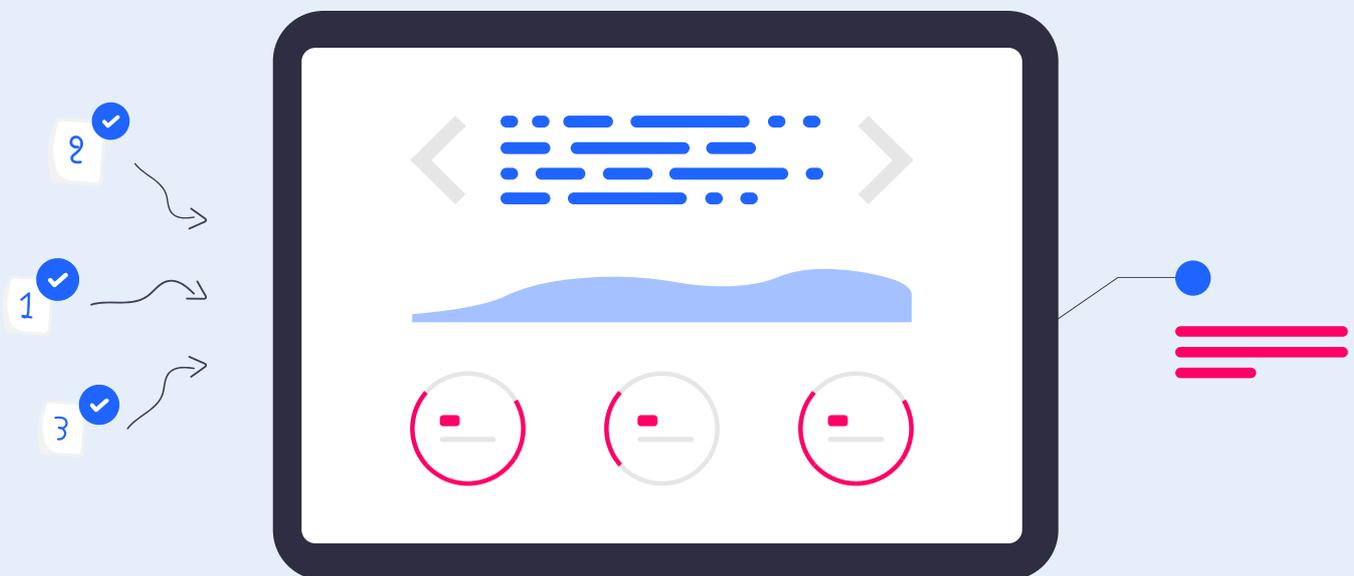
Diminuição da Carga: O cache reduz a pressão sobre os servidores de IA, liberando recursos para outras inferências ou tarefas mais complexas.

Paralelismo

Para lidar com grandes volumes de dados ou modelos de alta complexidade, o paralelismo viabiliza a divisão da carga de trabalho, acelerando a execução. Frameworks como Ray ou Dask simplificam a distribuição de tarefas de treinamento ou inferência em vários nós, ampliando a eficiência e diminuindo o tempo de processamento.

> **Frameworks de Computação Distribuída:** Ray e Dask oferecem APIs acessíveis para paralelizar operações sem a necessidade de lidar diretamente com a comunicação entre nós.

> **Escalonamento Horizontal:** Com o aumento do número de nós, o desempenho tende a crescer linearmente, possibilitando execução de workloads progressivamente maiores



Monitoramento e Manutenção

Para manter a eficiência e a disponibilidade de soluções de GenAI ou Agentic AI em um ambiente on-premise, é fundamental estabelecer procedimentos sólidos de monitoramento e manutenção. Essas práticas permitem a detecção de gargalos de desempenho, o acompanhamento de logs e a atualização dos modelos, garantindo que as soluções continuem atendendo às necessidades do negócio.

Ferramentas de Monitoramento

Uma arquitetura de monitoramento abrangente facilita a identificação e a resolução ágil de problemas. Prometheus e Grafana são frequentemente utilizados em conjunto para coletar métricas (CPU, memória, uso de GPU, latência etc.) e exibir painéis que retratam a saúde geral do sistema.

