

WHITE PAPER · 11 DE MAY DE 2026

Como escolher a plataforma de agentes de IA da sua empresa em 2026

Inteligência operacional, lock-in estratégico e o mapa das principais plataformas

Para · CEOs e Founders B2B de Serviços, Brasil

Data · 11 de May de 2026

Como escolher a plataforma de agentes de IA da sua empresa em 2026

Cliente: TheAgent

Audiência: CEOs/Founders de empresas B2B de serviços, 50-500 funcionários, Brasil

Escopo: Brasil, serviços B2B, 2024-2026

Data: Maio de 2026

Frameworks aplicados: Porter's Five Forces · Technology Adoption Lifecycle (Moore) · VRIO (Barney)

01. Executive Summary

Em maio de 2026, OpenAI e Anthropic anunciaram joint ventures de prestação de serviços enterprise, entrando diretamente no mercado de implementação de IA operado pelos mesmos clientes que usavam suas APIs. Esse evento transforma uma questão que parecia técnica — qual plataforma de agentes de IA escolher — em uma decisão estratégica com consequências de dois anos.

A tese deste white paper: A escolha de plataforma de agentes de IA não é uma decisão técnica. É uma aposta estratégica sobre onde vai estar a inteligência operacional da empresa nos próximos dois anos — e essa aposta fica mais arriscada quando o fornecedor do modelo decide prestar o mesmo serviço que você entrega para os clientes que conquistou.

O que os dados mostram: O mercado de agentes de IA cresce 44% ao ano e deve dobrar para USD 10,8 bilhões até 2026. Sete players concentram 56% do mercado. O custo médio de migração entre plataformas é de USD 315.000, e apenas 25% das empresas que migraram atingiram o valor esperado no primeiro ano. No Brasil, a adoção de IA por empresas industriais cresceu 148% entre 2022 e 2024 — mas 72% das empresas ainda estão no início da jornada, tomando decisões de plataforma sem critério estratégico.

O padrão que se repete: Amazon usou dados de sellers para concorrer com eles. Apple capturou margem de desenvolvedores. Salesforce adquiriu parceiros de sucesso. OpenAI e Anthropic, em maio de 2026, lançaram braços de serviços enterprise. O conceito tem nome acadêmico — *platform envelopment* — e trajetória documentada.

O que distingue as empresas que ganham: A Cresta construiu USD 100 milhões de ARR não pelo modelo de IA que usa — um modelo de terceiros — mas pela inteligência operacional proprietária acumulada sobre ele. A vantagem competitiva sustentável está nos dados raros, difíceis de imitar, organizados para capturar valor — não na plataforma que qualquer concorrente também pode contratar.

O mapa de plataformas disponíveis — analisado na Seção 07b — mostra que há uma correlação quase direta entre facilidade de uso (no-code) e lock-in estratégico. Flowise, Dify e n8n (self-hosted, low-code) oferecem o ponto de entrada mais equilibrado para PMEs brasileiras: controle proprietário dos dados sem exigir time de engenharia especializado. Stack AI, Relevance AI, Copilot Studio e AWS Bedrock elevam o lock-in progressivamente em troca de conveniência — um custo que precisa ser explicitado antes da assinatura.

Cinco recomendações concretas:

1. Mapear dependências de IA existentes na empresa (90 dias).
2. Incluir quatro critérios estratégicos em toda avaliação de plataforma (imediato).
3. Definir arquitetura de armazenamento da inteligência operacional proprietária (Q3/2026).
4. Exigir cláusula de sobreposição de mercado em contratos com fornecedores de IA (imediato).
5. Estabelecer critério de maturidade antes de escalar qualquer solução de agentes (próximo piloto).

02. O Momento

Em maio de 2026, dois eventos redefiniram o tabuleiro. A Anthropic — criadora do Claude, o modelo de IA que alimenta dezenas de plataformas de agentes usadas por empresas de serviços B2B — anunciou uma joint venture de prestação de serviços enterprise, avaliada em USD 1,5 bilhão e financiada por Blackstone, Goldman Sachs e Hellman & Friedman. Uma semana depois, a OpenAI lançou "The Development Company", captando USD 4 bilhões de 19 investidores com declaração explícita de competir com as maiores consultorias do mundo — as mesmas consultorias que, até então, eram seus principais parceiros de go-to-market.

Tradução direta: os dois maiores fornecedores de modelos de linguagem do mundo decidiram, no mesmo mês, prestar os mesmos serviços que seus clientes entregam.

Isso não é especulação. É fato com nome, valor de mercado e fundadores identificados.

Para o CEO de uma empresa brasileira de serviços B2B que está avaliando qual plataforma de agentes de IA adotar — ou que já adotou — esse momento muda o cálculo. A decisão que parecia técnica revela-se estratégica: uma aposta sobre onde estará a inteligência operacional da empresa nos próximos dois anos, e sobre com quem ela vai competir.

O mercado global de agentes de IA movimentou USD 5,4 bilhões em 2024 e deve dobrar para USD 10,8 bilhões até 2026, crescendo 44% ao ano (Fortune Business Insights / GM Insights, 2025). No Brasil, o investimento em automação inteligente deve atingir R\$ 12,4 bilhões em 2026 (TI Inside, 2026). O crescimento da adoção de IA por empresas industriais brasileiras foi de 148% entre 2022 e 2024 — de 16,9% para 41,9% das empresas (IBGE / Pintec Semestral, 2024).

Esse crescimento está acontecendo sem que a maioria dos CEOs tenha feito a pergunta correta. A janela para fazer essa pergunta — e ainda ter tempo de agir — é agora.

03. A Tese

A escolha de plataforma de agentes de IA não é uma decisão técnica. É uma aposta estratégica sobre onde vai estar a inteligência operacional da empresa nos próximos dois anos — e essa aposta fica mais arriscada quando o fornecedor do modelo decide prestar o mesmo serviço que você entrega para os clientes que conquistou.

Três premissas sustentam essa afirmação.