Em paralelo, uma stack de registros, como o ELK (Elasticsearch, Logstash, Kibana), unifica a gestão de logs em um único ponto. Isso simplifica o diagnóstico de incidentes, permitindo identificar rapidamente comportamentos anormais ou falhas, além de manter um histórico detalhado para auditoria ou futuras investigações.



Prometheus + Grafana: Monitoramento e visualização em tempo real das métricas mais críticas do cluster.



ELK Stack: Centralização e análise de logs, viabilizando correção de problemas e verificação de histórico de modo mais eficiente.



Retreinamento e Atualizações do Modelo

À medida que os dados de entrada mudam e surgem novos insights, manter o modelo atualizado é crucial para preservar a precisão e a relevância da solução de AI. Integrar pipelines de CI/CD (Continuous Integration e Continuous Delivery) ao processo de retreinamento possibilita automatizar etapas como carregamento de dados, verificação de performance do novo modelo e, eventualmente, implantação em produção.



argo



Jenkins

Ferramentas como ArgoCD ou Jenkins ajudam a controlar versões de modelo, liberando atualizações incrementais e executando rollbacks quando ocorrem inconsistências ou perdas significativas de desempenho. Esse ciclo de melhoria contínua acelera o desenvolvimento, minimiza riscos e garante que o modelo em produção permaneça adequado às demandas do negócio.



Pipelines de CI/CD: Permitem retreinamento frequente com dados atualizados e testes automatizados de validação.



Automação de Implantação: Soluções como ArgoCD ou Jenkins facilitam o delivery de modelos revisados, assegurando rastreabilidade e controle de versões.



Visão Geral do Fluxo de Implementação de Agentic AI / Gen AI On Premise

ETAPA	TÓPICO	SUBTÓPICO	POSSÍVEIS FERRAMENTAS/ FRAMEWORKS UTILIZADOS
Hardware	Poder Computacional	Baixa Demanda	CPUs de alto throughput
		Alta Demanda	GPUs de última geração (Exemplo: Kubernetes)
	Armazenamento	Baixa Demanda	SSDs locais ou um NAS simples
		Alta Demanda	NVMe SSD
	Rede	Baixa Demanda	Redes de 1GbE ou 10GbE (segmentação VLAN/VPN)
		Alta Demanda	Conexões ≥25GbE ou Infiniband
Tech Stack	Sistema Operacional	---	Linux (Ubuntu ou CentOS)
	Frameworks	AI	PyTorch, TensorFlow ou Keras
		Gen AI	LangChain, OpenAI e Hugging Face
		ML	Scikit-Learn, XGBoost ou LightGBM
	Containerização, Orquestração e Empacotamento	---	Docker, Kubernetes
	Seleção do Modelo	---	Llama 3 ou Mixtral
	Otimização do Modelo	---	Quantização e/ou Paralelismo
Implementação	---	Arquivos YAML ou Gráficos Helm (Helm Charts)	
Segurança e Conformidade	Proteção de Dados	---	RBAC
	Conformidade	---	---
	Segurança de Rede	---	VPC
Escalabilidade e Otimização de Desempenho	Balanceamento de Carga	---	NGINX ou Traefik
	Cache	---	Redis
	Paralelismo	---	Ray ou Dask
Monitoramento e Manutenção	Ferramentas de Monitoramento	---	Prometheus + Grafana ou ELK Stack
	Retreinamento e Atualizações do Modelo	---	ArgoCD ou Jenkins

Glossário

A - G

Agentic AI

Um tipo de inteligência artificial que adiciona autonomia e capacidade de tomada de decisão aos sistemas, permitindo que ajam de forma proativa, aprendam com a experiência e executem tarefas complexas com mínima intervenção humana.

Análise Preditiva

Técnica de análise de dados que utiliza estatísticas e algoritmos para prever tendências e padrões futuros com base em dados históricos.

API (Application Programming Interface)

Conjunto de regras e ferramentas que permite a integração de aplicações com sistemas de AI, facilitando comunicação e troca de dados.

Automação de Processos

Uso de tecnologia para realizar tarefas repetitivas e operacionais de forma eficiente, reduzindo a necessidade de intervenção humana.

Chatbot

Programa baseado em AI que interage com usuários em linguagem natural para responder perguntas, fornecer informações ou realizar ações automatizadas.