Primeira: A plataforma de agentes não é uma ferramenta intercambiável. Cada interação processada por um agente, cada workflow aprovado ou corrigido, cada exceção tratada, cada padrão de comunicação aprendido — tudo isso é inteligência operacional. Quando essa inteligência fica armazenada dentro de um sistema controlado por um fornecedor externo, a empresa não está acumulando um ativo. Está financiando a vantagem competitiva do seu fornecedor.

Segunda: O custo de uma decisão errada não é abstrato. O custo médio de migração entre plataformas de tecnologia empresarial é de USD 315.000 por projeto — e 57% dos líderes de TI reportam ter gasto mais de USD 1 milhão em migrações no último ano (The Register / CloudBees, 2025-2026). Apenas 25% das empresas que concluíram migrações de plataforma nos últimos dois anos atingiram o valor esperado no primeiro ano (CloudBees, 2025). O custo não é só financeiro: 70% das empresas reportam aumento de burnout entre desenvolvedores durante migrações, e 61% dizem que suas equipes evitaram novos projetos por mais de seis meses após a transição (CloudBees / Integrate.io, 2025).

Terceira: Os fornecedores de modelos estão subindo na cadeia de valor. Não como estratégia declarada — como fato consumado. Anthropic e OpenAI lançaram joint ventures de serviços enterprise em maio de 2026. O Google estruturou seu protocolo A2A (Agent-to-Agent) como infraestrutura corporativa nativa. A Microsoft integrou GPT-4 em toda sua stack — Office, Teams, Azure, Copilot — absorvendo verticalmente os casos de uso que startups de IA haviam estabelecido como mercados independentes.

O argumento não é que toda empresa de serviços B2B será destruída por OpenAI. O argumento é que as condições para esse cenário — dependência de plataforma, dados operacionais no controle do fornecedor, custo de migração proibitivo — estão sendo criadas agora, nas decisões que parecem pequenas e técnicas.

04. Evidências Quantitativas

O MERCADO ESTÁ SE CONCENTRANDO

Sete players — Amazon, Anthropic, Google, IBM, Meta, Microsoft e OpenAI — concentravam aproximadamente 56% do mercado de plataformas de agentes de IA em 2024. Microsoft liderava com cerca de 39% de market share; Google detinha 15% (IoT Analytics / MarketsandMarkets, 2024). A camada de software — onde vivem os agentes — já responde por mais de 77% da receita do mercado (MarketsandMarkets, 2025).

Concentração de mercado não é problema por definição. É problema quando cria assimetria de poder entre fornecedor e cliente — e o histórico de mercados de plataforma mostra que essa assimetria se manifesta em termos contratuais piores, mudanças unilaterais de precificação, e eventualmente competição direta. O mercado já percebe o risco: 67% das organizações estão ativamente trabalhando para evitar dependência de um único fornecedor de IA (Swfte AI / The Register, 2026).

O BRASIL ESTÁ ACELERANDO SEM ESTRATÉGIA DE PLATAFORMA

O crescimento da adoção de IA pelas empresas brasileiras é real e acelerado. O percentual de empresas industriais usando IA foi de 16,9% em 2022 para 41,9% em 2024 — crescimento de 148% em dois anos (IBGE / Pintec Semestral, 2024). Setenta e oito por cento das empresas nacionais planejavam ampliar investimentos em IA até o final de 2025, com gastos esperados acima de USD 2,4 bilhões — crescimento de 30% sobre 2024 (PwC Brasil, 2025).

Porém, apenas 13% das empresas brasileiras declaravam usar aplicações de IA em operações em 2024, com concentração em grandes empresas (38%) e no setor de TIC (CETIC.br / TIC Empresas 2024, 2025). Entre PMEs, o percentual é significativamente menor, com predominância de "soluções de prateleira" — exatamente as que criam maior dependência de fornecedor e menor controle sobre inteligência operacional. O resultado é um mercado que está acelerando a adoção sem estrutura de governança de plataforma: compras feitas por urgência, não por estratégia.

O CUSTO DE MIGRAR É SUBESTIMADO

O custo médio de migração entre plataformas de tecnologia empresarial é de USD 315.000 por projeto, e 57% dos líderes de TI reportam ter gasto mais de USD 1 milhão em migrações no último ano (The Register / CloudBees, 2025-2026). Setenta e cinco por cento dos projetos de migração de ERP ficam fora do prazo ou do orçamento; dois terços têm ROI negativo (McKinsey, citado em Curiosity Software, 2024). Apenas 42% das organizações que tentaram migrar entre plataformas de IA reportaram que o processo ocorreu sem problemas — 58% disseram que o processo falhou ou demandou esforço muito maior do que o esperado (Zapizer, 2025).

Há um agravante específico para agentes de IA que não existe em migrações de ERP. Kai Waehner, arquiteto enterprise da Confluent, chama de "behavioral lock-in" o conjunto de comportamentos, padrões de decisão, vocabulário especializado e workflows de exceção que o agente aprende ao longo do tempo e que não são exportáveis como dado (MindStudio / Kai Waehner, 2026). Em setores B2B com alta densidade de especialização — jurídico, financeiro, contábil, consultoria — esse lock-in comportamental é mais valioso do que qualquer integração técnica, e invisível nos benchmarks de portabilidade que os fornecedores publicam.

OS FORNECEDORES ESTÃO ENTRANDO NO MERCADO DE SERVIÇOS

Em outubro de 2025, a OpenAI passou a exigir investimento mínimo de USD 10 milhões para acesso à sua própria divisão de consultoria — competindo diretamente com as empresas de implementação que até então eram seus parceiros de go-to-market (Marketing AI Institute, 2025). Em fevereiro de 2026, lançou o "Frontier", plataforma de agentes para empresas posicionada explicitamente como "AI coworkers" — o produto que centenas de startups B2B estavam vendendo usando o ChatGPT como back-end.

Em maio de 2026, a Anthropic foi explícita: sua nova joint venture de serviços "competirá com as maiores consultorias do mundo" (Fortune / CIO.com, maio 2026). A Anthropic também fechou acordos com a Deloitte e a Accenture, com 30 mil profissionais treinados em Claude para competir no mesmo mercado de transformação digital corporativa que empresas brasileiras de serviços B2B entregam para seus clientes.

O Gartner prevê que 40% das aplicações empresariais estarão integradas com agentes especializados até o final de 2026, contra menos de 5% em 2025 (Gartner, 2025). Esse crescimento está criando um mercado de implementação que os próprios fornecedores de modelos decidiram capturar.

05. Perspectivas e Casos

O PADRÃO QUE SE REPETE: FORNECEDOR VIRA CONCORRENTE

A dinâmica tem nome acadêmico: *platform envelopment* — a estratégia pela qual uma plataforma expande suas capacidades para absorver mercados adjacentes onde seus parceiros operam. O economista Matteo Motta formalizou o conceito no *Journal of Economics & Management Strategy* em 2025, documentando o mecanismo como "kill zone": o espaço de mercado adjacente ao fornecedor onde novos entrantes — e parceiros existentes — são sistematicamente absorvidos ou eliminados.

Os precedentes são documentados e replicados.

A Amazon usou sistematicamente dados de vendas de sellers terceiros para identificar produtos de sucesso e lançar versões próprias (Amazon Basics), com preferência algorítmica nos resultados de busca — prática investigada por autoridades antitruste nos EUA e Europa (Fuqua School of Business,

Duke, 2024). O seller que construiu seu negócio sobre a plataforma da Amazon financiou, sem saber, a inteligência de mercado usada contra ele.

A Apple cobrou comissões de até 30% de desenvolvedores que construíram negócios dentro de seu ecossistema e lançou funcionalidades nativas que tornaram obsoletos apps de terceiros — prática que ficou conhecida como "Sherlocking". O UK Competition Appeal Tribunal determinou que a Apple abusou de sua posição dominante entre 2015 e 2020. A Comissão Europeia multou a Apple em €1,8 bilhão por práticas anticompetitivas (TechCrunch / MacRumors, 2023-2025).

A Salesforce adquiriu a Vlocity — que havia construído aplicações verticais de sucesso sobre o Salesforce CRM — por USD 1,33 bilhão em 2020, rebatizando-a Salesforce Industries. Comprou também a SteelBrick (CPQ) e a MuleSoft (integração). A lógica é consistente: identificar onde os parceiros criam mais valor e absorver.

A Microsoft executou o que ficou documentado como "embrace, extend, extinguish" — adotar padrões abertos, adicionar extensões proprietárias, tornar concorrentes incompatíveis. O mesmo movimento agora ocorre em IA: integração de GPT-4 em toda a stack Microsoft absorveu verticalmente casos de uso que startups haviam estabelecido como mercados independentes.

O QUE ACONTECE COM EMPRESAS QUE APOSTAM NO PARCEIRO ERRADO

A Inflection AI levantou mais de USD 1,5 bilhão para construir modelos fundacionais e aplicações enterprise. Em março de 2024, a Microsoft pagou USD 650 milhões para licenciar sua tecnologia e contratar praticamente toda a equipe, incluindo o fundador Mustafa Suleyman. Para os clientes enterprise que haviam apostado na Inflection como fornecedora estratégica independente, a posição de negociação desapareceu. Caso similar ocorreu com a Adept AI, adquirida pela Amazon por USD 430 milhões em julho de 2024 (ProMarket, abril de 2025). Em 2025, o mercado registrou 782 aquisições de empresas de IA — mais de 1,5 vez o volume de 2024.

O QUE ACONTECE COM EMPRESAS QUE CONSTROEM CAMADA PROPRIETÁRIA

A Cresta — startup de IA para vendas e suporte B2B — construiu sua plataforma sobre modelos de fundação de terceiros, mas investiu na camada de dados proprietários. Cada interação retroalimenta modelos da empresa, criando um loop de aprendizado que os concorrentes não replicam. Em abril de 2026, atingiu USD 100 milhões de ARR com backing de a16z e Sequoia (Axios, abril de 2026). O diferencial não era o modelo base — era a inteligência operacional acumulada sobre ele.

Empresas brasileiras de serviços B2B que integraram agentes de IA ao CRM e à base de dados histórica reportaram até 35% de aumento em conversões e redução de 50% no CAC. Os resultados sustentáveis vieram exclusivamente das empresas que integraram os agentes a dados proprietários — não das que instalaram ferramentas prontas sem personalização. A diferença entre as duas abordagens é precisamente onde a inteligência operacional fica armazenada (Nautis Tech, 2026).

O QUE DIZEM QUEM PENSA ESTRATÉGIA DE PLATAFORMA

O BCG, em "Rethinking Vendor Strategy: Age of AI Acceleration" (novembro de 2025), elenca quatro critérios para decisões de vendor em IA: **modularidade** (capacidade de trocar modelos sem reescrever workflows), **vantagem de dados** (prefira vendors que ajudam a construir moat proprietário, não os que absorvem dados na plataforma), **transformação de workflow** (não apenas automação de processos existentes) e **tempo até impacto** (resultados em trimestres, não anos fiscais).

Simon Wardley, criador da metodologia Wardley Maps, alerta que organizações que não mapeiam a trajetória evolutiva de tecnologias de IA estão inadvertidamente se posicionando para depender de plataformas fechadas — e perdem o controle sobre seus próprios roadmaps operacionais (Swarm Blog, 2025).

Benedict Evans, no relatório anual "AI Eats the World" (dezembro de 2025), argumenta que os modelos fundacionais estão se tornando commodity — o que significa que a vantagem competitiva migra para as camadas acima e abaixo do modelo: dados proprietários, integração com workflows, e experiência acumulada de uso. Para empresas de serviços: quem apostar apenas no modelo aposta na commodity; quem construir inteligência operacional proprietária sobre o modelo tem um ativo defensável.

Ethan Mollick (Wharton, 2025) observa que a maioria das grandes empresas ficou em território incremental — usando IA para automatizar tarefas legadas em vez de reimaginar workflows. A questão central não é qual modelo usar, mas se a organização desenvolveu a capacidade de experimentar, aprender e escalar continuamente — o que só é possível com arquitetura de plataforma que não crie lock-in irremovível.