Compliance

Conjunto de normas e regulamentos que as empresas devem seguir para garantir a conformidade legal e regulatória, especialmente no setor financeiro.

Criptografia

Método de proteção de dados por meio de codificação, garantindo segurança e privacidade nas transações digitais.

Dados Não Estruturados

Informações que não seguem um formato predefinido, como textos, imagens, áudios e vídeos, tornando seu processamento mais complexo.

Detecção de Fraudes

Uso de inteligência artificial e análise de dados para identificar atividades suspeitas e prevenir crimes financeiros, como lavagem de dinheiro e fraudes bancárias.

GenAI (Inteligência Artificial Generativa)

Subcampo da AI focado na criação de novos conteúdos, como texto, imagens, código e áudio, a partir de modelos treinados em grandes volumes de dados.

Governança de Dados

Conjunto de processos e políticas que garantem a qualidade, segurança e conformidade do uso de dados dentro de uma organização.

Modelos de Linguagem

Algoritmos de AI treinados para entender, processar e gerar texto com base em um vasto conjunto de dados de linguagem natural.

GUI (Graphical User Interface)

Interface visual que facilita a interação com IA agentiva, usando elementos gráficos como botões, menus e painéis para navegação intuitiva.

Onboarding de Clientes

Processo de integração de novos clientes em uma instituição financeira, garantindo que atendam a requisitos regulatórios e tenham acesso a produtos e serviços personalizados.

Inteligência Artificial (AI)

[Artificial Intelligence] Campo da ciência da computação que desenvolve sistemas capazes de realizar tarefas que normalmente exigem inteligência humana, como aprendizado, tomada de decisões e reconhecimento de padrões.

Personalização Avançada

Uso de AI para adaptar produtos, serviços e experiências com base nas necessidades e preferências individuais dos clientes.

LLMs (Large Language Models)

Modelos de Linguagem de Grande Escala que utilizam redes neurais avançadas para compreender e gerar linguagem natural, permitindo aplicações como chatbots e assistentes virtuais.

RAG (Retrieval-Augmented Generation)

Técnica utilizada em sistemas de inteligência artificial generativa, onde o modelo combina geração de texto com recuperação de informações relevantes de uma base de dados ou documentos externos para aumentar a precisão e relevância das respostas.

Machine Learning (Aprendizado de Máquina)

Técnica de AI que permite que sistemas aprendam a partir de dados e melhorem seu desempenho sem necessidade de programação explícita.

RPA (Robotic Process Automation)

Tecnologia que usa softwares para automatizar processos empresariais repetitivos, aumentando a eficiência operacional.

RPA (Robotic Process Automation)

Tecnologia que usa softwares para automatizar processos empresariais repetitivos, aumentando a eficiência operacional.

Transformação Digital

Processo de integração de tecnologias digitais em todas as áreas de uma empresa para melhorar operações, eficiência e experiência do cliente.

Segurança de Dados

Práticas e tecnologias utilizadas para proteger informações sensíveis contra acessos não autorizados, vazamentos ou ataques cibernéticos.

XAI (Explainable AI)

A Inteligência Artificial Explicável é uma corrente da Inteligência Artificial que busca trazer compreensão aos resultados dos modelos e sistemas inteligentes, de maneira a justificar e viabilizar a auditabilidade dos processos com AI embedada.

Links & Referências

[1] Artigo

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] Pesquisa

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

[3] Material Institucional

Artefact. Conhecimento interno.

[4] Pesquisa

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

[5] Artigo

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

[6] Artigo

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] Vídeo | Evento

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

[8] Report

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] Artigo

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/>

Materiais técnicos e estudos internos

Compilado a partir de MVPs e demonstrações internas de IA generativa aplicadas ao setor financeiro, com foco em automação de documentos, APIs conversacionais, chatbots e copilotos.

Artefact – One Pagers e protótipos funcionais aplicados:

DraftAI – MVP / Atendimento com ChatBot Cognitivo / Bot conversacional com integração bancária / Business Intelligence Agente Aumentado / Processamento Inteligente de Documentos.

Estudos sobre GenAI e NLP aplicados em:

Mapeamento de riscos emergentes / Compreensão de chamadas com NLP / Deep dive em documentação de Private Equity / Aceleração do processo de Due Diligence / Automação no crédito via Intelligent Document Processing.