06. Contra-Argumentos

Um white paper que não considera suas objeções mais fortes não é análise — é panfleto. Quatro objeções merecem resposta direta.

"A TECNOLOGIA MUDA RÁPIDO DEMAIS PARA APOSTAS DE LONGO PRAZO"

O argumento: ciclos de obsolescência em IA são de 12 a 18 meses, não de dois a três anos como em ERPs. Quem analisou estrategicamente GPT-4 em 2023 viu o cenário mudar com Claude 3, Gemini 1.5 e LLaMA 3 em menos de 18 meses. McKinsey chama de "pilot purgatory" o estado em que empresas ficam presas na análise sem chegar à produção. A Antler documenta que 93% das empresas que aceleraram implementação reportaram ganhos de execução, com quase metade relatando aumento de velocidade de até 5x.

A objeção é legítima — mas endereça o objeto errado. A camada de modelos (qual LLM usar) muda rápido e pode ser trocada com uma troca de credencial de API. A camada de orquestração e dados — onde vivem os workflows, os dados de feedback acumulados, os processos operacionais incorporados

nos agentes — cria dependências estruturais que não se resolvem com velocidade de implementação.

Qualificação da tese: manter modularidade na camada de modelos (facilidade de trocar GPT-4 por Claude) enquanto se faz escolha cuidadosa na camada de orquestração e dados. A "aposta estratégica" é sobre a segunda camada, não a primeira.

"NÃO EXISTE LOCK-IN REAL — DADOS E WORKFLOWS PODEM SER MIGRADOS"

O ecossistema de agentes de IA converge para padrões abertos. A Agentic AI Foundation (AAIF), lançada em dezembro de 2025 com Anthropic, OpenAI e Block como membros fundadores, tem como missão garantir interoperabilidade. Frameworks como LangChain e CrewAI são agnósticos a modelos. Mais de 75% das organizações já operam com múltiplos modelos em produção simultaneamente.

Portabilidade técnica existe. Lock-in organizacional também existe — e é mais difícil de quantificar. Migrar credenciais é trivial. Migrar o conhecimento operacional codificado em prompts, os dados de feedback acumulados, os fluxos de aprovação internos, e a familiaridade da equipe com a plataforma tem custo de tempo e reorganização que não aparece em nenhum benchmark técnico. Em setores B2B com alta especialização, esse custo organizacional é a maior barreira real de migração — e é o que torna 58% das tentativas de migração problemáticas (Zapizer, 2025).

Qualificação da tese: o risco não é "lock-in de API". É "custo de migração organizacional" — mais difícil de quantificar, mais fácil de subestimar, e mais lento de desfazer.

"PARA PMES BRASILEIRAS, A PRIORIDADE É IMPLEMENTAR RÁPIDO, NÃO PROTEGER ESTRATÉGIAS"

Esta é a objeção mais honesta para o público deste documento. Apenas 25% das empresas brasileiras estavam plenamente preparadas para implementar IA em 2025, com lacunas críticas em governança de dados e infraestrutura (INSEAD Knowledge). A Bain & Company aponta que 67% das empresas brasileiras consideram IA prioridade estratégica — mas apenas 12% tinham projetos ativos em 2024. O gargalo não é análise — é execução.

Para o CEO de uma empresa de 80 funcionários sem time de TI próprio, a escolha entre "análise estratégica de plataforma" e "colocar um agente de atendimento funcionando esta semana" tem uma resposta racional: implementar e aprender.

Qualificação da tese: este white paper é para quem já passou do limiar de implementação — ou está prestes a ultrapassá-lo. Antes desse limiar, a orientação correta é: implemente rápido, documente suas dependências, e decida sobre plataforma quando tiver dados reais de uso. A análise estratégica torna-se necessária no momento em que o primeiro agente vai para produção com impacto mensurável.

"FORNECEDORES DE MODELOS VIRAREM CONCORRENTES É IMPROVÁVEL NO CURTO PRAZO"

O argumento: o modelo de negócio de OpenAI e Anthropic é vender acesso via API — receita recorrente de alta margem que não depende de entrar no mercado de serviços. Cada cliente de consultoria que migra para o concorrente representa receita de API perdida. Ambos os CEOs repetiram

publicamente, até recentemente, que seu modelo é ser "infraestrutura" — não prestadores de serviço. Era uma objeção razoável em 2024. Em maio de 2026, é factualmente incorreta. Anthropic anunciou joint venture de serviços enterprise avaliada em USD 1,5 bilhão, com Blackstone, Goldman Sachs e Hellman & Friedman. OpenAI lançou "The Development Company" com USD 4 bilhões captados, competindo diretamente com clientes que usam sua API para entregar esses mesmos serviços. O Brookings Institution publicou análise específica sobre o que acontece quando empresas de IA passam a competir com seus clientes — e a conclusão não é tranquilizadora.

A objeção identificou corretamente o incentivo de longo prazo dos fornecedores (escalar API). O que não previu foi a pressão de curto prazo de valuation, expectativas de investidores e oportunidade de mercado que tornaram a expansão para serviços irresistível.

07. Análise Estratégica

Três frameworks estruturam a análise desta tese. Cada um ilumina uma dimensão diferente do problema.

PORTER'S FIVE FORCES — O PODER DO FORNECEDOR ESTÁ MUDANDO DE NATUREZA

O modelo de Porter identifica cinco forças competitivas. Para empresas de serviços B2B que adotam plataformas de agentes de IA, a força mais relevante em 2026 é o poder de barganha dos fornecedores — e sua extensão mais crítica: a ameaça de integração vertical.

Poder dos fornecedores: Sete players concentram 56% do mercado. Microsoft (39%) e Google (15%) dominam a camada de plataforma. Essa concentração cria assimetria de poder: quando há poucos fornecedores de uma tecnologia crítica, o cliente perde capacidade de negociar termos, preços e continuidade de funcionalidades.

Integração vertical — o movimento de maio de 2026: A teoria de Porter prevê que fornecedores poderosos invadem mercados de seus clientes quando identificam que a margem ali é superior à margem na posição atual. OpenAI e Anthropic fizeram exatamente isso. A pergunta para o CEO não é "isso vai acontecer?" — já aconteceu. A pergunta é: quanto da inteligência operacional da empresa está dentro de sistemas que esses fornecedores controlam?

Ameaça de novos entrantes: A concentração em hyperscalers cria barreiras que consolidam o poder dos atuais líderes. Startups de plataformas independentes são suscetíveis de aquisição — e com elas vão os clientes que haviam apostado em independência (caso Inflection AI, caso Adept AI).

A análise de Porter aponta para uma conclusão direta: a posição de poder do comprador está se deteriorando à medida que o mercado se concentra e os fornecedores sobem na cadeia de valor. Reverter essa posição exige construir ativos que reduzam a dependência: dados proprietários, arquitetura modular, e inteligência operacional que não vive dentro da plataforma do fornecedor.

TECHNOLOGY ADOPTION LIFECYCLE (MOORE) — A JANELA DE VANTAGEM ESTÁ FECHANDO

Geoffrey Moore identificou que tecnologias emergentes percorrem fases distintas de adoção, com um "abismo" entre adotantes iniciais e a maioria. Plataformas de agentes de IA, em 2026, estão cruzando esse abismo no mercado global — e no Brasil, estão na fronteira entre innovators e early adopters.

Onde estamos no ciclo: O Gartner prevê que 40% das aplicações empresariais estarão integradas com agentes especializados até o final de 2026, contra menos de 5% em 2025. Esse salto é a definição de crossing the chasm — a transição da adoção por entusiastas para adoção pela maioria.

O que isso significa para a decisão de plataforma: No momento em que uma tecnologia cruza o abismo, os padrões de mercado se solidificam. Os defaults de 2026 — qual plataforma é "a padrão" para determinado setor, qual integração é "a que todo mundo usa" — serão muito mais difíceis de desafiar em 2028. Empresas que estabelecem posição agora, com arquitetura bem estruturada, acumulam dados operacionais que os entrantes tardios não conseguem replicar rapidamente.

O custo de esperar: Empresas que adotam tarde em mercados de plataforma não pagam apenas o custo de oportunidade. Pagam o custo de entrar em um mercado onde os concorrentes já acumularam dois anos de dados operacionais, e onde a margem de negociação com o fornecedor é menor porque a dependência do ecossistema já está estabelecida.

Para o CEO brasileiro: em 2026, estar na vanguarda não significa ser pioneiro imprudente. Significa não ser o último a decidir quando as opções ainda existem.

VRIO (BARNEY) — O QUE DEFINE VANTAGEM COMPETITIVA SUSTENTÁVEL EM IA

O framework VRIO identifica quatro critérios para que um recurso seja fonte de vantagem competitiva sustentável: Valor, Raridade, Imitabilidade difícil, e Organização capaz de capturar o valor.

Valor (V): Inteligência operacional — comportamentos, padrões de decisão, e conhecimento acumulado pelo sistema de agentes ao longo do tempo — tem valor direto em eficiência e qualidade de serviço. Esse valor existe. A questão é: quem captura?

Raridade (R): Modelos de linguagem disponíveis via API são acessíveis a qualquer empresa com cartão de crédito. Inteligência operacional construída sobre dados proprietários de uma empresa específica — seu histórico de clientes, seus padrões de exceção, seu vocabulário especializado — é rara por definição. Não existe em nenhum outro lugar do mercado.

Imitabilidade (I): Um concorrente pode assinar a mesma plataforma de agentes que você usa. Não pode replicar os dois anos de dados de feedback que seus agentes acumularam. O modelo base é imitável; a camada de dados proprietários construída sobre ele não é — desde que essa camada fique sob controle da empresa.

Organização (O): Aqui está o ponto crítico. Se a inteligência operacional fica armazenada dentro da plataforma do fornecedor — nos logs, nos embeddings, nas conexões entre sistemas que só o fornecedor controla — a empresa não consegue capturar o valor que ajudou a criar. A vantagem existe, mas pertence ao fornecedor.

A conclusão do VRIO: A vantagem competitiva sustentável em IA não vem do modelo escolhido. Vem de construir inteligência operacional proprietária — dados, fluxos, comportamentos aprendidos — em uma arquitetura onde a empresa mantém controle e portabilidade. Plataformas que tornam isso impossível destroem valor estratégico mesmo quando entregam valor operacional de curto prazo.

07b. Panorama de Plataformas: As Principais Opções e Onde Cada Uma te Posiciona

A análise estratégica dos frameworks anteriores é abstrata sem um mapa concreto das opções disponíveis. O mercado de plataformas de agentes de IA em 2026 segmenta-se em quatro perfis distintos — cada um com implicações diferentes para governança, custo, controle e risco de lock-in.

PERFIL 1 — FRAMEWORKS DE CÓDIGO ABERTO PARA DESENVOLVEDORES

LangChain / LangGraph (Python/JS, open-source)

O ecossistema LangChain é o mais adotado entre desenvolvedores para construção de pipelines agênticos. LangChain organiza chamadas a modelos, ferramentas e memória em cadeias (chains); LangGraph adiciona orquestração de múltiplos agentes com ciclos e estado persistente. Exige desenvolvedor Python com pelo menos conhecimento intermediário. O framework em si é gratuito, mas o custo real vem da infraestrutura e do custo de desenvolvimento. O lock-in não é do fornecedor — é do próprio framework: migrar um pipeline LangChain para outra arquitetura exige reescrita significativa. Vantagem crítica: os dados ficam onde você escolhe hospedar; a inteligência operacional é totalmente proprietária.

CrewAI (Python, open-source)

Especializado em sistemas multi-agente com papéis definidos (crew de agentes com funções específicas). Mais opinativo que LangChain — impõe uma arquitetura de "agente com papel + objetivo + ferramentas". Curva de aprendizado menor para equipes que querem multi-agente sem implementar toda a orquestração do zero. Comunidade crescente; menos maduro que LangChain para casos de produção complexos. Dados 100% proprietários.

PERFIL 2 — PLATAFORMAS LOW-CODE SELF-HOSTED

Flowise (open-source, self-hostable)

Interface visual de drag-and-drop sobre LangChain — permite construir pipelines agênticos sem escrever código Python, usando nós conectáveis (LLM, memória, ferramentas, documentos). Ideal para times com um desenvolvedor que conhece LangChain mas quer velocidade de construção. Self-

hosted: dados ficam no servidor da empresa. Gratuito para self-host; versão cloud disponível. Comunidade ativa, updates frequentes. Lock-in baixo — os fluxos podem ser exportados como JSON e reimplementados. Limitação: não foi projetado para workflows com lógica de negócio muito complexa ou loops de aprovação humana sofisticados.

n8n com nós de IA (open-source, self-hostable)

Plataforma de automação de workflows que evoluiu para suportar agentes de IA como parte de fluxos maiores. Ideal quando o agente de IA precisa se integrar a CRM, ERP, e-mail, webhooks e outros sistemas — n8n já tem centenas de integrações nativas. A lógica de agente fica embutida em um workflow maior. Self-hosted com Docker, dados na empresa. Lock-in médio — a lógica de workflow fica presa no formato n8n, mas os dados são livres. Preço: gratuito (self-host) ou \$20-\$50/mês (cloud). Para empresas de serviços B2B com muitas integrações operacionais, n8n é frequentemente a opção mais pragmática.

PERFIL 3 — PLATAFORMAS NO-CODE / LOW-CODE EM CLOUD

Dify (open-source core, SaaS e self-hosted)

Plataforma de LLM ops + construtor de workflows agenticos com UI amigável. Suporta prompt engineering visual, RAG (Retrieval-Augmented Generation) com bases de conhecimento nativas, e orquestração de múltiplos modelos. Pode ser self-hosted (Docker) ou usado como SaaS (\$159/mês no plano profissional). Tem forte adoção no mercado asiático e crescimento rápido na Europa/LATAM. Boa opção para times que querem controle sem precisar de desenvolvedor full-time. Lock-in baixo-médio: dados e configurações exportáveis no formato self-hosted; dependência de plataforma no SaaS.

Stack AI (SaaS, enterprise)

Plataforma no-code enterprise para construção de workflows agenticos. Interface visual com suporte a múltiplos modelos (GPT-4, Claude, Llama), RAG nativo, conectores para Salesforce, HubSpot, Google Drive, Slack, e APIs personalizadas. Foco em enterprise com governança, controles de acesso e auditoria. Preço a partir de \$499/mês — posicionado para empresas médias e grandes que precisam de agilidade sem time de engenharia. Dados ficam na plataforma Stack AI (cloud), com opções de compliance (SOC2, HIPAA). Lock-in alto: workflows visuais não são portáteis para outras plataformas. Para a PME brasileira, o preço é o principal obstáculo.

Relevance AI (SaaS, no-code)

Plataforma australiana focada em "AI workers" — agentes configuráveis via UI que executam tarefas repetitivas (pesquisa, redação, qualificação de leads, suporte). Interface mais simples que Stack AI, com templates prontos. Preço: \$19-\$599/mês (planos flexíveis). Boa opção para times não-técnicos que querem implementar rapidamente. Lock-in alto: dados e agentes ficam no SaaS Relevance AI; exportação limitada. Risco de sobreposição de mercado médio — a Relevance AI não é um fornecedor de modelos, mas depende de OpenAI/Anthropic como modelos base.

PERFIL 4 — PLATAFORMAS ENTERPRISE DE HYPERSCALERS

Microsoft Copilot Studio (SaaS, enterprise)

Plataforma no-code da Microsoft para construção de agentes integrados ao ecossistema Microsoft 365 (Teams, SharePoint, Dynamics). Ideal para empresas que já têm Azure Active Directory, Dynamics CRM ou Microsoft 365 como infraestrutura central. Lock-in máximo: os agentes só funcionam dentro do ecossistema Microsoft; dados ficam no Azure. Preço: ~\$200/mês + consumo. Para empresas que já estão no ecossistema Microsoft e não pretendem sair, é a opção de menor fricção de integração — e a de maior dependência estratégica.

AWS Bedrock Agents (PaaS, enterprise)

Plataforma da Amazon para construção de agentes sobre múltiplos modelos fundacionais (Claude, Llama, Titan) via API, com gerenciamento de memória, RAG nativo (Knowledge Bases) e orquestração. Exige desenvolvimento (Python/SDK), mas oferece mais controle que Copilot Studio. Dados ficam no AWS (S3, RDS, OpenSearch). Lock-in alto — não pelo agente em si, mas pela integração profunda com serviços AWS (Lambda, Bedrock, S3). Para empresas que já têm infraestrutura AWS consolidada, é uma extensão natural. Para quem não tem, o custo de setup é alto.

COMO MAPEAR ESSAS OPÇÕES PELA LENTE ESTRATÉGICA

A tabela abaixo organiza as principais plataformas pelos quatro critérios estratégicos definidos neste white paper:

Plataforma	Governança dos dados	Nível técnico	Custo mensal base	Lock-in estratégico
LangChain/LangGraph	Proprietária (self-hosted)	Pro-code	Gratuito + infra	Baixo
CrewAI	Proprietária (self-hosted)	Pro-code	Gratuito + infra	Baixo
Flowise	Proprietária (self-hosted)	Low-code	Gratuito + infra	Baixo
n8n	Proprietária (self-hosted)	Low-code	Gratuito-\$50	Baixo-médio
Dify	Proprietária (self-hosted) ou fornecedor	Low-code	Gratuito-\$159	Baixo-médio
Relevance AI	Fornecedor (SaaS)	No-code	\$19-\$599	Alto
Stack AI	Fornecedor (SaaS)	No-code	\$499+	Alto
Copilot Studio	Microsoft (Azure)	No-code	\$200+ consumo	Muito alto
AWS Bedrock Agents	Amazon (AWS)	Pro-code/API	Pay-as-you-go	Muito alto

Leitura estratégica da tabela: Há uma correlação quase direta entre facilidade de uso (no-code) e lock-in estratégico. As plataformas mais acessíveis para times não-técnicos são exatamente as que criam maior dependência de fornecedor e menor portabilidade de inteligência operacional. Isso não significa que plataformas no-code são erradas — significa que o CEO precisa entender o custo estratégico embutido na conveniência.

Para a PME brasileira de serviços B2B com 50-500 funcionários, o ponto de entrada mais equilibrado entre controle e viabilidade é Flowise ou Dify em self-hosted — ambas permitem times pequenos com um desenvolvedor júnior ou analista técnico construírem agentes funcionais, mantendo os dados e a inteligência operacional sob controle da empresa. n8n é o complemento natural quando o agente precisa se integrar ao ecossistema operacional existente.

Stack AI e Relevance AI fazem sentido quando o time não tem capacidade técnica alguma e precisa de resultado em 30 dias — com consciência explícita de que o custo de saída aumenta a cada mês de uso. Copilot Studio e AWS Bedrock Agents só se justificam quando a empresa já está profundamente integrada ao ecossistema do hyperscaler e essa integração é estrategicamente intencional, não acidental.

08. Implicações para o Leitor

Para o CEO ou founder de uma empresa de serviços B2B com 50 a 500 funcionários no Brasil, esta análise tem três implicações concretas.

Você provavelmente já está tomando essa decisão, mesmo que não saiba.

Toda vez que uma área da empresa contrata uma ferramenta de IA — um agente de atendimento, um copilot de vendas, um gerador de propostas — está sendo feita uma escolha de plataforma. Essas escolhas fragmentadas acumulam dependências que, somadas, criam o mesmo lock-in que uma decisão formal de plataforma criaria. A diferença é que decisões fragmentadas são feitas sem critério estratégico e sem visibilidade do risco acumulado. A primeira implicação é mapear o que já existe: quais fornecedores de IA estão ativos hoje, onde estão os dados que esses sistemas processam, e quem controla acesso e exportação.

O critério de seleção de plataforma está errado na maioria das avaliações.

Empresas brasileiras tendem a avaliar plataformas de IA por preço, facilidade de implementação, e qualidade das demonstrações de produto. Esses são critérios válidos para a fase de piloto. Não são os critérios corretos para a decisão de escala. Os critérios estratégicos — que determinam onde a empresa vai estar em dois anos — são: (1) portabilidade de dados, (2) modularidade de modelo, (3) sobreposição de mercado entre o fornecedor e os clientes da empresa, e (4) custo de migração quantificado. Fornecedores que se recusam a responder questões sobre portabilidade de dados estão comunicando a resposta, mesmo que não verbalmente.

A inteligência operacional precisa ser tratada como ativo estratégico, não como subproduto.

A maioria das empresas B2B trata os dados gerados por seus agentes de IA como subproduto operacional — logs a serem deletados, históricos a serem arquivados. As empresas que vão construir vantagem competitiva real tratam esses dados como o ativo mais valioso que estão produzindo. Isso requer decisões arquiteturais deliberadas: onde os dados ficam armazenados, quem tem acesso, como são estruturados para retroalimentar os sistemas, e como garantir que esse ativo pertence à empresa — não ao fornecedor.

09. Recomendações

RECOMENDAÇÃO 1 — MAPEAR DEPENDÊNCIAS DE IA EXISTENTES NA EMPRESA

Ação: Inventariar todos os fornecedores de IA ativos, identificando: nome do fornecedor, tipo de dados processados, localização dos dados (servidor do fornecedor vs. servidor próprio), condições contratuais de exportação, e custo estimado de migração.

Prazo: 90 dias.

Por quê: Você não pode gerenciar o risco de uma dependência que não sabe que existe. Empresas que fizeram esse mapa antes de escalar evitam a situação — documentada em 58% das migrações — em que o custo de sair é maior do que o custo de continuar em uma relação insatisfatória (Zapizer, 2025).

RECOMENDAÇÃO 2 — INCLUIR QUATRO CRITÉRIOS ESTRATÉGICOS EM TODA AVALIAÇÃO DE PLATAFORMA

Ação: Antes de qualquer novo contrato de plataforma de agentes de IA, conduzir avaliação formal nos quatro critérios: (1) portabilidade de dados — é possível exportar, em formato aberto, todos os dados processados pelos agentes? (2) modularidade de modelo — a plataforma permite trocar o modelo de linguagem subjacente sem reescrever a lógica de negócio? (3) sobreposição de mercado — o fornecedor está ou poderia estar prestando serviços para os clientes que minha empresa atende? (4) custo de migração quantificado — se precisar sair desta plataforma em 18 meses, qual é o custo real?

Prazo: Imediato — aplicável ao próximo contrato a ser renovado ou firmado.

Por quê: Fornecedores que se recusam a responder questões de portabilidade de dados ou que não têm SLA de exportação documentado estão, implicitamente, comunicando a resposta. A avaliação formal revela posições que não aparecem em demonstrações de produto.

RECOMENDAÇÃO 3 — DEFINIR ONDE VAI FICAR A INTELIGÊNCIA OPERACIONAL PROPRIETÁRIA ATÉ Q3/2026

Ação: Decidir explicitamente — como decisão de arquitetura, não como consequência de contratações fragmentadas — onde os dados operacionais gerados pelos agentes vão ser armazenados, estruturados e governados. Isso inclui: definir um repositório sob controle da empresa, estabelecer política de retenção e exportação, e nomear um responsável por essa governança.

Prazo: Q3/2026.

Por quê: O framework VRIO mostra que a vantagem competitiva sustentável vem de ativos raros, difíceis de imitar, e organizados para capturar valor. Dados operacionais de IA preenchem os três primeiros critérios — mas só geram vantagem se a empresa consegue capturar esse valor, o que exige governança explícita.

RECOMENDAÇÃO 4 — EXIGIR CLÁUSULA DE SOBREPOSIÇÃO DE MERCADO EM CONTRATOS COM FORNECEDORES DE IA

Ação: Incluir, em novos contratos com fornecedores de plataformas de agentes de IA, cláusula que estabeleça obrigação de notificação prévia (mínimo 90 dias) caso o fornecedor decida entrar em mercados onde a empresa opera, ou adquirir empresas que atuam nesses mercados.

Prazo: Imediato — para próximos contratos; renegociar nos contratos existentes até Q4/2026.

Por quê: Os eventos de maio de 2026 mostram que a transição de fornecedor para concorrente pode ocorrer em semanas. Notificação prévia não elimina o risco, mas dá tempo para ajuste estratégico antes que o lock-in operacional torne a migração proibitiva.

RECOMENDAÇÃO 5 — ESTABELECECR CRITÉRIO DE MATURIDADE ANTES DE ESCALAR

Ação: Definir internamente um critério de "escala autorizada" para qualquer solução de agentes de IA: a solução só recebe mais investimento e integração profunda após demonstrar ROI mensurável em piloto de 60 dias, com dados de uso exportáveis para análise interna.

Prazo: A partir do próximo piloto de IA.

Por quê: A Bain & Company (2025) aponta que 67% das empresas brasileiras consideram IA prioridade estratégica, mas apenas 12% tinham projetos ativos em 2024. O gargalo não é análise — é execução. O critério de escala autorizada resolve dois problemas simultâneos: força a implementação rápida (piloto de 60 dias) e impede o scale-up sem evidência de valor (ROI mensurável com dados exportáveis).

10. Conclusão

Em maio de 2026, a decisão sobre qual plataforma de agentes de IA usar deixou de ser uma escolha técnica com consequências revertíveis. Tornou-se uma aposta sobre com quem você vai competir, onde vai estar a inteligência operacional que sua equipe acumulará nos próximos dois anos, e qual será o custo de mudar de ideia.

Os fornecedores de modelos não são vilões. São empresas racionais respondendo a pressões de mercado, de investidores, e de oportunidades que nenhum conselho de administração competente ignoraria. O problema não está na intenção dos fornecedores — está na estrutura de incentivos que torna a integração vertical um movimento previsível, e na desatenção com que a maioria das empresas de serviços B2B está construindo dependências sobre essas estruturas.

A proteção não exige infraestrutura sofisticada nem equipe de dezenas de engenheiros. Exige clareza sobre três perguntas que cada CEO pode responder agora:

Onde fica a inteligência operacional que minha empresa está acumulando?

Se meu principal fornecedor de IA decidir competir com meus clientes, o que acontece?

Qual é o custo real de sair desta plataforma em 18 meses?

Empresas que respondem essas perguntas antes de assinar contratos são as que, em 2028, terão construído ativos proprietários defensáveis. As que não respondem estarão, naquele momento, financiando a vantagem competitiva dos fornecedores que escolheram — às suas próprias custas.

A janela para fazer as perguntas certas — e ainda ter opções para agir — é agora.

11. Notas Metodológicas

Escopo e período da pesquisa

Este white paper cobre o período de 2024 a 2026, com ênfase no mercado brasileiro de serviços B2B e contextualização com dados globais. A pesquisa foi conduzida em maio de 2026 por três agentes especializados em paralelo: (A) evidências quantitativas, (B) perspectivas e casos, (C) contra-argumentos e nuances.

Frameworks aplicados

- **Porter's Five Forces** (Porter, 1979): análise do poder de fornecedores e ameaça de integração vertical no mercado de plataformas de agentes de IA.
- **Technology Adoption Lifecycle** (Moore, 1991 / "Crossing the Chasm"): posicionamento das plataformas de agentes de IA no ciclo de adoção e implicações para o timing de decisão.
- **VRIO** (Barney, 1991): identificação de recursos que geram vantagem competitiva sustentável — aplicado à inteligência operacional como ativo estratégico.

Fontes principais

Dados de mercado: Fortune Business Insights / GM Insights (2025), MarketsandMarkets (2025), IoT Analytics (2024), Gartner (2025), IDC (2025), TI Inside (2026).

Dados Brasil: IBGE / Pintec Semestral (2024), CETIC.br / TIC Empresas 2024 (2025), PwC Brasil (2025), INSEAD Knowledge (2025), Bain & Company (2025), Exame (2025), Nautis Tech (2026).

Custo de migração: The Register / CloudBees (2025-2026), Integrate.io (2025), Zapizer (2025), McKinsey Global Institute (via Curiosity Software, 2024), MindStudio / Kai Waehner (2026).

Casos e perspectivas: Axios (Cresta, abril 2026), ProMarket (Inflection AI / Adept AI, 2025), TechCrunch / Fortune (OpenAI e Anthropic joint ventures, maio 2026), SiliconAngle / TechInformed (OpenAI Frontier Alliances, 2026), The Next Web (Google A2A, 2026), Brookings Institution (AI companies vs. customers), BCG "Rethinking Vendor Strategy" (novembro 2025), Benedict Evans "AI Eats the World" (dezembro 2025), Ethan Mollick / Exponential View (2025), Simon Wardley / Swarm Blog (2025), Fuqua School of Business / Duke (Amazon Basics), TechCrunch / MacRumors (Apple App Store, 2023-2025), Journal of Economics & Management Strategy / Wiley (Motta, 2025), ScienceDirect (generative AI platforms, 2025).

Governança e padrões: Agentic AI Foundation (dezembro 2025), SmythOS / CustomGPT (portabilidade LLM), Swfte AI (concentração de mercado), Marketing AI Institute (2025).

Limitações

Dados de mercado de plataformas de agentes de IA (Fortune Business Insights, MarketsandMarkets, GM Insights) são projeções comerciais com metodologias não totalmente transparentes — usados como ordem de magnitude, não como número preciso. Dados de migração de plataforma (CloudBees, Zapizer) são de pesquisas com amostras principalmente norte-americanas; custos brasileiros podem variar. Dados específicos sobre adoção de agentes de IA por PMEs brasileiras são limitados — CETIC.br / IBGE são as fontes mais robustas disponíveis, mas cobrem "IA" de forma ampla, não agentes especificamente. O movimento de fornecedores de IA para serviços (OpenAI, Anthropic, maio 2026) é recente; impactos completos no ecossistema estão em desenvolvimento.

Data de conclusão: Maio de 2026.